MAXIMIZING AND ASSESSING THE ACCURACY OF EQUILIBRIUM DISSOCIATION CONSTANTS FOR AFFINITY COMPLEXES

TONG YE WANG

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PHYSICS & ASTRONOMY
YORK UNIVERSITY
TORONTO, ONTARIO

July 2025

Abstract

Affinity interactions are fundamental to both biophysical technologies and biological systems, with the equilibrium dissociation constant (K_d) serving as a critical parameter that quantifies binding strength. Accurately determining K_d is essential for applications ranging from drug discovery and diagnostics to mechanistic studies in molecular biology. However, despite its importance, K_d values determined by different biophysical methods often vary significantly, highlighting unresolved measurement inaccuracies. Before this work, no comprehensive framework has been established to systematically identify the root causes of these inaccuracies or to provide accessible tools for reliably evaluating K_d accuracy.

This dissertation addresses these challenges through both experimental and theoretical advancements. On the experimental side, I confirmed the robustness and ruggedness of the Accurate Constant *via* Transient Incomplete Separation (ACTIS) method. Initially applicable only to protein–small molecule systems due to its reliance on diffusivity differences, ACTIS was extended to study protein–DNA complexes by optimizing instrumentation and experimental protocols. On the theoretical side, I performed a systematic analysis to identify the key determinants of K_d accuracy that are independent of measurement method. The study revealed that minimizing the concentration of the limiting component and reducing systematic errors in reagent concentrations and signal measurements are essential for achieving accurate K_d values. Recognizing practical constraints such as the limit of quantitation (LOQ), I investigated common sources of systematic error and proposed mitigation strategies.

To provide researchers with a practical tool for assessing K_d accuracy, I developed a computationally efficient algorithm to estimate the Accuracy Confidence Interval (ACI) for K_d

from a single binding isotherm, which serves as the foundation for the web-based tool ACI-Kd. Given the mathematical similarity between K_d and the Michaelis constant (K_m), the framework was further extended to enzyme kinetics, resulting in ACI-Km.

Additionally, I developed ACI-ITC, a Monte Carlo-based tool that evaluates the accuracy of K_d, binding enthalpy ($\Delta H°$), and stoichiometry (n) from isothermal titration calorimetry data. All tools are accessible through a unified web platform: https://aci.sci.yorku.ca. Collectively, this research provides rigorous, accessible methodologies for enhancing the reliability of molecular interaction measurements across a wide range of experimental systems.

Acknowledgments

Above all, I give glory and praise to God for His constant presence and for granting daily strength, wisdom, peace, and eternal hope.

I am profoundly grateful to my supervisor, Dr. Sergey Krylov, for his unwavering support, guidance, and encouragement throughout this journey. At the beginning, when I struggled to find my footing, his patience and belief in my potential gave me the time and space to grow. During moments when my research reached an impasse, his insightful advice, intellectual contributions, and steadfast encouragement consistently brought clarity and renewed motivation. His wisdom, reflected in thoughtful insights and discerning perspectives, continually guided me toward deeper understanding and more effective problem-solving. Dr. Krylov's dedication to rigorous scientific inquiry, unwavering integrity, and high standards has been an ongoing source of inspiration. His thoughtful feedback and committed mentorship have profoundly shaped my academic development and personal growth. I am genuinely thankful for the opportunity to learn under his guidance and for the lasting impact he has had on my journey.

To my committee members, Dr. Ozzy Mermut and Dr. Christopher Bergevin, thank you for your valuable feedback, constructive criticism, and generous investment of time. Your thoughtful contributions have significantly strengthened the quality and clarity of this work. My appreciation also goes to my internal examiner, Dr. Philip Johnson, and my external examiner, Dr. Juewen Liu from the University of Waterloo, for dedicating their time to evaluating my dissertation.

I am also grateful to all my past and current labmates and colleagues for their camaraderie and collaboration. The shared joys, challenges, breakthroughs, and everyday conversations made this journey richer and more meaningful. Special thanks to senior lab members and alumni, Dr.

Svetlana Krylova, Dr. Liang Hu, Dr. An Le Thi Hoai, Dr. Stanislav Beloborodov, Jean Luc Rukundo, Dr. Nikita Ivanov, and Dr. Sven Kochmann for their mentorship, generous help, and invaluable contributions that made this journey possible.

I would also like to extend my sincere gratitude to all my co-authors who contributed to the publications arising from this work. Although I am unable to name each of you individually here, your insights, technical contributions, and collaborative spirit have been invaluable throughout the research process. Your dedication and support have played a crucial role in shaping the scientific quality and impact of this dissertation, and I am truly grateful for the opportunity to have worked alongside you.

To my office mates, Dr. Hongchen Ji and Anirudh Krishnadas, I will always cherish our shared moments — our heartfelt conversations, mutual support, laughter, and even the tears along the way.

To my church family, thank you for your constant prayers, love, and encouragement. Your faith and support reminded me of God's goodness and faithfulness, especially during the most challenging times.

Finally, to my beloved family, your love, sacrifice, and unwavering support have been the foundation of my life and work. To my parents, Wei Dong Wang and Qing Wang, thank you for your endless love, profound patience, and tireless encouragement. Your constant faith in me has carried me farther than words can convey. To my wife, Hong Shen, thank you for journeying alongside me with patience, strength, and boundless love. Your quiet endurance, sacrifices, and unwavering belief in me have been crucial to this accomplishment. To my lovely children, Eddy and Effie, your radiant smiles, joyful laughter, and unconditional love continually remind me of

life’s deepest beauty and meaning. Your presence fills my heart with profound gratitude, endless motivation, and immeasurable joy.

I acknowledge the use of ChatGPT (OpenAI) to enhance the clarity and readability of this dissertation, as well as to support the development of the Python programs presented herein. All scientific analyses, data interpretation, and conclusions were independently conceived, executed, and validated by me.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACE	Affinity capillary electrophoresis
ACI	Accuracy Confidence Interval
ACTIS	Accurate Constant via Transient Incomplete Separation
BLI	Biolayer Interferometry
BSA	Bovine Serum Albumin
BSI	Back-Scattering Interferometry
CE	Capillary Electrophoresis
CE-FA	Capillary Electrophoresis-Frontal Analysis
CI	Confidence Interval
CRMs	Certified Reference Materials
ELISA	Enzyme-Linked Immuno-Sorbent Assays
EM	Equilibrium Mixture
ESI-MS	Electrospray Ionization Mass Spectrometry
FA	Fluorescence Anisotropy
FRET	Fluorescence Resonance Energy Transfer
HMC	Hamiltonian Monte Carlo
HPLC	High-Performance Liquid Chromatography
ISO	International Organization for Standardization

ITC	Isothermal Titration Calorimetry
KCE	Kinetic Capillary Electrophoresis
LIF	Laser-Induced Fluorescence
LOQ	Limit of Quantitation
L-S	Least-Squares
MCE	Microchip Capillary Electrophoresis
MCMC	Markov Chain Monte Carlo
MM	Michaelis-Menten
MS	Mass Spectrometry
MST	Microscale Thermophoresis
MW	Molecular Weight
NAAP	NECEEM Area Analysis Program
NECEEM	Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures
NMR	Nuclear Magnetic Resonance
PCI	Precision Confidence Interval
PEEK	Polyether Ether Ketone
PEG	Polyethylene Glycol
PVA	Polyvinyl Alcohol
R&D	Research and Development

RB	Running Buffer
RFU	Relative Fluorescence Unit
RSD	Relative Standard Deviation
S/N	Signal-to-Noise ratio
SAM	self-assembled monolayer
SB	Sample Buffer
SDS-PAGE	Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis
SOP	Standard Operating Procedure
SPR	Surface Plasmon Resonance
TIS	Transient Incomplete Separation
UV-Vis	Ultraviolet–Visible
WLS	Weighted Least Squares

List of Symbols

σ	standard deviation (uncertainty) of $K_{d,det}$ (or $K_{m,det}$) obtained with nonlinear regression
τ_L	characteristic time of diffusion of ligand
δR	random error of determined fraction of unbound ligand (or enzyme)
[C]	instantaneous concentration of complex in a reversible target–ligand interaction
[L]	instantaneous concentration of ligand in a reversible target–ligand interaction
[T]	instantaneous concentration of target in a reversible target–ligand interaction
μ_C	diffusion coefficient of complex
μ_L	diffusion coefficient of ligand
a	inner radius of a capillary
C	equilibrium concentration of the target–ligand (or enzyme–substrate) complex
C	target–ligand (or enzyme–substrate) complex
D_{app}	apparent diffusion coefficient
E	equilibrium concentration of the enzyme
E_0	total concentration of the enzyme
G	grating factor in fluorescence anisotropy
$I_\perp$	intensity of emitted light with a polarization orientation perpendicular to that of the excitation light
$I_\parallel$	intensity of light emitted by a fluorophore with a polarization orientation parallel to that of the excitation light

I_{max}	intensity the of light polarized in the direction that corresponds to maximum intensity
I_{min}	intensity of light polarized in the direction that corresponds to minimum intensity
K_a	acid dissociation constant for acids
k_{cat}	catalytic rate, *i.e.*, product-formation rate
K_d	both name the equilibrium dissociation constant and the accurate/true value of K_d
$K_{d,det}$	value of K_d determined with nominal ligand and target concentrations
$K_{m,det}$	value of K_m determined with nominal enzyme and substrate concentrations
$K_{d,det}'$	value of K_d determined for the ligand concentration with a large artificial systematic error
$K_{m,det}'$	value of K_m determined with the enzyme concentration with a large artificial systematic error
$K_{d,inp}$	inputted K_d value in simulations
K_L	Langmuir constant of adsorbent–adsorbate interactions
K_m	both name the Michaelis constant and the accurate/true value of K_m
k_{off}	rate constant of complex dissociation
k_{on}	rate constant of complex formation
K_s	stability constant of metal–ligand complexes
L	equilibrium concentration of the ligand
L	ligand
L_0	total concentration of the ligand
$L_{0,fit}$	the value of L_0 used in nonlinear regression (fitting the binding isotherm); assumed to be equal to the nominal value of L_0

$L_{0,\mathrm{fit}}'$ the value of L_0 used in nonlinear regression (fitting the binding isotherm); chosen to be much smaller than the nominal value of L_0, *i.e.*, with a large artificial systematic error

n the stoichiometry of the interaction between the ligand and the target; in ITC, it is often considered as a correction factor for the ligand concentration in the sample cell

p parameter; true/accurate value of parameter

P polarization, or the cumulative polarization measured from the equilibrium mixture

p_{det} parameter determined with nominal values of variables

p_{det}' parameter determined by introducing large artificial systematic errors into the values of variables

Q volumetric flow rate

Q_i total heat absorbed/released after i-th injection in the context of ITC

Q_{rinse} volumetric flow rate used in the rinsing process

Q_{s} volumetric flow rate of the sample-injection process

Q_{TIS} volumetric flow rate for optimal transient incomplete separation (TIS)

R fraction of unbound ligand/enzyme at the equilibrium/pseudo-equilibrium state

r anisotropy, or the cumulative anisotropy determined from the equilibrium mixture

S cumulative signal from the unbound ligand and the target-bound ligand in the equilibrium mixture; equilibrium concentration of the substrate

S_0 total concentration of the substrate

S_{C} signal from the target–ligand complex

S_C^*	signal from pure complex with a concentration of L_0
S_L	signal from the unbound ligand
S_L^*	signal from pure ligand with a concentration of L_0
T	target
T	equilibrium concentration of the target
T_0	total concentration of the target
t_{eq}	characteristic equilibration time under the condition of pseudo-equilibrium in the binding process
t_{inc}	incubation time of a mixture of ligand and target
t_{sep}	characteristic separation time
v_0	initial reaction velocity in the context of Michaelis–Menten kinetics
V_0	ITC sample cell volume
$v_{0,max}$	maximum reaction velocity in the context of Michaelis–Menten kinetics
v_{av}	average flow velocity
V_{inj}	injection volume per step in ITC
v_{max}	maximum flow velocity
x_c	the fixed (constant) independent variable in a parameter-determination method
x_v	the varied independent variable in a parameter-determination method
Γ	the molar ratio between the target and the ligand molecules
$\mathbf{\Lambda}$ and $\mathbf{\Lambda'}$	two different vectors of the systematic errors in variables
δ	the "artificial" relative systematic error added to the "experimental" relative systematic error of x_c
$\delta(\Delta H)$	random error in measured enthalpy change per mole of injected target

$\Delta(\Delta H°)$	systematic error in the ITC-derived enthalpy change for the formation of one mole of complex
$\mathbf{\Delta^*}$ and $\mathbf{\Delta^{*\prime}}$	two different vectors of relative systematic errors of variables
ΔE_0	deviation of nominal E_0 from true (unknown) E_0
δE_0	random error in stock enzyme concentration
ΔG^o	Gibbs free energy
ΔH	enthalpy change per mole of injected target in ITC
$\Delta H_{cum,i}$	cumulative ΔH after the i-th injection in the context of ITC
$\Delta H_{cum,max}$	the maximum cumulative ΔH value in the context of ITC
ΔH^o	enthalpy change associated with the formation of one mole of the target–ligand complex
ΔK_d	the deviation of determined K_d ($K_{d,det}$) to its (unknown) accurate/true value
δL_0	random error in the concentration of stock ligand solution
ΔQ_i	systematic error in the detected total heat absorbed/released after i-th injection in the context of ITC
ΔS_0	deviation of nominal S_0 from true (unknown) S_0
δS_0	random error in stock substrate concentration
ΔS^o	the change in entropy associated with the formation of one mole of the target–ligand complex
δT_0	random error in the concentration of stock target solution
$\Delta \Gamma$	systematic error in the molar ratio between the target and the ligand molecules
ε	molar extinction coefficient
φ	factor by which the signal of ligand changes when it binds the target

$\boldsymbol{\Delta^*_{max}}$ the vector of relative systematic errors of variables that leads to the largest systematic error of the determined parameter

$\boldsymbol{\Delta^*_{min}}$ the vector of relative systematic errors of variables that leads to the smallest systematic error of the determined parameter

ΔL_0 deviation of nominal L_0 from true (unknown) L_0

Δp systematic error of the determined parameter

ΔR systematic error in the fraction of unbound ligand (or enzyme)

ΔS systematic error of measured signal

ΔT_0 deviation of nominal T_0 from true (unknown) T_0

Δx_c systematic error of the fixed (constant) independent variable

Δx_v systematic error of the varied independent variable

Chapter 1. Introduction

Equilibrium constants are essential parameters in understanding chemical and biological processes. These constants provide a measure of the reaction's tendency to proceed and help determine the relative stability of the involved species.[1, 2] Common equilibrium (or pseudo-equilibrium) constants include the equilibrium dissociation constant (K_d) of biomolecular complexes, the acid dissociation constant (K_a) for acids, the stability constant (K_s) of metal–ligand complexes, the Michaelis constant (K_m) of enzymatic reactions, and Langmuir constant (K_L) of adsorbent–adsorbate interactions.[3, 4, 5, 6, 7, 8, 9] Each of these constants offers valuable insights into the behaviour of its corresponding system, from quantifying the strength of molecular interactions to predicting the strength of acids and bases.[3, 4, 5, 6, 7, 8, 9] Equilibrium constants are indispensable across diverse disciplines, including environmental chemistry, pharmacology, and industrial applications, offering a foundational framework for characterizing and optimizing reaction dynamics.

Among the commonly studied equilibrium constants, the equilibrium dissociation constant (K_d) of a reversible molecular interaction is arguably the most frequently measured thermodynamic physicochemical parameter in molecular science, reported in over 1,000 peer-reviewed papers per month.[10, 11] The most common type of reversible molecular interaction in biological and technological contexts is the 1:1 binding model, in which a ligand (L) binding noncovalently and reversibly to a target (T), resulting in the formation of a ligand–target complex (C):

$$\mathrm{T} + \mathrm{L} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \mathrm{C} \qquad (1.1)$$

Here, k_{on} and k_{off} are rate constants of complex formation and dissociation, respectively. The K_d value represents the affinity between the two binding partners and can be defined through either the rate constants or equilibrium concentrations of the target (T), ligand (L), and complex (C), as follows:

$$K_d = \frac{k_{off}}{k_{on}} = \frac{TL}{C} \tag{1.2}$$

The lower the K_d value, the higher the affinity between the molecules, which is usually desired in the development of drugs and diagnostic probes.[12]

In this dissertation, I focus on the 1:1 reversible binding model, as described in **Eq (1.1)** to streamline the analysis and establish a solid foundation for studying molecular interactions. This model represents the simplest and most prevalent form of binding in biological and technological systems, making it an ideal starting point for method development and validation. More complex binding models, such as 1:n or cooperative interactions, require additional parameters and more sophisticated fitting models, which can introduce ambiguity and are reserved for future work. To ensure that the studied systems conform to 1:1 stoichiometry, experimental confirmation using established analytical techniques is essential. Methods such as capillary electrophoresis (CE) and isothermal titration calorimetry (ITC) provide high-confidence assessments of binding stoichiometry by independently resolving bound and unbound species or by quantitatively analyzing the thermodynamic signatures associated with complex formation. The use of such orthogonal techniques helps ensure the validity of applying the 1:1 binding model and strengthens the reliability of subsequent thermodynamic analysis.

1.1. Historical Development of K_d

The history of K_d determination is deeply intertwined with the development of chemical kinetics and thermodynamics. In the 1860s, Cato Guldberg and Peter Waage established the Law of Mass Action, the foundational principle for equilibrium constants, including K_d.[13] They proposed that the rate of a chemical reaction is proportional to the product of the concentrations of the reactants, leading to the concept of chemical equilibrium.

According to the Law of Mass Actio n, the rates of reaction for each species in **Eq (1.1)** are:

$$\frac{d[\mathrm{C}]}{dt} = -\frac{d[\mathrm{T}]}{dt} = -\frac{d[\mathrm{L}]}{dt} = k_{\mathrm{on}}[\mathrm{T}][\mathrm{L}] - k_{\mathrm{off}}[\mathrm{C}] \tag{1.3}$$

where [T], [L], and [C] represent the instantaneous concentrations of target, ligand, and complex, respectively. When the reversible binding process demonstrated in **Eq (1.1)** reaches equilibrium, [T] = T, [L] = L, and [C] = C remain constant, leading **Eq (1.3)** equal to zero, which is:

$$k_{\mathrm{on}}TL - k_{\mathrm{off}}C = 0 \Rightarrow k_{\mathrm{on}}TL = k_{\mathrm{off}}C \tag{1.4}$$

From this relationship, the equilibrium dissociation constant K_d is defined through **Eq (1.2)**, *i.e.*, $K_d = k_{off}/k_{on} = TL/C$.

In the late 19th century, the development of thermodynamics by scientists like Josiah Willard Gibbs and Hermann von Helmholtz provided a theoretical framework for understanding the equilibrium constants in terms of free energy changes.[14] The connection between K_d and the standard Gibbs free energy change (for the binding process **Eq (1.1)**), ΔG^{o}, is given by the following equation:[14]

$$\Delta G^{\circ} = RT\ln\left(K_{\mathrm{d}}\right) \tag{1.5}$$

Here, R is the gas constant (8.314 J/mol·K) and T is the temperature in Kelvin (K).

Because ΔG° is dependent on both enthalpy and entropy, the dissociation constant K_d is inherently sensitive to temperature. The Gibbs free energy equation shows the thermodynamic relationship between free energy, enthalpy, and entropy as:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \tag{1.6}$$

This equation indicates that changing the temperature alters the relative contributions of enthalpy (ΔH°) and entropy (ΔS°) to the free energy (ΔG°), and therefore affects the equilibrium position. Specifically, for exothermic binding reactions ($\Delta H^\circ < 0$), increasing the temperature reduces the favorability of complex formation, leading to a larger K_d (weaker binding). Conversely, for endothermic reactions ($\Delta H^\circ > 0$), increasing the temperature makes binding more favorable, resulting in a smaller K_d (stronger binding). These trends can be understood intuitively as the system shifting to oppose the effect of added heat, releasing heat discourages further complex formation when the system is externally heated, while heat-absorbing reactions are enhanced by additional thermal energy.

The temperature dependence of K_d can also be interpreted from a kinetic perspective. Since K_d is the ratio of the dissociation and association rate constants (**Eq (1.2)**), and both k_{on} and k_{off} typically follow temperature-dependent Arrhenius behavior of the form:

$$k = Ae^{-E_a/RT} \tag{1.7}$$

Temperature can influence both rates, though not necessarily to the same degree. In **Eq (1.7)**, k is the rate constant that quantifies how quickly a process occurs (*e.g.*, k_{on} and k_{off} in **Eq (1.2)**). The pre-exponential factor A (also known as the frequency factor) reflects the frequency and orientation of molecular collisions that are conducive to reaction, while the activation energy E_a represents the minimum energy barrier that must be overcome for the process to proceed.

As temperature increases, both the association rate constant k_{on} and k_{off} generally increase. However, if k_{off} increases more rapidly, due to a higher activation energy, then K_d also increases, indicating weaker binding at elevated temperatures. Conversely, if k_{on} is more sensitive to temperature, K_d decreases, reflecting stronger binding. These kinetic trends are consistent with thermodynamic expectations, as the overall enthalpy change of binding is approximately governed by the difference in activation energies between the forward and reverse reactions.

Taken together, these thermodynamic and kinetic considerations demonstrate that K_d values are inherently temperature-dependent. Accurate determination of K_d, therefore, requires careful control and reporting of temperature, especially when comparing affinities across different systems or applying *in vitro* measurements to *in vivo* conditions.

Eqs (1.2) and **(1.5)** provide three possible methods to determine K_d: (*i*) measuring k_{on} and k_{off}; (*ii*) measuring equilibrium concentrations (or the relation between the equilibrium concentrations), and (*iii*) measuring standard Gibbs free energy of the binding reaction.

Since Gibbs free energy cannot be measured directly, K_d values are typically determined using one of two methods: either by measuring k_{on} and k_{off} through kinetic approaches, such as Surface Plasmon Resonance (SPR) and Biolayer Interferometry (BLI), or by determining the ratio of unbound (or bound) ligand concentrations at equilibrium using equilibrium approaches like Fluorescence Anisotropy (FA), Enzyme-Linked Immuno-Sorbent Assays (ELISA), Microscale Thermophoresis (MST).[15] Note that, when K_d is determined, Gibbs free energy ΔG^o can be calculated through **Eq (1.5)**.

1.2. Significance of Accurate K_d Determination Across Scientific Fields

The equilibrium dissociation constant (K_d) is arguably one of the most frequently measured thermodynamic and physicochemical parameters in molecular science. While a substantial

number of K_d values are reported in the scientific literature, this reflects only a fraction of the field's true extent. Many additional K_d determinations are carried out internally by pharmaceutical and biotechnology companies using highly parallelized, high-throughput screening methods that are typically not published.[10, 11] These platforms enable the rapid assessment of binding affinities across thousands of candidate compounds, accelerating the identification of promising drug leads. The continued prevalence of K_d measurements in both published and proprietary settings highlights their central role in advancing both fundamental research and applied biomedical discovery.

The accurate determination of K_d values is crucial for understanding molecular interactions, which are foundational to many biological and chemical processes.[16, 17] Critical decisions in both academic research and industrial production often hinge on K_d values. In the pharmaceutical industry, for example, K_d values are vital for the drug discovery and development process.[18, 19] They are used to rank and select drug candidates based on their binding affinity to specific targets, ensuring that only the most promising compounds are pursued in the costly and time-consuming stages of clinical development.[12] In the field of biotherapeutics, K_d measurements are essential for evaluating the potency of antibodies and aptamers, which are engineered to bind selectively to disease-related targets.[20, 21] Understanding the binding affinity of these molecules allows researchers to optimize their therapeutic properties and improve their efficacy in treating diseases such as cancer and autoimmune disorders.[22, 23] Similarly, K_d values are critical in the design and calibration of biosensors, which detect specific molecules in complex biological samples.[24] High sensitivity and specificity in biosensors are often achieved by selecting or engineering recognition elements with appropriate K_d values. Furthermore, K_d values play a key role in deciphering biological pathways. By quantifying the binding affinities of various

molecular interactions within a cell, scientists can map out the complex network of interactions that regulate cellular functions and responses to environmental changes.[25] This information is crucial for understanding the mechanisms of diseases at a molecular level and for identifying potential targets for therapeutic intervention. Therefore, the accurate determination of K_d values is not only fundamental to advancing our understanding of molecular interactions but also pivotal in translating this knowledge into practical applications across diverse fields such as drug discovery, biotherapeutics, biophysical technology, and the study of complex biological systems.

1.3. Challenges in Accurate Determination of K_d

1.3.1. Precision VS Accuracy

Precision and accuracy are both essential components of reliable K_d determination, but they address fundamentally different aspects of measurement quality. Precision refers to the consistency and reproducibility of experimentally determined K_d values. A method is considered highly precise if it produces similar K_d results across repeated experiments under identical conditions. Enhancing precision primarily involves minimizing random errors, which may stem from fluctuations in detected signals, variability in sample preparation, or changes in environmental conditions such as temperature, humidity, and atmospheric pressure. High precision reflects a stable and well-controlled experimental setup, indicating that data collection and analysis processes are robust and reproducible. However, precision alone does not guarantee that the measured K_d values are close to the true value.

In contrast, accuracy pertains to how close a determined K_d value is to its true, though typically unknown, thermodynamic value. While precision concerns reproducibility, accuracy concerns correctness. Accurate K_d values depend on minimizing or correcting systematic errors, which may arise from methodological limitations, concentration misassignments, detector

calibration issues, or biases introduced during data analysis. An accurate measurement reflects the actual binding affinity between the ligand and its target under physiological or experimentally relevant conditions. Therefore, high accuracy is essential for drawing valid biological or pharmacological conclusions from experimental data.

Importantly, precision and accuracy are not interchangeable. A method may be highly precise but systematically biased, consistently yielding incorrect values. Conversely, a less precise method may, on average, produce accurate results if systematic errors are well controlled. Thus, both high precision and high accuracy are required for trustworthy K_d determination: precision ensures consistency, and accuracy ensures validity.

To clarify the terminology used throughout this dissertation, it is important to distinguish between signal, noise, and error, and to introduce the concept of the signal-to-noise ratio (S/N). These definitions are applied consistently in all subsequent sections/chapters.

The signal refers to the measurable physical response that conveys information about molecular binding, such as detected fluorescence intensity, calorimetric heat pulses, or absorbance changes. The signal represents the deterministic component of the measurement that would be observed under ideal conditions.

Noise, in contrast, refers specifically to random, unpredictable fluctuations superimposed on the signal. These fluctuations may arise from electronic or thermal disturbances in detectors, stochastic photon emission and detection events, or other inherent sources of physical variability. In experimental practice, noise is commonly quantified as the standard deviation of the background signal in the absence of a specific binding event. Throughout this dissertation, noise is defined strictly as the random component of the measured signal and does not include systematic deviations such as baseline drifts, calibration errors, or concentration inaccuracies.

Error is a more general term encompassing all deviations between a measured or inferred value and its true value. In this work, error is understood to have two main components: random error, which arises from noise and contributes to variability across replicates (thus affecting precision), and systematic error, which introduces a consistent bias across measurements (thus affecting accuracy). The distinction between these two types of error is central to interpreting and evaluating K_d measurements.

Within this framework, the signal-to-noise ratio (S/N) serves as a useful metric for evaluating measurement quality. S/N quantifies the relative strength of the true signal to the magnitude of the noise. In general, a higher S/N leads to improved precision, as random fluctuations have less influence on parameter estimation. However, a high S/N does not by itself ensure accuracy: systematic biases, such as incorrect model assumptions or concentration errors, can still produce inaccurate results. Conversely, even low-S/N data may lead to accurate K_d estimates if analyzed using robust methods that effectively mitigate systematic error.

In summary, precision is governed primarily by the influence of random noise and error, while accuracy depends on the detection, correction, or elimination of systematic biases. Throughout this dissertation, the terms signal, noise, error, precision, accuracy, and S/N are used consistently to support a clear and rigorous evaluation of the reliability and limitations of various K_d determination approaches.

1.3.2. Determining the Target K_d Accuracy: Application-Dependent Considerations

No single level of accuracy suits every K_d-determination experiment; the "right" targeted K_d accuracy depends on the scientific or practical decision the number will inform. In exploratory high-throughput screens, the goal is usually to rank hundreds of ligands rather than to pin down their exact affinities. A rough tolerance of two- to three-fold is often sufficient to pick out the

handful of compounds worth carrying forward. By contrast, once chemists begin structure-activity-relationship work, they need to determine whether a subtle change to a functional group genuinely improves binding. Here, an error band of twenty percent or less becomes desirable so that small but meaningful affinity gains are not masked by measurement noise.

Biophysical research that feeds thermodynamic or kinetic models benefits from even tighter control, because any systematic uncertainty in determined K_d will propagate into calculated free energies, rate constants, or stoichiometries. Investigators commonly aim for accuracies better than ten percent to keep these downstream parameters reliable. Clinical contexts impose their own accuracy requirements. In diagnostic applications, calibrators used in binding-based assays must be accurate enough to ensure that key medical decision points, such as diagnostic thresholds, are reliably maintained. While perfect accuracy is not essential, especially in light of biological variability among patients, the K_d values used must still support consistent and clinically meaningful interpretation. In therapeutic development, particularly for monoclonal antibodies, accuracy becomes even more critical. A modest error in K_d can significantly affect predicted receptor occupancy, which directly influences dose selection for first-in-human trials. As a result, regulators often expect a level of confidence comparable to that demanded in detailed mechanistic studies, where accurate quantification is essential for modeling drug–target interactions.

Whatever the application, researchers should begin by asking, “What consequence will an inaccurate K_d have on my conclusion?” and then design the experiment so that the projected systematic uncertainty falls comfortably below that threshold. They should still pursue the best accuracy their resources allow, recognising that an absolute floor exists. Pipetting tolerances, temperature drift, protein heterogeneity, and the simplifying assumptions baked into every

binding model place a practical limit, rarely better than five percent, on even the most diligent effort. When estimates of systematic error suggest the attainable accuracy lies above the requirement, it may be wiser to combine orthogonal techniques or refine the experimental setup than to accept a number whose uncertainty clouds the decision it is meant to guide.

1.3.3. Difficulties Faced in Determining Accurate K_d

Determining accurate physicochemical parameters is relatively straightforward for quantities that can be directly measured against established reference standards, such as length, mass, time, temperature, or electric charge.[26, 27, 28, 29] However, determining accurate K_d values for molecular interactions presents a fundamental challenge. Unlike physical constants, K_d lacks a direct reference standard, and its true value is highly sensitive to experimental conditions, including temperature, pH, buffer composition, *etc*.[30, 31] These factors inherently affect the binding equilibrium, rendering the K_d value context-dependent. Even under identical chemical conditions for the same molecular pair, measured K_d values often vary significantly due to a combination of method-dependent and method-independent inaccuracies.[32, 33]

Method-dependent inaccuracies arise from limitations or artifacts intrinsic to the measurement technique, such as molecule immobilization, labelling, or unintended alterations of experimental conditions (*e.g.*, temperature or pH fluctuations) during the process.[32, 33] Consequently, K_d values for the same molecular pair under the same conditions can differ by orders of magnitude when determined using different approaches.[32] Method-independent inaccuracies arise from systematic errors in experimental variables, such as signal detection and analyte concentration measurements. These errors propagate through all measurement methods, with their magnitude and impact influenced by the specific experimental design. When combined with differences in experimental conditions, these errors can significantly amplify

variability in measured K_d values, even when the same method is employed.[34] As a result, discrepancies in reported K_d values that are determined with the same method can also reach orders of magnitude,[35] underscoring the pervasive influence of method-independent factors. Together, all the sources of systematic errors discussed above underscore the complexity of accurately determining K_d and highlight the critical importance of rigorous experimental design, meticulous control of variables' errors, and careful interpretation of results.

1.4. Comprehensive Literature Review on Efforts to Enhance the Accuracy of K_d Determination

The accurate determination of K_d is fundamental to molecular interaction studies, but it remains challenging due to both method-dependent and method-independent inaccuracies. Over the past decades, extensive efforts have been made to overcome these challenges, leading to significant advancements in measurement techniques and analytical approaches. This section provides a comprehensive review of some mainstream developments, highlighting their achievements, inherent limitations, and areas requiring further improvement. By citing relevant literature, it offers a critical perspective on recent progress in enhancing K_d determination accuracy and identifies key gaps that must be addressed in future research.

1.4.1. (Method-Dependent) Technique Innovations and Improvements

Method-dependent inaccuracies in K_d determination primarily arise from artifacts and limitations inherent to specific measurement techniques, with molecular immobilization and labelling being the most common sources of error.[11] These modifications can perturb native binding kinetics, introduce systematic biases, and distort affinity measurements, leading to deviations from true K_d values. Additionally, unintended changes in experimental conditions

(*e.g.*, temperature, pH, *etc*.) during the measurement process may further exacerbate these inaccuracies.[36]

Immobilization and labelling are widely used in biophysical techniques for studying molecular interactions, while they frequently introduce significant artifacts that compromise the accuracy of K_d measurements.[37, 38] Fluorescent or isotopic labelling can alter a molecule's native conformation or binding dynamics, while immobilization on solid surfaces may induce steric hindrance, nonspecific binding, or other surface-related artifacts.[37, 38] These factors can distort observed binding affinities and kinetic profiles, leading to erroneous K_d determinations.

To mitigate these effects, researchers have developed innovative approaches over the past decades to minimize the impact of labelling and/or immobilization on binding measurements. These advancements include label- and/or immobilization-free detection techniques, optimized surface chemistries, and computational corrections for experimental artifacts. This section explores modern techniques for binding affinity measurement, emphasizing strategies that preserve biomolecular integrity, enhance the precision and accuracy of K_d determination, and address key limitations. It should be noted that qualitative assays, such as immunology-based methods (*e.g.*, ELISA), and certain conventional techniques like fluorescence anisotropy (FA) are excluded from this analysis.

1.4.1.1. Surface Plasmon Resonance (SPR)

Surface Plasmon Resonance (SPR) is a widely employed optical technique for studying biomolecular interactions, offering real-time, label-free analysis. SPR relies on the excitation of surface plasmons, which are oscillations of free electrons at the interface of a gold-coated sensor and a dielectric medium, induced by polarized light.[39, 40] Binding events at the sensor surface cause changes in the refractive index, leading to detectable shifts in the SPR angle.[39, 40] The

magnitude of these shifts is directly proportional to the quantity of target molecules bound to the sensor surface, enabling precise quantification of binding kinetics, including the association and dissociation rate constants (k_{on} and k_{off}), as well as the equilibrium dissociation constant K_d.[39, 40]

Although SPR is a label-free technique, the immobilization of molecules remains the primary source of inherent inaccuracy in the SPR method. Besides immobilization, another major source of error in SPR-based K_d determination is the nonspecific binding of target molecules to the solid surface.[41, 42] These issues can lead to overestimated or underestimated SPR-measured parameters, affecting the reliability of kinetic and equilibrium analyses.[43]

Over the years, significant advancements have been made, particularly in optimizing surface chemistry to mitigate nonspecific interactions. Mrksich et al. (1997) showed that self-assembled monolayers (SAMs) with oligo(ethylene glycol) groups effectively suppress nonspecific binding, improving the accuracy in SPR measurements.[44] Similarly, Pasche et al. (2005) demonstrated that enhanced surface coatings, such as PEGylated linkers, reduce nonspecific binding and maintain the biological activity of immobilized ligands.[45] More recently, Ye et al. (2020) explored zwitterionic polymer-based surface modifications, which further minimized background noise and improved the signal-to-noise ratio in SPR assays.[46] These developments highlight the importance of surface chemistry optimization in achieving more accurate and reproducible SPR-based K_d determination.

Another significant challenge in SPR-based K_d determination is mass transport limitation, where the diffusion rate of analytes to the sensor surface influences observed binding kinetics. Schuck and Zhao (2010) introduced mathematical models to correct these effects, improving the reliability of K_d measurements.[47] Instrumental advancements have also played a crucial role, with Mayer and Hafner (2011) emphasizing the importance of higher sensitivity detectors and

precise fluidics in modern SPR platforms, which enable the detection of low-affinity interactions and enhance K_d accuracy.[48]

Together, these improvements in surface chemistry and instrumentation have strengthened SPR as a cornerstone technique for precise K_d determination. However, its accuracy remains constrained by the need for molecular immobilization, which can obstruct binding sites and potentially disrupt native molecular interactions.

1.4.1.2. Biolayer Interferometry (BLI)

Biolayer Interferometry (BLI) is a label-free optical biosensing technology that enables real-time measurement of molecular interactions, providing a robust alternative to other techniques such as SPR. BLI functions by detecting changes in the optical interference pattern of light reflected from a biosensor surface.[49, 50] Binding events between the immobilized ligand and the target molecules cause measurable shifts proportional to the amount of bound target.[49, 50] Since its introduction, BLI has gained popularity for determining K_d due to its simplicity, robustness, and compatibility with high-throughput workflows.

In addition to being label-free, a key advantage of BLI is its operational flexibility. Unlike SPR, BLI is relatively insensitive to bulk refractive index changes, allowing for accurate measurements in complex sample matrices, such as crude extracts or serum.[51] This robustness, coupled with the use of disposable sensors in platforms like the Octet system, reduces cross-contamination risks and simplifies assay preparation.[49] These features make BLI particularly useful in early-stage drug discovery, antibody characterization, and biomolecular screening applications.

Recent advancements in BLI technology have aimed at enhancing the reliability of BLI-derived K_d values. Similar to SPR, improvements in sensor surface chemistries, such as

biocompatible coatings and site-specific ligand immobilization, have significantly minimized nonspecific binding while preserving ligand integrity.[49, 50] These innovations have established BLI as a versatile method for studying biomolecular interactions across diverse experimental conditions. However, like SPR, the immobilization of ligand molecules remains a potential limitation, as it can introduce artifacts that affect the accuracy of determined K_d values. Addressing this challenge is critical for further refining BLI's applications in quantitative binding studies.

1.4.1.3. Isothermal Titration Calorimetry (ITC)

Isothermal Titration Calorimetry (ITC) is widely regarded as the gold-standard method for studying molecular interactions, providing direct, label- and immobilization-free measurements of binding thermodynamics.[52, 53, 54] By quantifying heat released or absorbed during binding events, ITC generates comprehensive thermodynamic profiles, including equilibrium dissociation constants (K_d), binding enthalpy ($\Delta H°$), entropy ($\Delta S°$), Gibbs free energy ($\Delta G°$), and stoichiometry (n).[55, 56] Introduced in the 1960s, ITC has evolved into a reliable and versatile approach for analyzing diverse biomolecular interactions, from protein–small-molecule complexes to intricate macromolecular assemblies.[57, 58, 59]

Initially, ITC instrumentation was limited by low sensitivity, high sample consumption, and insufficient measurement precision, which restricted its use primarily to high-affinity interactions.[60] However, technological advancements in calorimeter design have significantly enhanced sensitivity and reduced sample requirements, enabling precise measurement of weaker interactions previously challenging to characterize. Modern high-sensitivity calorimeters, such as those highlighted by Krainer et al. (2012), precisely detect minute heat changes, expanding ITC's applicability across a broader range of affinities.[61] The reduced sample volumes

requirements have also facilitated research on scarce or expensive biomolecules, such as therapeutic antibodies and rare enzymes.[62]

Complementary integration of ITC with analytical methods like SPR and BLI further validates K_d measurements by combining methodological strengths and mitigating individual limitations.[63] Wu and Piszczek (2002) emphasized that such hybrid approaches enhance confidence in the results and provide comprehensive insights into molecular interactions.[64] Additionally, innovations like microfluidic ITC, demonstrated by Jia et al. (2017), have further reduced sample usage and improved sensitivity, allowing characterization of interactions previously inaccessible due to technical constraints.[65]

Despite these advancements, ITC continues to face challenges, including relatively high sample consumption, lengthy experiments, and difficulty to accurately examine very weak or ultra-tight affinities due to residual sensitivity limitations.[53] Systematic issues such as heat signal variability, buffer mismatch, and baseline corrections also complicate data interpretation.[53] Ongoing developments in calorimeter technology, advanced data processing algorithms, and strategic combinations with other analytical techniques promise continued enhancement of ITC's precision, reliability, and practicality.[53] Its continuous evolution ensures that ITC remains invaluable in both fundamental research and applied sciences, such as drug discovery and biotechnology.

1.4.1.4. Microscale Thermophoresis (MST)

Microscale Thermophoresis (MST) is a highly sensitive, immobilization-free technique for determining K_d by measuring the thermophoretic movement of molecules in a temperature gradient.[66] First introduced as a quantitative binding assay by Duhr and Baaske (2010), MST has gained significant traction in biophysical and pharmaceutical research due to its rapid, low-

volume analysis and ability to study molecular interactions in solution without immobilization.[66, 67, 68]

MST is particularly advantageous for studying a wide range of biomolecular interactions, including protein–small-molecule, protein–protein, and nucleic acid involved interactions.[69] Its ability to precisely measure binding affinities across a broad dynamic range, from picomolar to millimolar, makes it a versatile tool for drug discovery, structural biology, and enzymology.[69] This technique has been successfully applied in high-throughput screening, particularly in assessing small molecule inhibitors, antibody-antigen interactions, and fragment-based drug discovery.[70]

The commercial development and continued optimization of MST technology have been led by NanoTemper Technologies, which pioneered the instrumentation and software that improved detection sensitivity and throughput.[70] Modern MST platforms incorporate fluorescence-labelling and label-free detection modes, allowing broader applications across biomolecular research and drug development.[71]

Despite its strengths, MST has notable limitations. The technique is highly sensitive to buffer composition, viscosity, and the presence of stabilizing agents, which can introduce variability in K_d measurements.[67] Additionally, temperature changes during MST measurements can potentially lead to inaccuracies in K_d determination.[72] Furthermore, the utility of MST is limited in studying very weak interactions or highly hydrophobic molecules that may unpredictably influence thermophoretic behaviour.[73] Nevertheless, MST remains a valuable tool in quantitative biophysical studies, particularly for applications requiring rapid affinity measurements. As advancements in automation and data processing continue, MST is expected to play an increasingly integral role in biomolecular interaction studies and drug discovery.

1.4.1.5. Nuclear Magnetic Resonance (NMR)-Based Methods

Nuclear Magnetic Resonance (NMR) is an immobilization-free technique that enables precise K_d determination by directly analyzing molecular interactions. Unlike calorimetric or optical methods, this technique provides atomic and molecular-level insights, making it particularly useful for weak to moderate affinity interactions and complex biomolecular systems.[35]

NMR spectroscopy measures molecular interactions through chemical shift perturbations, relaxation rates, and nuclear Overhauser effects.[74, 75] Chemical shift perturbation analysis, a widely used approach, tracks stepwise changes in peak positions upon target titration.[76] By fitting these shifts to binding models, researchers expect to derive precise K_d values, particularly for low-affinity interactions, typically spaning from millimolar to micromolar range.[35] Shi and Zhang (2021) indicate that advanced techniques such as saturation transfer difference and Carr-Purcell-Meiboom-Gill relaxation dispersion have enhanced NMR's utility in fragment-based drug discovery and transient interaction studies.[77] Hyperpolarization methods further improve sensitivity, facilitating analyses of low-abundance biomolecules.[78]

However, certain NMR-based methodologies for determining K_d have inherent limitations. Heteronuclear experiments, such as those involving ^{15}N or ^{13}C nuclei, require isotopic labelling to enhance sensitivity,[35] which increases cost and experimental complexity. This labelling process may introduce perturbations that could affect the accuracy of the determined K_d values. Additionally, NMR experiments often demand large sample quantities, typically in the millimolar concentration range, making the technique impractical for scarce biomolecules.[79] Furthermore, spectral overlap and signal broadening present challenges when studying large macromolecules.[80] As molecular size increases, slower rotational correlation times lead to

reduced spectral resolution, necessitating advanced labelling strategies and multidimensional techniques for accurate data interpretation.[80]

1.4.1.6. Mass Spectrometry (MS)-Based Methods:

Mass Spectrometry (MS) is a highly sensitive and direct method for K_d determination, capable of detecting ligand–target complexes in the gas phase with minimal sample requirements.[81] Unlike NMR, MS avoids spectral overlap and does not require isotopic labelling, making it particularly advantageous for high-throughput binding studies.

Native electrospray ionization mass spectrometry (native ESI-MS) preserves the native biomolecular structures during ionization, enabling precise titration analyses of target–ligand interactions, including transient and multi-site binding events.[82, 83] This approach allows for the characterization of weak and dynamic interactions that might be challenging to detect using other biophysical techniques. Additionally, hydrogen-deuterium exchange mass spectrometry provides structural insights by monitoring amide hydrogen exchange upon ligand binding, making it a powerful tool for epitope mapping and drug discovery applications.[84]

Despite these advantages, MS-based K_d determination faces several challenges. Ionization biases can lead to variations in signal intensity, potentially affecting quantification accuracy.[85] Maintaining physiological solution conditions in the gas phase remains difficult, and nonspecific binding artifacts can complicate data interpretation.[86] Furthermore, while MS excels in detecting binding interactions, it lacks the atomic-level resolution provided by NMR, limiting its ability to precisely map binding interfaces.

Specialized instrumentation and expertise are essential for accurate data acquisition and interpretation in MS-based approaches. However, compared to NMR, MS is generally more accessible due to its lower cost and shorter experimental times.[87] As a result, MS serves as a

complementary technique to traditional methods like ITC, SPR, and BLI, offering a high-resolution, label-free alternative for K_d determination.[50] Advances in hardware, computational modeling, and automation continue to expand the role of MS in quantitative molecular interaction studies, further strengthening its impact in drug discovery, structural biology, and biotechnology.

1.4.1.7. Capillary Electrophoresis (CE)-Based methods

Capillary electrophoresis (CE) is a widely used analytical technique known for its high resolution, minimal sample consumption, and rapid analysis. Its capability in determining K_d of molecular complexes is well established. CE-based K_d determination has been applied across diverse systems, making it an invaluable tool for quantitative affinity studies.

Affinity capillary electrophoresis (ACE) remains one of the most widely used CE-based methodologies for studying molecular interactions. ACE relies on monitoring the variations in the electrophoretic mobility of an analyte as a function of its binding partner's concentration,[88] offering a direct and homogeneous (immobilization-free) approach to determining K_d with high precision. This technique has been widely employed in studying protein–small molecule, protein–aptamer, and protein–protein interactions, demonstrating its versatility across diverse biomolecular systems.[89, 90, 91]

Another significant advancement is capillary electrophoresis–frontal analysis (CE-FA), which enables direct measurement of free and bound ligand concentrations at equilibrium.[92] CE-FA is especially valuable for characterizing high-affinity interactions, where conventional electrophoretic methods may struggle due to the long analysis times that can lead to complex dissociation, even for slowly dissociating complexes.[93] In parallel, microchip capillary electrophoresis (MCE) has garnered interest for its suitability in high-throughput screening.[94] By

integrating microfluidic technology with CE, MCE facilitates rapid and efficient affinity measurements while minimizing reagent consumption.[95] This makes it particularly well-suited for pharmaceutical and biomolecular screening applications, where automation and miniaturization are essential.

The Krylov Lab at York University (Toronto, Canada) has been at the forefront of developing CE-based approaches to enhance K_d determination. A key innovation is nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM), which enables simultaneous determination of K_d and k_{off} in a single experiment.[96] This technique has been effectively applied to study protein–small-molecule, protein–DNA, and DNA–small-molecule interactions, providing detailed binding profiles from minimal sample quantities.[97, 98] Building upon NECEEM, Krylov lab introduced kinetic capillary electrophoresis (KCE), a conceptual framework that extends CE applications to both affinity and kinetic analyses.[99, 100] KCE allows for precise characterization of binding and dissociation dynamics, enhancing the reliability of biomolecular interaction studies.[99, 100] Additionally, the Krylov Lab has developed computational tools to improve CE data analysis, addressing the inverse problem of KCE and enabling more accurate kinetic parameter extraction.[101]

Despite its advantages, CE-based K_d determination faces several challenges, including sensitivity to buffer composition and ionic strength, as well as nonspecific adsorption to capillary walls, all of which can contribute to measurement variability.[102] Recent innovations, such as microfluidic-based CE and MS-coupled CE, have improved detection sensitivity and throughput.[91, 103, 104] These advancements continue to refine CE's role in characterizing biomolecular interactions, solidifying its position as a powerful tool for quantitative binding studies.

Table 1.1 provides a comparative overview of the K_d-determination methods discussed above, evaluating each technique against key operational and analytical criteria, including labeling or immobilization requirements, sample volume, measurable K_d range, throughput, susceptibility to artifacts, and recent technological advancements. As reflected in both the table

Table 1.1. Comparative characteristics of experimental K_d-determination methods.

Technique	Labeling Required	Immobilization Required	Sample Volume	K_d Range	Throughput	Main Systematic Error Sources	Recent Advancements
SPR	No	Yes	Low–Moderate	pM–mM	High	• Immobilization • Nonspecific binding • Mass transport limits	• Optimized surface chemistry • Improved detectors
BLI	No	Yes	Low	pM–mM	High	• Immobilization • Nonspecific binding	• Improved sensor coatings • Disposable tips
ITC	No	No	High	nM–mM	Low	• High ligand concentration requirements • Heat drift	• Microfluidic ITC • Sensitivity enhancement
MST	Yes/ Optional	No	Low	pM–mM	Medium	• Buffer sensitivity • Temperature effects • Hydrophobicity bias	• Label-free modes • Automation • Improved software
NMR-based methods	Yes/ Optional	No	High	μM–mM	Low	• Isotope labeling • Spectral overlap	• Hyperpolarization • Advanced 2D/3D NMR • Better resolution
MS-based methods	No	No	Low	nM–μM	High	• Ionization artifacts • Non-physiological conditions	• Native ESI-MS • HDX-MS • Improved instrumentation
CE-based methods	Yes/ Optional	No	Very Low	pM–μM	Medium	• Buffer sensitivity • Capillary-wall adsorption • Joule heating	• MCE, NECEEM, KCE, • MS-coupling • Improved computational modeling

and preceding discussion, each method offers distinct advantages and limitations, particularly in terms of labeling, immobilization, dynamic range, and sources of systematic error.

Among commonly used techniques, only ITC and MS-based techniques are truly free from both labeling and immobilization, enabling measurements under conditions that better reflect native biological environments. ITC remains a gold standard for label-free analysis in solution, though its requirement for high ligand concentrations can amplify experimental uncertainties and propagate e into the determined K_d values. MS-based approaches, enhanced by innovations such as native ESI and HDX-MS, offer high sensitivity but may be affected by ionization artifacts and non-physiological conditions that compromise K_d accuracy. Ultimately, the choice of method should be guided by the specific experimental context, characteristics of the target–ligand system, and the desired balance between throughput, accuracy, and physiological relevance.

1.4.2. Method-Independent Strategies for Enhancing K_d Accuracy

While method-dependent factors contribute to inaccuracies in K_d determination, method-independent strategies focus on improving measurement accuracy across all techniques. Two primary areas that significantly impact K_d-determination accuracy are the accurate determination of reagent concentrations and the application of robust data-analysis methods. Systematic errors in analyte concentrations can lead to systematic deviations in binding curves, affecting the accuracy of K_d determination. Standardized concentration determination methods and best laboratory practices help mitigate these issues. Additionally, rigorous data-analysis techniques, such as accounting for concentration and signal uncertainties, applying appropriate fitting models, and using advanced statistical approaches, enhance the reliability of determined K_d values. Although signal accuracy is largely method-dependent, universally accepted practices, such as instrument calibration and baseline subtraction, further contribute to minimizing

measurement artifacts. This section conducts a detailed literature review of existing strategies aimed at mitigating errors arising from inaccurate reagent concentrations and suboptimal data-analysis methodologies.

1.4.2.1. Strategies for Enhancing the Accuracy of Reagent Concentrations

Accurate reagent concentrations are essential for accurate K_d determination, as even minor systematic errors in analyte concentrations can propagate through calculations, leading to significant deviations of the determined K_d from their true values. Studies have shown that miscalculations in ligand and/or target concentrations can lead to artificial shifts in K_d values across different techniques.[105, 106] Common sources of systematic concentration errors include reagent impurities, liquid-handling errors, and fluctuations in experimental conditions such as temperature, air pressure, and humidity.[107, 108] To mitigate these issues, researchers utilize quantification techniques complemented by standardized experimental protocols. This section reviews these strategies in detail, highlighting their role in enhancing concentration accuracy for reliable K_d determination in molecular binding studies.

1.4.2.1.1. Analytical Techniques for Measuring Reagent Purity and Concentration

A wide range of analytical techniques has been employed to ensure reagent quality, each offering distinct strengths, limitations, and suitability for different sample types. Detailed discussions for each method are provided in the subsequent sections.

High-Performance Liquid Chromatography (HPLC): HPLC is an analytical technique used to separate, identify, and quantify components in a mixture based on their interactions with a stationary phase and a liquid mobile phase.[109] In reagent purity measurements, HPLC is widely used to detect and quantify impurities in chemical and biological samples.[110, 111, 112] It plays a

critical role in assessing the purity of pharmaceuticals, proteins, and small molecules by identifying degradation products, contaminants, and residual solvents.[110, 111, 112] By comparing chromatographic profiles with reference standards, HPLC ensures batch-to-batch consistency in manufacturing and quality control processes.[110, 111, 112]

Despite its advantages, HPLC has some limitations, including high operational costs and the need for specialized training. Method development can be complex, requiring careful optimization of mobile phase composition, column (stationary phase) selection, and detection parameters.[113] Additionally, HPLC often relies on large volumes of organic solvents, leading to high waste generation and environmental concerns.[114]

Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE): SDS-PAGE is a widely used technique for assessing protein purity by separating proteins based on their molecular weight. In this method, proteins are denatured and coated with SDS, an anionic detergent that imparts a uniform negative charge.[115] When an electric field is applied, the proteins migrate through a polyacrylamide gel, with smaller proteins moving faster than larger ones.[115] After separation, staining techniques such as Coomassie Brilliant Blue or silver staining allow visualization of protein bands, providing a qualitative assessment of purity and potential contaminants.[116]

SDS-PAGE is commonly used in reagent purity analysis to detect protein impurities, degradation products, and post-translational modifications. It is particularly valuable in assessing the homogeneity of recombinant proteins, antibodies, and enzyme preparations.[117] By comparing sample bands with molecular weight markers, researchers can estimate protein sizes and identify unexpected bands that may indicate contamination or incomplete purification.[118]

Although SDS-PAGE is a valuable technique, its limitations include its qualitative nature and reduced resolution for proteins with similar molecular weights.[119] The intensity of stained bands may not always correlate directly with protein concentration, making it less accurate than chromatographic or spectrometric methods.[119] Additionally, SDS-PAGE is unsuitable for detecting non-protein impurities and requires careful gel preparation and staining procedures.[120] While it remains an essential tool for routine protein purity assessments, it is often complemented by more quantitative techniques such as HPLC or MS for a more comprehensive analysis.

Ultraviolet-visible (UV-Vis) spectroscopy: UV-Vis spectroscopy is a fundamental technique for protein quantification, primarily leveraging the absorbance properties of amino acids such as tryptophan, tyrosine, and phenylalanine.[121] Proteins typically exhibit strong absorbance at 280 nm due to these aromatic residues.[121] This property enables the estimation of protein concentration using UV-Vis spectroscopy in conjunction with the Beer–Lambert law, expressed as:[121]

$$A = \varepsilon l c \tag{1.8}$$

where A represents absorbance (unitless), ε is the molar extinction coefficient (in the unit of $M^{-1} \cdot cm^{-1}$), l denotes the path length (in the unit of cm), and c represents the concentration of the examined protein solution (in the unit of Molar (M)). By measuring the absorbance and knowing the values of ε and l, researchers can calculate the protein concentration c.

Accurate determination of protein concentration using UV-Vis spectroscopy relies on the knowledge of accurate molar extinction coefficient ε, which varies among proteins due to differences in amino acid composition and structural conformation.[122] Factors such as pH and ionic strength of the solution can further influence ε by altering protein structure and charge

distribution.[122, 123] Therefore, relying on standard ε values without considering specific protein properties and solution conditions can introduce significant systematic errors in concentration assessments. To address these challenges, researchers have developed several strategies.

One approach involves empirically determining ε by measuring the absorbance of a protein solution with a known concentration under specific experimental conditions, thereby accounting for variations in buffer composition, ionic strength, and pH.[124] Alternatively, ε can be predicted based on the protein's amino acid composition, particularly the content of tryptophan, tyrosine, and cysteine residues, though this method may not account for environmental factors affecting ε.[125] Maintaining consistent buffer composition, pH, and ionic strength during spectrophotometric measurements minimizes variability in absorbance readings. By implementing these strategies, researchers can enhance the accuracy of protein concentration measurements, thereby improving the reliability of subsequent analyses.

MS and NMR: MS offers high sensitivity and specificity for molecular quantification. By ionizing chemical compounds to generate charged molecules and measuring their mass-to-charge ratios, MS can effectively determine the concentrations of analytes, including proteins and small molecules.[126] Techniques such as electrospray ionization and matrix-assisted laser desorption/ionization have enhanced the applicability of MS in quantitative analysis.[127] Additionally, the use of internal standards, compounds chemically similar to the analyte of interest, is common in MS to correct for variability in sample preparation and instrument response, thereby improving quantification accuracy.[127]

NMR spectroscopy, on the other hand, provides detailed information about molecular structures and dynamics by observing the magnetic properties of atomic nuclei. Quantitative NMR (qNMR) allows for the determination of absolute concentrations without the need for

calibration curves, provided that an internal standard of known concentration is used.[128] This technique is particularly useful for small molecules and metabolites, offering high reproducibility and accuracy.

However, both MS and NMR have limitations, particularly concerning cost and accessibility. MS instruments are expensive, often requiring significant investment, and their operation and maintenance necessitate specialized training. Additionally, the complexity of MS data analysis can be a barrier for laboratories lacking experienced personnel.[129] Similarly, NMR spectrometers are costly and demand substantial infrastructure, including shielding and vibration isolation, making them less accessible to smaller laboratories.[87] Moreover, NMR typically requires larger

Table 1.2. Comparison of techniques for purity and concentration analysis.

Technique	Primary Application	Quantitative Accuracy	Sample Type Compatibility	Key Strengths	Main Limitations
HPLC	Purity and concentration analysis	High	Small molecules, proteins	• High resolution • High sensitivity • Detecting both impurities and degradation products	• High cost • Complicated method development • Solvent waste generation
SDS-PAGE	Protein purity analysis	Low–Moderate	Proteins	• Simple setup • Visualizing protein homogeneity • Detecting degradation • Estimating molecular weight	• Limited quantification accuracy • Low resolution for similar molecular weight • Cannot detect non-protein impurities
UV-Vis Spectroscopy	Protein concentration quantification	Moderate–High	Proteins with aromatic residues	• Rapid • Low-cost • Widely accessible • Label-free	• Sensitive to pH and buffer • Requires accurate ε • Subject to interference from co-absorbing species
MS	Quantification and purity verification	Very High	Proteins, peptides, small molecules	• Extremely high sensitivity • Detecting trace impurities • Label-free • Supporting internal standards	• Expensive instrumentation • Expert operation required • Potential for ionization artifacts
NMR (qNMR)	Absolute concentration of small molecules	High	Small molecules, metabolites	• High reproducibility • No calibration curve needed • Supporting internal standard-based quantification	• Expensive infrastructure • Large sample requirements • Lower sensitivity than MS

sample quantities and may have lower sensitivity compared to MS, limiting its applicability for low-abundance proteins.[130] Consequently, while both MS and NMR offer reliable protein quantification, their high costs and technical requirements can restrict their widespread use, particularly in resource-limited settings.

As summarized in **Table 1.2** and elaborated in the preceding discussions, each analytical technique for assessing reagent purity and concentration — ranging from HPLC and SDS-PAGE to UV-Vis spectroscopy, MS, and NMR — offers unique advantages and inherent limitations. HPLC remains the benchmark for high-resolution impurity profiling across diverse sample types, while SDS-PAGE serves as a practical tool for visually assessing protein homogeneity and detecting degradation. UV-Vis spectroscopy provides a rapid, cost-effective, and widely accessible means of protein quantification, albeit with susceptibility to buffer conditions and the need for accurate extinction coefficients. MS and NMR techniques offer the highest levels of quantitative accuracy and molecular insight, supported by internal standards and recent technological innovations; however, their broader application is constrained by high costs, complex instrumentation, and technical expertise requirements. Ultimately, no single method is universally optimal. The choice of technique should be guided by the analyte's chemical nature, the required sensitivity, available instrumentation, and the specific goals of the study. In many cases, combining complementary techniques yields the most reliable and comprehensive assessment.

1.4.2.1.2. Use of Standard Reference Materials and Calibration Standards

Certified Reference Materials (CRMs) and calibration standards play a vital role in enhancing the accuracy of reagent concentrations in binding experiments, particularly for proteins, small molecules, and nucleic acids. Inaccuracies in calibration reagents can shift the calibration curve,

leading to systematic errors in all reagent concentrations derived from it.[131] Consequently, accurate quantification of these reagents is essential to ensure reliable determination of binding affinities and kinetics, minimizing errors and increasing confidence in experimental results.

CRMs, produced under strict quality control, provide certified values with traceability, ensuring measurement consistency across different experiments and laboratories.[131] The International Organization for Standardization (ISO) has established guidelines such as ISO 17034:2016 to define the requirements for CRM production, ensuring high-quality materials that minimize systematic errors in concentration measurements.[132]

Despite the benefits of CRMs and standardized reagents, accurate concentration determination also depends on careful sample handling and storage. To ensure the reliability of binding experiments and enhance data precision and accuracy in quantitative affinity studies, researchers should adopt best laboratory practices and effectively utilize certified materials.

1.4.2.1.3. Techniques to Minimize Liquid-handling Errors

Accurate liquid handling, such as pipetting and dilution, is crucial for obtaining reliable analyte concentrations in K_d-determination experiments. Errors in liquid-handling processes can lead to significant deviations in experimental outcomes, reducing the precision and accuracy of determined K_d values. Variability in volume measurements and accumulated dilution errors can introduce biases that impact quantitative binding studies. To minimize these errors, researchers have developed various best practices and systematic approaches aimed at improving liquid-handling precision and accuracy.

Proper Pipetting Techniques: Proper pipetting techniques, such as slow aspiration and dispensing, vertical pipette positioning, and minimal tip immersion, help prevent volume inaccuracies.[133] Pre-wetting the pipette tip stabilizes liquid uptake, improving accuracy,

especially for viscous or volatile solutions.[134] Regular calibration and maintenance, typically every 3 to 6 months or per manufacturer guidelines, are crucial for sustaining pipette performance.[135] Adhering to these best practices minimizes systematic errors, enhances reproducibility, and ensures the accuracy of analyte concentration measurements in experimental studies.

Advancements in pipetting technology have introduced methods like reverse pipetting, which is particularly useful for viscous or foaming liquids.[136] In reverse pipetting, the pipette aspirates a larger volume than needed and dispenses the desired volume, leaving some liquid in the tip to account for any residual liquid that may adhere to the tip walls, thereby improving accuracy.[136] Furthermore, the adoption of electronic and connected pipettes has enhanced precision and accuracy.[136] These devices reduce user-dependent variability by automating pipetting actions and offer features like data logging and protocol integration, which streamline workflows and improve traceability.

Controlling Evaporation Effects: Evaporation poses a significant challenge in maintaining accurate solute concentrations, particularly when handling small volumes or volatile solvents, where even minor solvent loss can result in substantial concentration deviations.[137] To minimize these effects, maintaining stable environmental conditions, including controlled temperature and humidity, is essential, as fluctuations in these parameters have been linked to solvent evaporation rates.[138] Proper sealing of containers, such as using tightly capped vials and low-permeability materials, has also been shown to significantly reduce solvent loss and prevent concentration drift over time, ensuring greater accuracy and reproducibility in experimental measurements.[139]

1.4.2.2. Advanced Statistical Approaches for Improving K_d Accuracy

Accurate determination of K_d requires robust data-analysis methods that incorporate appropriate binding models while accounting for both random and systematic experimental errors. Careful model selection, grounded in a clear understanding of the underlying molecular interaction mechanism, is essential to avoid misinterpretation. Conventional nonlinear regression methods based on the least-squares (LS) algorithm remain widely used but have important limitations. While some implementations allow for the incorporation of systematic uncertainties in input variables, this aspect is frequently overlooked in practical applications. As a result, the fitted parameters may appear precise but lack true accuracy if key sources of experimental error are overlooked.

Standard LS regression typically assumes constant variance across all data points and accounts only for errors in the dependent variable (e.g., fraction bound), ignoring uncertainties in independent variables such as ligand or target concentrations, despite these being common sources of systematic error. Furthermore, if the data do not adequately span the full binding curve, particularly the low- and high-concentration regions, the estimation of K_d becomes poorly constrained. The LS method is also sensitive to initial parameter guesses and susceptible to overfitting, especially when applied to unnecessarily complex models. These factors can lead to biased or unstable parameter estimates and create a false impression of model validity.

To address these pitfalls, experiments should be carefully designed to sample a broad and balanced concentration range. More advanced statistical approaches, such as weighted least squares (WLS), Monte Carlo simulations, and Bayesian inference, explicitly incorporate both random and systematic uncertainties, thereby improving the reliability and interpretability of the results. This section provides an overview of these approaches and outlines best practices for enhancing the accuracy and credibility of K_d determination.

1.4.2.2.1. Weighted Least Squares Fitting

Weighted Least Squares (WLS) fitting has emerged as a statistical approach for improving the accuracy of K_d determination in binding studies, particularly in datasets characterized by varying levels of uncertainty. Unlike ordinary least-squares regression, which assumes homoscedasticity across all measurements, WLS accounts for heteroscedasticity by assigning weights inversely proportional to the variance of each data point.[143] This method reduces the impact of less reliable observations while prioritizing data points with higher precision, leading to more reliable parameter estimations.[144] In all biophysical experiments, where measured signals fluctuate due to experimental conditions and instrumental limitations, WLS provides a more refined approach to data fitting.[145] By effectively incorporating variable noise levels, WLS enhances model robustness, allowing for more precise and accurate determination of K_d values, particularly in weak binding interactions where precision in signal detection is critical.

However, despite its strengths, WLS has inherent limitations that must be considered. Its efficacy depends on the accurate estimation of data variances, which are often unknown or must be approximated, introducing potential sources of error.[146] Incorrectly assigned weights may lead to biased parameter estimates, potentially distorting the fitted model. Furthermore, WLS assumes that data points are independent, an assumption that may not hold in datasets where measurement errors exhibit correlation.[147] In cases where correlated noise is present, alternative statistical approaches, such as Bayesian inference or Monte Carlo simulations, may be required to complement WLS. Despite these challenges, WLS remains a widely accepted and effective method for refining binding affinity measurements, particularly when experimental uncertainties are carefully accounted for and appropriate variance estimation techniques are applied.

1.4.2.2.2. Monte Carlo Simulations

Monte Carlo simulations, particularly bootstrap resampling, have become essential for quantifying uncertainty in biophysical binding experiments by addressing errors in concentration and signal measurements.[148] Unlike analytical error propagation, bootstrap resampling generates empirical probability distributions of K_d by iteratively resampling experimental data, allowing robust estimation of confidence intervals without assuming a predefined error model.[149]

Bootstrap methods are valuable for all binding-affinity studies, where pipetting errors, instrument noise, and reagent variability can introduce systematic biases.[150] By propagating these uncertainties through repeated Monte Carlo simulations, bootstrap resampling provides a more reliable representation of variability in binding affinity estimates compared to conventional nonlinear regression approaches.

While bootstrap resampling is a powerful tool, it also has inherent limitations. It assumes that the resampled dataset sufficiently represents the true underlying distribution, which may not hold for small datasets.[151] Additionally, large-scale resampling is computationally intensive, particularly for complex nonlinear binding models requiring iterative fitting.[151] However, as computational power increases, Monte Carlo-based methods continue to improve the reliability of K_d determination, complementing Bayesian inference and other probabilistic frameworks.

1.4.2.2.3. Bayesian Inference

Bayesian inference is a powerful statistical framework for determining K_d in binding studies, integrating prior knowledge with experimental data to improve parameter estimation and uncertainty quantification.[141, 142] Unlike frequentist methods, which provide point estimates and confidence intervals, Bayesian inference generates a posterior probability distribution for K_d, offering a more comprehensive uncertainty assessment.[152, 153]

This approach is particularly valuable in binding studies with noisy or sparse data, where traditional regression may provide unreliable uncertainty estimates.[153, 154] By incorporating prior information, such as previously measured affinities, Bayesian models enhance robustness against measurement noise and systematic errors.[153] This is especially useful for experiments with limited data points, where frequentist methods may yield unstable estimates.

Bayesian inference relies on Markov Chain Monte Carlo (MCMC) sampling methods, such as Metropolis-Hastings and Hamiltonian Monte Carlo (HMC), to approximate posterior distributions efficiently.[155, 156] Compared to conventional nonlinear regression, Bayesian methods provide richer insights by generating full probability distributions rather than single-point estimates.

However, Bayesian inference has challenges, including the computational cost of MCMC sampling and the sensitivity of results to prior selection.[157] Rigorous convergence diagnostics, such as the Gelman-Rubin statistic, are necessary to ensure reliable posterior estimates.[158]

With ongoing advancements in computational methods, Bayesian inference is becoming an essential tool for quantitative affinity studies, complementing Monte Carlo-based techniques like bootstrap resampling to enhance the accuracy and reproducibility of K_d determinations.

Table 1.3. Comparison of data-analysis strategies in K_d determination.

Approach	Ability to Handle Systematic Errors	Requires Prior Info	Computational Cost	Accuracy Output	Interpretability
Nonlinear Regression	No	No	Low	Point estimate	High
Weighted Least Squares	Partially	No	Low	Point estimate	High
Monte Carlo Resampling	Yes	No	Moderate	Confidence intervals	Medium
Bayesian Inference	Yes	Yes	High	Posterior distribution	Medium to Low

Table 1.3 provides a comparative summary of statistical approaches in K_d-determination data analysis, highlighting their capabilities in handling systematic errors, reliance on prior information, computational demands, types of output, and ease of interpretability. As discussed at the beginning of this section, conventional nonlinear regression methods offer simplicity and high interpretability but lack the ability to account for experimental uncertainties. WLS improves accuracy by incorporating variance-based weighting, though it depends on reliable error estimates. Monte Carlo resampling methods, such as bootstrap, enhance robustness by empirically capturing variability through repeated simulations, offering more realistic confidence intervals at a moderate computational cost. Bayesian inference provides the most comprehensive uncertainty quantification via posterior distributions but requires prior information and significant computational resources. Together, these approaches form a toolkit that researchers can tailor based on data quality, computational resources, and the desired balance between rigor and interpretability in K_d determination.

1.4.2.3. Universal Strategies for Improving Signal Accuracy

Accurate signal measurement is fundamental to the accurate determination of K_d in molecular binding studies. Variability in detected signals due to instrumental drift, background noise, or inconsistent normalization can introduce significant random and/or systematic errors in data analysis, ultimately affecting the reliability of K_d determination. To mitigate these issues, universally accepted strategies such as rigorous instrument calibration and baseline subtraction have been widely implemented.

1.4.2.3.1. Instrument Calibration

Implementing stringent instrument calibration protocols is crucial for minimizing systematic errors in K_d determinations. Inadequate or irregular calibration can lead to erroneous signal

detection and instrument-specific biases, resulting in significant deviations in experimentally determined K_d values.[15] Therefore, rigorous calibration procedures are essential to ensure the reliability and accuracy of K_d measurements.

For instance, in fluorescence-based techniques, improper calibration of fluorophore quantum yield can significantly skew determined K_d values.[159] Regular calibration of excitation sources and detectors ensures consistent signal detection, reducing measurement variability.[160] In SPR experiments, improper instrument calibration can lead to baseline instability and erroneous refractive index measurements.[161] Researchers address these issues by using calibration standards to construct calibration curves and correct for bulk refractive index differences, along with routine instrument maintenance and validation using standard samples.[162]

Enhancing the accuracy of K_d determination requires the systematic implementation of rigorous instrument calibration protocols, standardized measurement procedures, and comprehensive documentation of calibration parameters.[163, 164] Routine calibration of experimental instruments is a universal requirement to minimize systematic errors that may compromise the reliability of binding measurements. Standardized pre-experiment instrument verification and cross-validation using known reference standards have been widely adopted to improve data consistency and facilitate comparability across different studies.[165] Additionally, transparent and detailed reporting of calibration methodologies in the literature allows for more accurate assessment and replication of K_d measurements across laboratories. Emphasizing stringent calibration practices and proactively addressing potential sources of systematic error contributes to improved measurement accuracy and reproducibility, ultimately enhancing the reliability of K_d determinations in molecular interaction studies.

1.4.2.3.2. Compensation for Baseline and Background Noise

Accurate K_d determination requires an accurate distinction between true binding signals and baseline drift or background noise, making effective compensation techniques a critical requirement across biophysical methods. Uncorrected baseline shifts and background fluctuations can obscure real binding events, leading to systematic and random errors in K_d determination. To address this, researchers have developed signal correction approaches tailored to different experimental platforms, ensuring more reliable data interpretation.

For instance, in fluorescence-based assays, baseline subtraction methods, such as ratio-metric correction and reference dye normalization, are widely employed to account for fluctuations in excitation intensity and detector sensitivity.[166] Similarly, in SPR, reference channel subtraction is a standard practice to eliminate non-specific refractive index changes, allowing for more accurate kinetic and affinity calculations.[167] ITC benefits from rigorous negative controls and automated baseline correction algorithms that minimize heat signal drift, improving the accuracy of enthalpy quantification in binding interactions.[168]

These advancements underscore the necessity of baseline and noise correction in all K_d-determination techniques. Whether through mathematical modeling, hardware optimization, or experimental controls, compensating for unwanted signal variations enhances the reproducibility and accuracy of affinity measurements. By implementing rigorous correction protocols, researchers can reduce the systematic errors in detected signals and improve the accuracy of obtained K_d values.

1.5. Research Gaps in Maximizing and Assessing the Accuracy of Experimentally Determined K_d

Despite extensive advancements in K_d-determination techniques, inherent limitations remain across all methodologies, as each technique introduces specific sources of systematic errors in

determined K_d values. For instance, SPR and BLI offer real-time, label-free measurements, but rely on surface immobilization, which may induce steric constraints, mass transport limitations, or non-specific interactions that affect the accuracy of binding kinetics.[50, 169] ITC, a widely used label- and immobilization-free method, measures heat exchange upon binding but is constrained by its high sample consumption, low sensitivity for weak interactions, and sensitivity to buffer mismatches that can introduce background heat variations.[53] Other emerging techniques also exhibit method-dependent challenges. MST, while requiring minimal sample volumes and being compatible with complex biological matrices, often relies on fluorescence detection, necessitating labeling, which can alter the native binding properties of molecules.[66] Additionally, MST measurements involve the introduction of a temperature gradient using an infrared laser, which can influence molecular interactions and introduce systematic errors in K_d determination due to temperature-dependent changes in binding affinity.[67] CE-based methods, such as ACE and NECEEM, provide high-resolution binding analysis in solution but require stringent electrophoretic conditions to maintain resolution and prevent excessive Joule heating.[170] These constraints can cause deviations from physiological conditions, potentially leading to large discrepancies between CE-derived K_d values and those relevant to pharmaceutical applications. NMR-based techniques allow direct, immobilization-free measurement of molecular interactions but often require isotopic labeling for enhanced sensitivity and are limited by large sample requirements and spectral overlap in complex macromolecules.[130]

Recognizing the fundamental limitations of existing K_d-determination methods, Krylov Lab introduced Accurate Constant *via* Transient Incomplete Separation (ACTIS), the first inherently accurate approach for K_d measurement.[171] ACTIS leverages transient incomplete separation (TIS) in a capillary under pressure-driven laminar flow, which occurs when two species differ in

diffusion coefficients. By integrating MS as a detection method, ACTIS enables true label-free and immobilization-free analysis while maintaining low sample consumption, similar to MST due to its capillary-based setup.[171] Unlike conventional CE-based methods, ACTIS operates under pressure-driven conditions, eliminating buffer restrictions and concerns about temperature fluctuations caused by Joule heating. More importantly, unlike traditional approaches that rely on reference reactions, ACTIS is a deterministic method whose accuracy and robustness can be rigorously validated through computational modeling based on partial differential equations.[171, 172]

The robustness of the virtual ACTIS instrument has previously been validated through *in-silico* studies;[172] however, the robustness and ruggedness of the physical ACTIS instrumentation have not been evaluated. Furthermore, past applications of TIS have primarily focused on analytes exhibiting diffusion coefficients differing by at least one order of magnitude.[173, 174, 175] Consequently, ACTIS has predominantly demonstrated accurate K_d determination for protein–small-molecule interactions, where substantial diffusion differences naturally exist.[171] Therefore, verifying the robustness and ruggedness of the physical ACTIS instrumentation and extending its applicability to molecular systems with comparable diffusion coefficients are crucial next steps for enhancing the versatility and broader impact of ACTIS in quantitative affinity measurements.

Beyond the method-dependent limitations discussed, a broader challenge in K_d determination is the lack of a systematic framework for identifying and addressing method-independent sources of systematic errors. While various strategies have been proposed to improve K_d accuracy, such as rigorous instrument calibration, baseline correction, and advanced data-fitting techniques, these approaches remain largely empirical. No analytical study has comprehensively

investigated the fundamental sources of systematic errors in K_d determination, leaving existing strategies without a unified, generalizable methodology applicable across different experimental techniques. A rigorous examination of how systematic errors originate and propagate through experimental workflows is necessary to establish effective, method-independent strategies for improving K_d accuracy.

Traditional nonlinear regression remains widely used for K_d determination, but Monte Carlo-based methods offer greater reliability by incorporating experimental uncertainties into the analysis.[140, 141, 142] However, these advanced approaches primarily address random errors in variables such as concentration and signal intensity, while systematic errors, particularly in concentration measurements, are often overlooked. Although some studies have attempted to account for such systematic errors,[141, 176] they generally lack a clear framework for their identification and quantification. Moreover, many of the existing computational methods remain inaccessible to most in the biochemical and biological research community due to their reliance on complex programming or high computational demands.[140, 141, 142] This limits their practical utility for experimentalists with limited computational expertise. The absence of a user-friendly, computationally efficient tool that integrates systematic error analysis into K_d determination has hindered the broader adoption of rigorous data analysis practices in biophysical and biochemical research. To address this gap, there is a pressing need for an accessible tool that not only evaluates the accuracy of experimentally determined K_d values with consideration of systematic uncertainties in experimental variables but also offers practical guidance for identifying and quantifying these uncertainties.

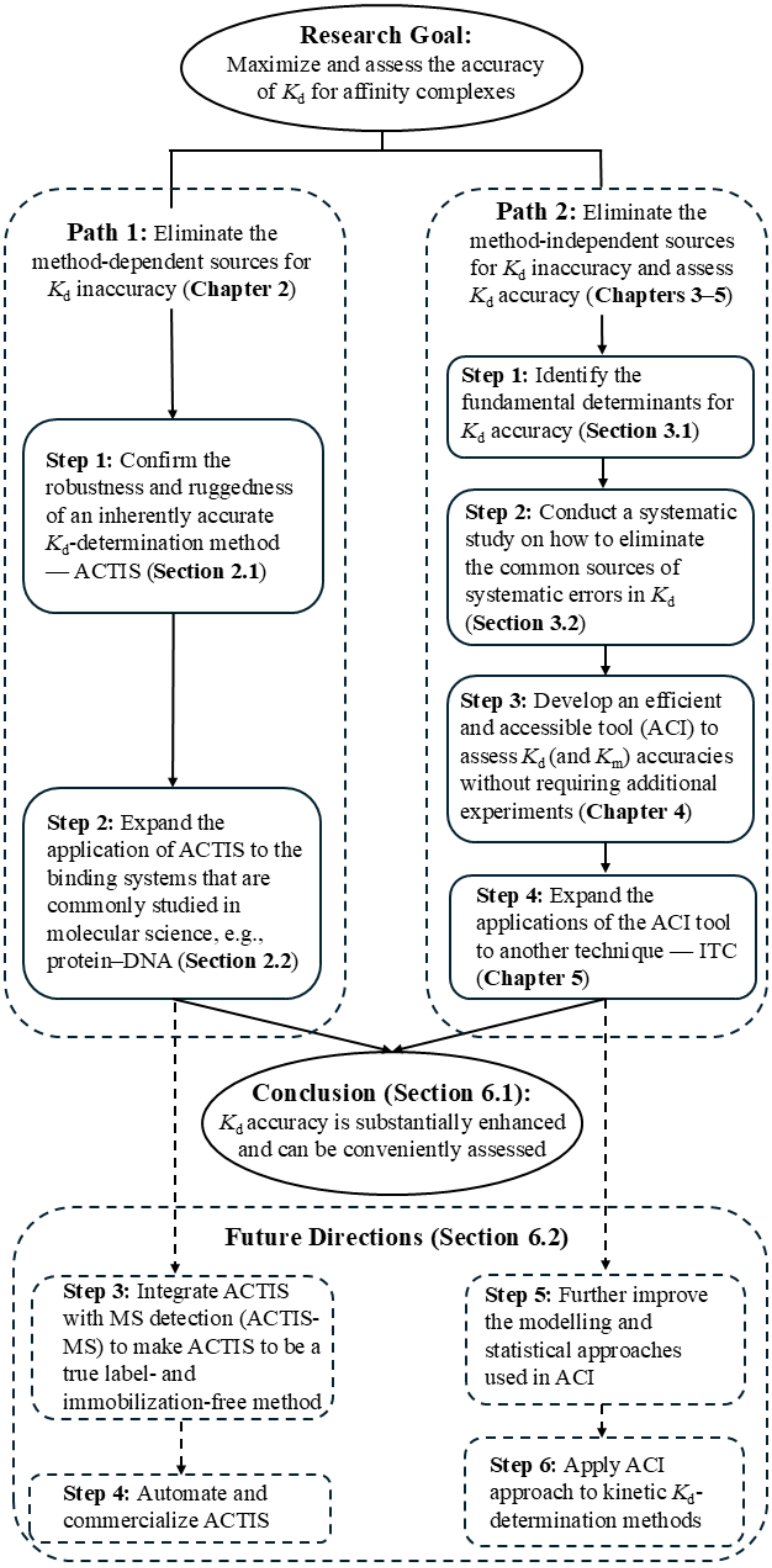


Figure 1.1. Flowchart outlining the structure and key components of the dissertation.

1.6. Dissertation Overview

This dissertation addresses the longstanding research gaps identified in **Section 1.5** by advancing both the experimental methodologies and theoretical frameworks essential for maximizing and assessing the accuracy of K_d. The work is structured along two primary paths: (*i*) eliminating method-dependent sources of error, and (*ii*) addressing method-independent sources of error and introducing an effective approach to assessing K_d accuracy, as outlined in **Figure 1.1.** In the following, I describe the key milestones achieved in each area, indicate the corresponding chapters and sections where these developments are detailed, summarize the main conclusions, and discuss potential future directions stemming from this research.

1.6.1. Path 1: Eliminating the Method-Dependent Sources for K_d Inaccuracy (Chapter 2):

This path focuses on eliminating inaccuracy caused by method-dependent artifacts such as labeling, immobilization, and surface effects. To address these issues, this work builds on ACTIS — a label-free, potentially immobilization-free method grounded in transient incomplete separation (TIS) theory. Two key objectives are pursued: (*i*) experimentally validating the robustness and ruggedness of ACTIS instrumentation, and (*ii*) expanding its applicability to molecular systems with minimal diffusivity differences. These efforts establish ACTIS as a practical and reliable technique for accurate K_d determination across a wider range of binding pairs.

1.6.1.1. Step 1: Confirming the Robustness and Ruggedness of ACTIS (Section 2.1)

ACTIS is a theoretically rigorous Transient Incomplete Separation (TIS) based method for determining K_d, and its accuracy and robustness had previously been validated *in silico*.[171, 172] To

promote broader application, it was necessary to experimentally confirm the robustness and ruggedness of ACTIS instrumentation. To achieve this, two physical ACTIS instruments were assembled and employed simultaneously by two different experimenters to determine K_d values for identical binding pairs over multiple days.[177] The results confirmed the robustness and ruggedness of ACTIS instrumentation, as consistent K_d values were obtained across different instruments and operators.[177]

1.6.1.2. Step 2: Expanding the Application of ACTIS (Section 2.2)

It was widely assumed that TIS-based methods were only applicable to particles with large differences in diffusivity,[173, 174, 175] which consequently restricted early ACTIS research primarily to protein–small-molecule complexes.[171] In this work, through intensive instrumental and experimental optimizations, I successfully expanded ACTIS applicability to accurately determine K_d values for molecular pairs whose free ligand and complex have close diffusivities.[178] The accuracy of these measurements was validated by comparison with a well-established reference method.[178] These findings significantly broaden the potential of ACTIS for practical applications in molecular science research.

1.6.2. Path 2: Eliminating the Method-Independent Sources for K_d Inaccuracy and Assessing K_d Accuracy (Chapters 3–5):

This pathway addresses method-independent sources of systematic error, such as inaccuracies in concentration measurements and signal quantification, which can distort determined K_d values even in ideal experimental setups. Through error identification, systematic evaluation of practical strategies, and dedicated tool development, this work finds the fundamental determinants of K_d accuracy and introduces the Accuracy Confidence Interval (ACI) framework as a quantitative means to assess and report it. This section details the derivation of error-

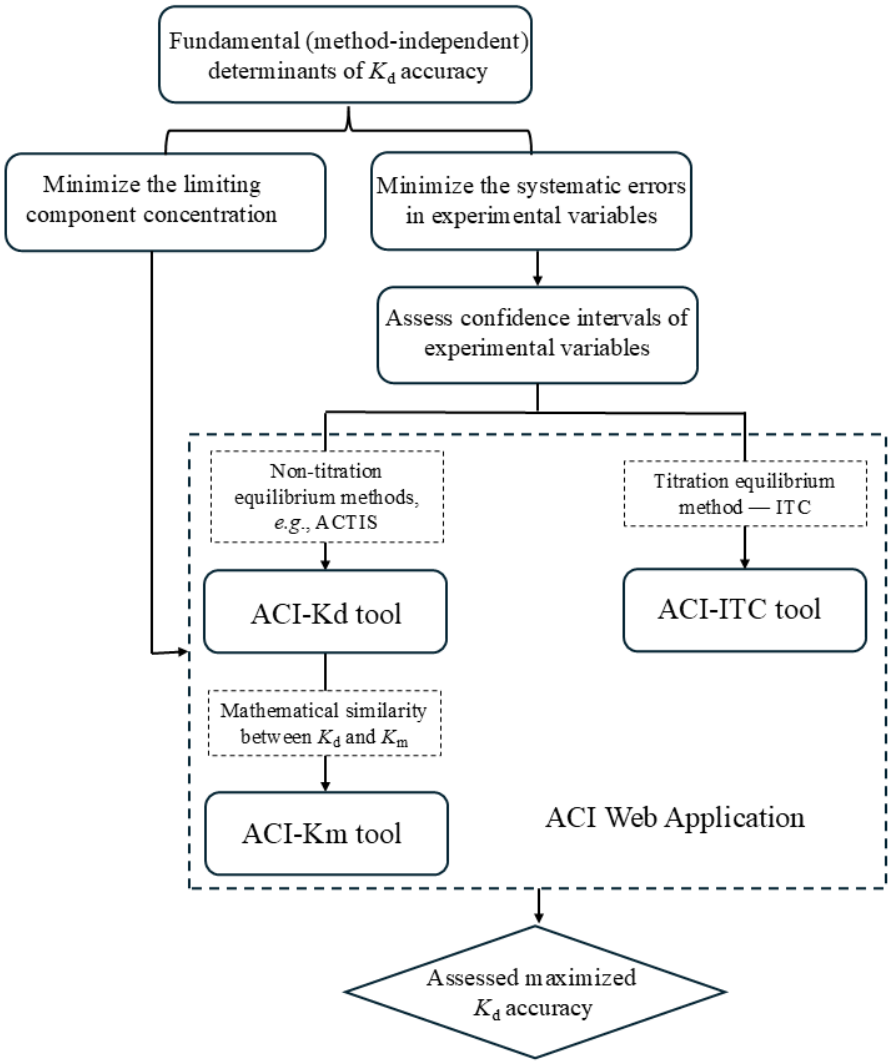


Figure 1.2. Method-independent conceptual framework outlining strategies to maximize and assess K_d accuracy.

propagation models, practical guidelines for optimizing K_d accuracy, and the development of accessible tools (ACI-Kd, ACI-Km, and ACI-ITC) for evaluating the accuracy of experimentally determined K_d. The conceptual workflow and intended outcomes of this pathway are illustrated in **Figure 1.2**, with detailed descriptions of each step provided in the following subsections.

1.6.2.1. Step 1: Identifying the Fundamental Determinants for K_d Accuracy (Section 3.1)

To enhance the accuracy of experimentally determined K_d values, it is essential first to identify the fundamental factors determining K_d accuracy and quantify the impact of each factor. To achieve this, we conducted a detailed error-propagation analysis, examining the theoretical relationship between K_d and experimental variables such as analyte concentrations and measured signals.[179] Our analysis revealed that maximizing the accuracy of experimentally derived K_d values requires experimentalists to minimize both the ligand concentrations used during the experiments and the systematic errors associated with the measurement of experimental variables.[179]

1.6.2.2. Step 2: Eliminating Common Error Sources to Maximize K_d Accuracy (Section 3.2)

Since the lowest possible ligand concentration in K_d determination experiments is constrained by the instrument's limit of quantitation (LOQ),[179] minimizing systematic errors in experimental variables becomes essential for maximizing K_d accuracy. Although previous comprehensive studies exist,[34, 180, 181] they did not analytically quantify how each common source of systematic error propagates into K_d inaccuracy, nor did they provide complete practical guidance. To address these gaps, we conducted a detailed analytical investigation, deriving explicit solutions that demonstrate how systematic errors from common experimental variables propagate into inaccuracies in determined K_d.[182] This analysis highlights the risks researchers face if potential systematic error sources are overlooked. Finally, building upon established prior research and incorporating our novel insights, we developed a comprehensive and practical checklist to guide researchers toward achieving maximized accuracy in K_d determination.[182] Additionally, we proposed a practical method for assessing confidence intervals for systematic errors in analyte concentrations, a crucial step toward evaluating K_d accuracy.[182]

1.6.2.3. Step 3: Developing an Efficient and Accessible Tool to Assess K_d and K_m Accuracies (Chapter 4)

Several numerical methods exist for determining the confidence interval of K_d based on confidence intervals of concentrations.[140, 141, 142] However, most existing methods only propagate random errors in experimental variables, resulting in confidence intervals that reflect the precision rather than the accuracy of K_d. Additionally, existing computational tools are often "black boxes" for molecular scientists, utilizing complex, difficult-to-interpret models and requiring extensive computational resources and programming expertise.[140, 141, 142] In this research, we combined an analytical error-propagation approach with a numerical regression-stability test (commonly employed experimentally by molecular scientists) to derive the Accuracy Confidence Interval (ACI) of the determined K_d from a single binding isotherm without the need for additional experiments.[10] Based on this methodology, we developed an intuitive, browser-based tool — ACI-Kd — that enables researchers to calculate the ACI of K_d quickly and easily, without programming experience or software installation.[10] Additionally, we provided practical guidance on determining critical input parameters, specifically the confidence intervals of systematic errors in experimental variables.[10]

Recognizing the mathematical similarity between K_d and K_m determinations, [183] we demonstrated that our validated approaches for maximizing and assessing K_d accuracy could be directly extended to K_m determinations with appropriate adjustments in terminology. Motivated by this insight and confirmed through numerical validations, we developed a complementary tool — ACI-Km — which enables users to conveniently and efficiently compute the ACI for experimentally determined K_m values. [183]

1.6.2.4. Step 4: Applying ACI Concept to ITC (Chapter 5)

Isothermal titration calorimetry (ITC) is often regarded as a precise and accurate technique for K_d determination due to its label-free and immobilization-free characteristics.[53, 184, 185] To achieve high accuracy, a *c*-value (the ratio of ligand concentration to true K_d value) between 10 and 100 has traditionally been recommended.[186] However, as demonstrated by our analytical findings in **Chapter 3**, this widely accepted *c*-value range may inadvertently amplify systematic errors in experimental variables, resulting in significant inaccuracies in the derived K_d values. To substantiate this concern, we conducted analytical error propagation analyses on the K_d determination process utilized in ITC.[187] Our results confirmed that even within the optimal *c*-value range, ITC-derived K_d values can be substantially affected by systematic errors,[187] highlighting the need for routine accuracy assessment of ITC-derived parameters.

To facilitate this critical evaluation practically and efficiently, we developed a Monte Carlo simulation-based calculator — ACI-ITC — to compute the Accuracy Confidence Interval (ACI) for key ITC-derived parameters, including K_d, enthalpy $\Delta H°$, and stoichiometry n.[188] To encourage widespread adoption and ease of access, this functionality has been incorporated into our existing browser-based ACI toolkit. This advancement can significantly enhance the reliability and interpretability of parameters derived from ITC experiments.

1.6.3. Conclusions and Future Directions (Chapter 6)

In **Chapter 6**, I summarize the key achievements of this dissertation and outline several promising future research directions. These directions include the improvement of ACTIS coupled with mass-spectrometry detection (ACTIS-MS), the automation of ACTIS instrumentation to enhance the usability of ACTIS and reduce human errors, improvements in the mathematical modeling and statistical methodologies underlying ACI calculations, and the extension of ACI calculations to kinetic measurement techniques. Pursuing these avenues will

further strengthen the accuracy, efficiency, and versatility of K_d determination methods, ultimately benefiting broader applications across molecular science research.

Chapter 2. Eliminating the Method-Dependent Sources for K_d Inaccuracy

2.1. Confirming the Robustness and Ruggedness of an Inherently Accurate K_d-Determination Method: Accurate Constant via Transient Incomplete Separation

A portion of the following material was previously published and is reprinted with permission from "Rukundo, J.-L.; Kochmann, S.; **Wang, T. Y.**; Ivanov, N. A.; Le Blanc, J. Y.; Gorin, B. I.; Krylov, S. N. Template Instrumentation for 'Accurate Constant via Transient Incomplete Separation'. *Analytical Chemistry* **2021**, *93* (34), 11654–11659." Copyright 2021 American Chemical Society. My contributions to the article included: (*i*) conducting a portion of the ACTIS experiments, (*ii*) assisting in analyzing and interpreting the results, and (*iii*) assisting in the preparation of figures and the writing of the manuscript. Co-author Jean-Luc Rukundo assembled the ACTIS instrument and conducted a portion of the ACTIS experiments.

2.1.1. Introduction: The ACTIS Concept and Its Significance in Accurate K_d Determination

As discussed in **Sections 1.4.1** and **1.5**, inherent limitations persist across most existing methodologies, as each technique introduces specific sources of systematic errors in K_d determination. Established K_d-determination methods fall into five major categories: spectroscopic, biosensoric, calorimetric, thermophoretic, and electrophoretic.[12, 189, 190, 191] Each category has distinct sources of inaccuracy, often leading to discrepancies in results obtained using different methods for the same complex. In some cases, these quantitative inconsistencies

are so pronounced that they lead to conflicting qualitative conclusions, where one group reports binding, while another claims no binding at all.[32]

The roots of the inaccurate K_d problem are not technical but fundamental. Unlike common physical values, such as mass, time, and length, which have reference standards, there are no reference reactions with reference K_d values. It appears to be an axiom that the lack of a reference reaction forbids creating a reference method for K_d determination. This axiom has been challenged recently by a demonstration that the accuracy of an analytical method can be tested and characterized *in silico* if the method is deterministic. A deterministic method is the one that can be described comprehensively by a system of partial differential equations and, therefore, a virtual instrument can be created, and virtual experiments can be conducted with virtual reference-standard samples.[192, 193]

The first deterministic method for finding K_d values of protein–small-molecule complexes has been recently introduced by Krylov Lab at York University (Toronto, Canada) and termed as "Accurate Constant *via* Transient Incomplete Separation" (ACTIS).[171] ACTIS is based on very fast (typically achieved in less than 1 min) transient incomplete separation (TIS) of the complex from the small molecule. TIS of two species always occurs in a pressure-driven laminar flow inside a capillary if their diffusion coefficients differ.[173, 174, 175] The concept of ACTIS is reiterated in **Figure 2.1**. An equilibrium mixture (EM) of the target and ligand is prepared in an incubation buffer outside of the capillary. The mixture contains T, L, and C at their equilibrium concentrations linked by K_d according to **Eq (1.2)**. A sample plug of this mixture is injected into a capillary prefilled with the pure buffer solution. TIS of C from L is achieved when this plug is propagated inside the capillary by a pressure-driven flow. Such a flow has a parabolic flow-velocity profile (**Figure 2.1a**). The maximum flow velocity, v_{max}, is in the center of the capillary,

the zero velocity is on the capillary walls, and the average flow velocity is $v_{av} = v_{max}/2 = Q/(\pi a^2)$, where Q is a volumetric flow rate and a is the inner radius of the capillary. The interplay

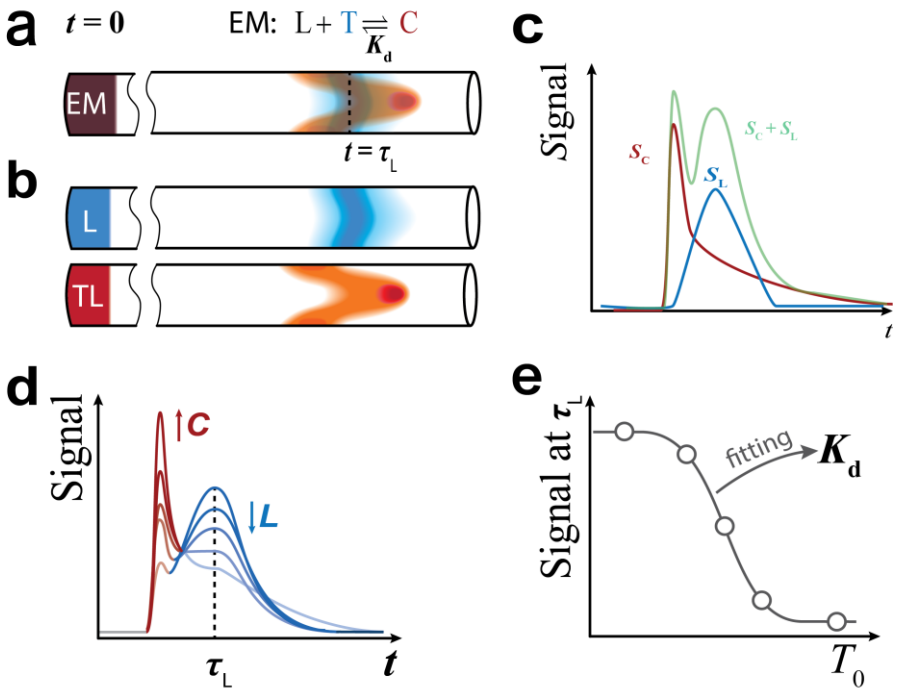


Figure 2.1. Concept of ACTIS **(a)** A sample plug of the equilibrium mixture (EM) of T, C and L is injected in a capillary prefilled with the pure buffer and propagated by a pressure-driven flow. **(b)** The different flow velocities across the capillary and different rates of transverse diffusion of C and L cause the separation of C from L in the longitudinal direction. **(c)** The detector registers both the free ligand and the target-bound ligand (complex) indiscriminately and produces their cumulative signal. The dependence of this signal on time (a separagram) is comprised of two overlapping peaks: a non-diffusive peak of target-bound ligand (complex) and a diffusive peak of free ligand **(d)** A set of separagrams corresponding to different values of T_0 is generated. **(e)** A classic binding isotherm "R vs T_0" is built. The values of the fraction of free ligand, $R = L/L_0$, are calculated from signals at time τ_L taken from the separagrams for different T_0. The K_d value is obtained from fitting the binding isotherm "R vs T_0" with the theoretical dependence of R on T_0, where K_d is the fitting parameter.

between different flow velocities across the capillary and different rates of transverse diffusion of C and L cause the separation of C from L in the longitudinal direction (**Figure 2.1b**). This separation is fast but incomplete: the zones of C and L always overlap. This separation is also transient: it extends to maximum when propagation time reaches the characteristic time of diffusion of ligand (τ_L) from the capillary center to its wall, and it gradually dissipates with longer propagation. If the diffusion coefficient of ligand μ_L is known, τ_L can be calculated as a^2/μ_L. To propagate the plug to the detector at time τ_L, Q is chosen so that v_{av} is linked to the distance l inside the capillary from the point of start of TIS to the detector as $v_{av} = l/\tau_L$. The detector is set to detect the ligand, and it produces a cumulative signal from both target-bound ligand and free ligand. The dependance of this signal on time is called a "separagram", which is comprised of two overlapping peaks: a non-diffusive peak of target-bound ligand complex and a diffusive peak of free ligand (**Figure 2.1c**, left). The detector does not distinguish target-bound ligand from free ligand and, therefore, the separagram is a sum of the diffusive and non-diffusive peaks (**Figure 2.1c**, right). To determine K_d, TIS is performed for a series of equilibrium mixtures with a constant total concentration of ligand, $L_0 = L + C$, and a varying total concentration of target, $T_0 = T + C$. Accordingly, a set of separagrams corresponding to different values of T_0 is generated (**Figure 2.1d**). Signals at time τ_L are taken from these separagrams to calculate fractions of free ligand, $R = L/L_0$, for different T_0. The R values are used to build a classic binding isotherm "R vs T_0" which, in turn, reveals the value of K_d when fitted with the theoretical dependence of R on T_0 (**Figure 2.1e**).[171] Its accuracy has been proven *in silico*, and it has been shown that an accurate ACTIS instrument could be constructed.[172]

In this work, we employed two physical ACTIS instruments designed to replicate the previously reported simple virtual instrument,[172] and demonstrated the stability of their operation.

Each instrument's fluidic system was integrated with a fluorescence detector. Bovine serum albumin (BSA) and fluorescein were selected as the protein–small molecule pair for K_d determination. To evaluate instrumental robustness and ruggedness, we conducted parallel experiments on different days using different operators. The K_d values obtained from both instruments for the same samples deviated by no more than 10%. These results demonstrate that the simple fluidic setup presented here serves as a robust and reliable template for ACTIS instrumentation, supporting its practical application in protein–small molecule binding studies.

Compared to SPR, one of the most widely used techniques for K_d determination, ACTIS offers several important advantages. As discussed in **Section 1.4.1.1**, SPR relies on the immobilization of one binding partner to a sensor chip, which may lead to conformational artifacts, steric hindrance, or nonspecific binding. Despite substantial improvements in surface chemistry and instrumental sensitivity, the need for immobilization remains an intrinsic limitation. In contrast, ACTIS is performed entirely in solution, thereby preserving the native conformations and minimizing experimental artifacts. While SPR excels in throughput and kinetic resolution, it may produce misleading K_d values when mass transport limitations are not properly addressed. ACTIS, by using a deterministic physical model under well-controlled solution-phase conditions, complements SPR by prioritizing accuracy over throughput.

Other common techniques, such as NMR and fluorescence spectroscopy, also have distinct advantages and limitations. As highlighted in **Section 1.4.1.4**, NMR offers atomic-resolution structural insights and is particularly useful for weak or transient interactions. However, its high sample requirements, complexity, and cost can limit accessibility. Fluorescence-based titrations, discussed in **Section 1.4.1.3**, are sensitive and relatively simple to implement but rely heavily on labeling strategies and assumptions about fluorescence response. ACTIS circumvents these

limitations by offering a label-free (when MS detection is applied) and immobilization-free platform, where the readout depends solely on molecular diffusion properties and physical transport.

Thus, ACTIS complements and enhances the current landscape of K_d-determination techniques. Its methodological transparency and solution-phase implementation make it a strong candidate for benchmarking K_d values obtained by SPR, NMR, and spectroscopic approaches, particularly in applications where accuracy and reproducibility are paramount.

2.1.2. Materials and Methods

2.1.2.1. Reagents and Solutions

All reagents were obtained from Sigma-Aldrich (Oakville, Ontario, Canada). The binding pair of a protein and small-molecule ligand used here was bovine serum albumin (BSA, catalog A2514) and fluorescein sodium salt (catalog F2456). New BSA solutions of 10 concentrations were prepared prior to each ACTIS experiment and used to prepare new EMs. The same stock solution of 1 μM fluorescein was used throughout the study. A single buffer, 30 mM ammonium acetate, pH 7.5, was utilized to prepare all solutions and used as the sample propagation buffer; accordingly, we simply refer to it as the buffer. Target−ligand EMs were prepared by mixing appropriate volumes of working solutions of T and L and incubating for 1 h before the start of ACTIS runs (longer incubation was shown not to change the determined value of K_d suggesting that 1 h was sufficient to approach equilibrium in the binding process (**Eq (1.1)**). All EMs had identical total ligand concentration L_0 = 100 nM, unless otherwise stated, while the total target (protein) concentrations in the EMs, T_0, varied within 4 orders of magnitude: from 0.1 μM to 1 mM. All sample handling and measurements were carried out at a room temperature of 20 ±

2 °C. Identical samples were utilized when experiments were conducted with two ACTIS setups in parallel.

2.1.2.2. Determination of K_d

The values of K_d were determined in two steps. First, we built a classic binding isotherm "fraction of free ligand R versus protein (target) concentration T_0" where R is defined as:

$$R = \frac{L}{L+C} = \frac{L}{L_0} \tag{2.1}$$

where L and C are the equilibrium concentrations of ligand of target–ligand complex, respectively, and L_0 is the total concentration of ligand.

Second, we conducted a nonlinear fitting of the experimental dependence of R on T_0 with the theoretical one:

$$R = -\frac{K_d + T_0 - L_0}{2L_0} + \sqrt{\left(\frac{K_d + T_0 - L_0}{2L_0}\right)^2 + \frac{K_d}{L_0}} \tag{2.2}$$

in which K_d is a fitting parameter, into the experimental isotherm.[194] To obtain the R values, we performed TIS on each of the 10 EMs (each with a different total concentration of the protein T_0) by running each EM in the ACTIS instrument. At least three replicated runs were made for each EM, and a signal value for each run was obtained by averaging points within a time window around the diffusive-peak-maximum position. This position corresponds approximately to the characteristic time of transverse diffusion of the protein-unbound ligand.[171] We took a total time window width corresponding to 8% of the diffusive-peak-maximum position, that is, ±4% around this position. The time window used in the determination of signal values was selected from the first EM at T_0 = 0.1 μM and applied subsequently to all other EMs at other T_0 values.

The signal values obtained from the replicated runs were used to find an average signal value, from which the values of R and corresponding standard deviations (obtained through simple error propagation) were calculated using equations described in a previous ACTIS publication.[171] The K_d value and its standard deviation were then obtained by fitting the experimental dependence of R versus T_0 with their theoretical relationship (**Eq (2.2)**) by performing a nonlinear fitting using the Levenberg−Marquardt algorithm in the OriginPro software. The standard deviation of K_d values indicates the goodness of the nonlinear fit.

2.1.3. Results and Discussion

In this work, we assembled two ACTIS instruments and conducted parallel experiments using identical samples, with two operators running the instruments independently. This approach provided a comprehensive evaluation of the robustness and ruggedness of the ACTIS method. The instruments were built of similar components, but no specific effort was undertaken to make the fluidic systems and detectors identical. As a result, the instruments likely differed in their ratios of the capillary length to flow velocity due to the finite tolerance of cutting the capillary and setting the flow rate via a pressure pump. However, ACTIS is robust to variations in this parameter;[172] accordingly, we did not dedicate time to perfect this parameter. The instruments also likely differed in focusing performance of the commercial detectors' optics due to the specifics of ball lenses.[195] These unintentional and hardly avoidable differences cumulatively were the reason for significant differences in the relative heights and positions of the nondiffusive peak (the left-hand side peak) between the two instruments (**Figure 2.2**, left panels). The same differences were observed between the results obtained with these instruments in all the following experiments (**Figures A1.1**−**A1.6** in **Appendix A**). Importantly, these differences in the separagram shapes did not lead to deviations in the K_d values between the two instruments.

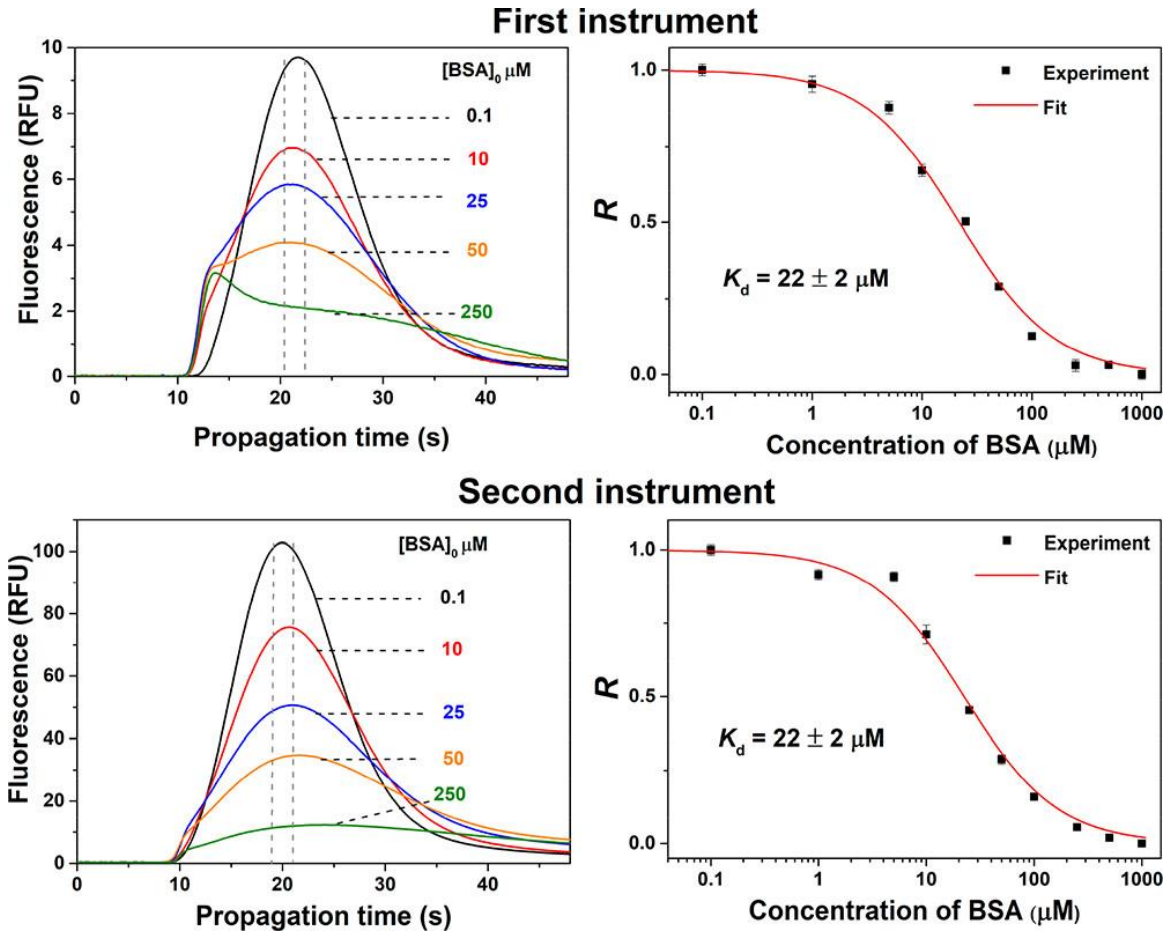


Figure 2.2. Robustness of K_d determination by ACTIS demonstrated with two ACTIS instruments used to run identical samples containing 100 nM fluorescein and varying concentrations of BSA. The left panels show representative separagrams (one of five for each BSA concentration), and the right panels show the corresponding binding isotherms. *R* is a fraction of unbound fluorescein.

All experiments performed in parallel with the two instruments returned K_d values equal within the limits of precision, provided that identical samples were used (**Figure 2.3**). New solutions were prepared every day and day-to-day variations were more significant; the largest K_d value was 28 μM and the smallest value was 20 μM. Our results clearly show the robustness and ruggedness of ACTIS experiments, suggesting that this simple ACTIS instrumentation can support accurate K_d determination.

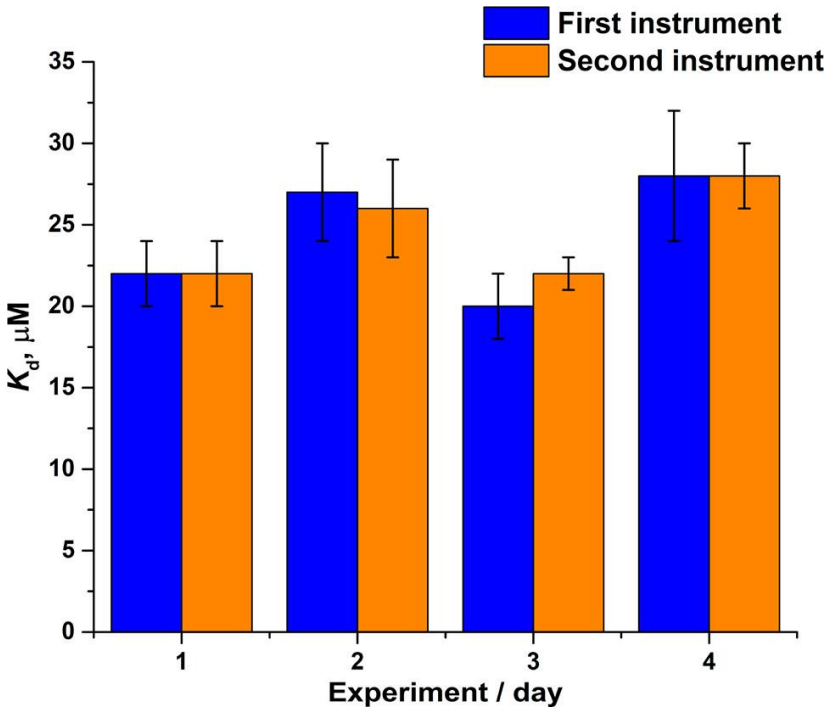


Figure 2.3. Ruggedness of K_d determination with two ACTIS instruments with new samples prepared for each experiment; identical samples were used for each of the two instruments in each experiment.

2.1.4. Concluding Remarks

Accurate determination of K_d remains a significant challenge due to the absence of reference standards and the inherent limitations of existing methodologies, which often introduce systematic errors. To address this, ACTIS provides a fundamentally accurate and deterministic approach that allows for validation *in silico*. In this study, we advanced the validation of ACTIS by experimentally evaluating the robustness and ruggedness of its physical implementation. We constructed two physical ACTIS instruments modeled after a previously validated virtual prototype and conducted parallel K_d-determination experiments using BSA and fluorescein as a model binding pair. Despite minor, unavoidable differences in the fluidic systems and detector optics between the two instruments, consistent K_d values were obtained for identical samples.

These results demonstrate that ACTIS is not only deterministic in theory but also robust and rugged in practice. Its tolerance to variations in experimental setup and operation by different users affirms its potential as a reliable and practical method for accurate K_d determination.

2.2. Expanding the Application of ACTIS: Transient Incomplete Separation of Species with Close Diffusivity to Study Stability of Affinity Complexes

The following material was previously published and is reprinted with permission from "**Wang, T. Y.**; Rukundo, J.-L.; Le, A. T.; Ivanov, N. A.; Le Blanc, J. Y.; Gorin, B. I.; Krylov, S. N. Transient Incomplete Separation of Species with Close Diffusivity to Study the Stability of Affinity Complexes. *Analytical Chemistry* **2022**, *94* (44), 15415–15422." Copyright 2022 American Chemical Society. My contributions to the article were: (*i*) performing all ACTIS experiments, (*ii*) preparing the figures, (*iii*) interpreting the results and (*iv*) writing the manuscript. Co-author Jean Luc Rukundo conducted the COMSOL simulations. Co-author An Le Thi Hoai conducted the NECEEM experiment.

2.2.1. Introduction:

The principle of Accurate Constant *via* Transient Incomplete Separation (ACTIS) is illustrated in **Figure 2.1.** ACTIS is a deterministic method that enables accurate determination of the equilibrium dissociation constant (K_d) for target–ligand interactions. Its theoretical accuracy has been rigorously validated *in silico* through the use of virtual instruments governed by partial differential equations,[171, 172] and its practical reliability has been demonstrated using physical instruments,[177] establishing the method's robustness and ruggedness under real-world experimental conditions.

Despite these advances, the broader application of ACTIS remains limited by a prevailing assumption in the field: that transient incomplete separation (TIS) is only effective when applied to molecular species that differ significantly in size, and therefore in diffusion coefficients. This assumption has been reinforced by all previously reported TIS-based studies, which have focused exclusively on systems in which the diffusion coefficients of the two species differ by a factor of 10 or more.[173, 174, 175] As a result, ACTIS has primarily been applied to protein–small-molecule complexes,[171] where the inherent size disparity between the small ligand and the larger protein–ligand complex naturally supports efficient separation during laminar flow.

However, this long-standing assumption excludes a wide range of biologically and pharmaceutically relevant systems involving larger ligands, such as peptides, oligonucleotides, or proteins, where the diffusion coefficients of the unbound and bound ligands are relatively close. In such systems, the difference in diffusivity may fall well below the conventional threshold presumed necessary for TIS to occur. If ACTIS is to become a broadly applicable analytical tool for binding studies, it is essential to examine whether TIS — and, by extension, ACTIS — can be successfully extended to complexes in which the diffusivity contrast between free and bound ligands is minimal.

The objective of the present study is to explore this open question by investigating the feasibility of applying ACTIS to a protein–large-ligand complex, where the diffusion coefficient difference between the free ligand and the complex is much smaller than in traditional ACTIS applications. Specifically, we aimed to determine whether TIS of such closely diffusing species can still be detected and whether the resulting separagrams provide sufficient resolution to allow accurate determination of K_d. Successful demonstration of this capability would represent a

significant expansion of the ACTIS method's utility, paving the way for its application to a broader range of biomolecular interactions.

2.2.2 Materials and Methods

2.2.2.1. Chemicals, Materials, and Solutions

All chemicals and proteins were purchased from Sigma-Aldrich (Oakville, ON, Canada) unless otherwise stated. Fused-silica capillary with an inner diameter of 100 μm and an outer diameter of 360 μm were purchased from Molex Polymicro (Phoenix, AZ, USA) and used throughout this work. His-tagged recombinant Thermus aquaticus MutS protein (MW = 92.8 kDa) was purchased from Prospec (Ness-Ziona, Israel). An Alexa488-labeled MutS-binding aptamer was synthesized by Integrated DNA Technologies (Coralville, IA, USA) and had the following sequence: 5'-Alexa488 CTT CTG CCC GCC TCC TTC CTG GTA AAG TCA TTA ATA GGT GTG GGG TGC CGG GCA TTT CGG AGA CGA GAT AGG CGG ACA CT-3'.[196] Its stock solution was subjected to annealing by incubating at 90°C for 2 min before cooling it to 20°C at a rate of 0.5°C/s, prior to dilution and preparation of equilibrium mixtures. The running buffer (RB) and sample buffer used in both ACTIS and nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) was 50 mM Tris-Acetate pH 8.2 supplemented with 0.05% Triton™ X100 to reduce adsorption of DNA and MutS protein to surfaces of vials, capillaries, and connectors.

2.2.2.2. COMSOL Virtual Experiments

Computer simulation of ACTIS was conducted with an ACTIS virtual instrument (see **Scheme A2.1** in **Appendix A** for the geometry) built in COMSOL Multiphysics software and described in detail earlier.[172, 177] The dimensions of the virtual components were chosen to be identical to those of the physical ACTIS instrument. The parameters used in the virtual ACTIS

experiments were chosen based on the properties of the molecular pair (protein–aptamer complex and unbound aptamer) to be separated by TIS. Most aptamers have a length of 40 to 80 nucleotides and accordingly have a diffusion coefficient μ_L varying in an approximate range of 50 to 100 $\mu m^2 s^{-1}$.[197, 198] The robustness of ACTIS to separation flow rates suggests ACTIS to be robust to the choice of μ_L.[171, 172] Hence, to simplify the calculations and reduce the computing time, the upper limit value of 100 $\mu m^2/s$ was chosen as μ_L in all virtual ACTIS experiments. The radius a of the injection/separation capillary was set to be 50 μm. The characteristic time of transverse diffusion of the ligand (aptamer) was then $\tau_L = a^2/\mu_L = 25$ s. The distance from the point of start of TIS to the detector was set to be 40 cm. The volumetric flow rate for optimal TIS was then calculated as $Q_{TIS} = \pi\mu_L l \approx 8$ μL/min. All other details on COMSOL settings can be found in Note A2.1.

2.2.2.3. Poly (vinyl alcohol) (PVA)-Coating

All capillaries used in this work were coated with PVA to reduce the adsorption of sample components to their inner walls. The PVA-coating procedure was modified from our previously published work.[199] PVA (5%, w/v) was prepared by dissolving the polymer in boiling deionized water. Each time, two 130-cm-long capillaries were coated simultaneously. The two uncoated fused-silica capillaries were sequentially flushed with 0.1 M NaOH and deionized water for 1 h each under a 6.7 psi flow of nitrogen gas. The pretreated capillaries were then flushed with the PVA solution for 10 min at 8.4 psi and emptied using a 5.6 psi nitrogen gas flow for 10 min. PVA was immobilized on the capillary surface by drying overnight in an oven set at 140 °C and continuously flushed with nitrogen gas at a pressure of 2.8 psi. The detection window on the capillary was made by removing the outer coating with a fuming solution of H_2SO_4.

2.2.2.4. ACTIS Instrumentation

The ACTIS instrumental setup was identical to the previously published one,[177] except for the following modifications (see schematics in **Figure A2.1** in **Appendix A**). First, the capillary for draining waste was replaced by a 7.0-cm-long "sample aspiration capillary". Second, the sample was introduced into the injection loop by suction: the syringe pump was set at "withdraw only mode" to aspire the sample directly from the sample vial that was placed at the end of the sample aspiration capillary. This modification allowed us not to fill syringes with sample and, thus, avoid adsorption of sample components to the inner wall of syringes. Third, the syringe pump was programmed to only be triggered and run during the injection step, allowing for a largely reduced volume of consumed sample. Finally, the main pump was operated in the "flow rate-control" mode facilitated by an in-line flow sensor.

2.2.2.5. ACTIS Experiments

The following changes were made to the previously published ACITS experimental procedure (**Figure A2.2**).[177] For Step 1, the valve was in position I, and the injection loop was filled with the sample using a syringe pump to aspirate the sample at a flow rate of Q_s = 50 µL/min for 15 s. Since there was a delay for the desired volume of sample to flow through the fluidic path, the valve was set to stay at position I for 40 s. Step 2 was not changed. For Step 3, the sample plug was propagated inside the capillary by the main pump at a flow rate of $Q_{TIS} \approx 9$ µL/min for 65 s. For Step 4, the separation capillary was rinsed with RB at a flow rate of $Q_{rinse} \approx 37$ µL/min (the highest flow rate we can reach with the current setup) for 2 min allowing for approximately 12 volumes of the separation capillary to pass through the fluidic system. The four-step experimental run took approximately 4.2 min. For the "between-sample" rinsing step, a 1.5 mL sample vial containing RB was placed at the end of sample aspiration

capillary, and the syringe pump was set to continuous withdrawing at a flow rate of 75 μL/min, and the main pump kept injecting the buffer at a flow rate of 37 μL/min. During the "between-sample" rinsing step, the valve was repeatedly switching between positions I and II for 60 times. For each switch, the valve spent 1 s in position I and 2 s in position II. During this rinsing, the buffer with approximately 37 volumes of the separation capillary passed through the fluidic system. This rinsing step took approximately 3 min. A total of three ACTIS runs were performed for each concentration of the protein. An ACTIS experiment with seven protein concentrations along with all the rinsing steps (excluding the sample preparation procedure) took approximately 3 h. All experiments were conducted at an ambient temperature of 24 ± 1 ºC.

2.2.2.6. Capillary Electrophoresis (CE) Instrumentation

All CE experiments were performed with a P/ACE MDQ apparatus (SCIEX, Concord, Ontario, Canada) equipped with a laser-induced fluorescence (LIF) detection system. Fluorescence was excited with a blue line (488 nm) of a solid-state laser and detected at 520 nm. A 50-cm-long capillary was used (39.8 cm from the inlet to the detector). Prior to each run, the capillary was rinsed with RB at a pressure of 40 psi for 3 min.

2.2.2.7. NECEEM Experiments

Following to the rinsing procedure, the equilibrium mixture was injected into the capillary with a pressure of 0.5 psi for 10 s, which yielded a 1.9-cm-long sample plug. Then, the injected sample plug was propagated through the uncooled region of the capillary at the inlet by injecting a 5.7-cm-long plug of RB with a pressure of 0.3 psi for 50 s. As a result, the total separation distance was 33.2 cm (from the center of the sample plug after its propagation to the detector). Electrophoretic separation was carried out at 25 kV with an anode at the outlet. The capillary-coolant temperature was set to 15 ºC. The duration of electrophoretic separation was 15 min.

2.2.3. Results and Discussion

2.2.3.1. Virtual Experiments for Complexes with Close μ_C and μ_L

As we mentioned, ACTIS is a deterministic method allowing for virtual ACTIS experiments to be conducted *in silico*. Therefore, we used *in-silico* experiments to reasonably predict if ACTIS was suitable for the separation of a protein–aptamer complex from unbound aptamer. The ACTIS virtual instrument built in COMSOL was used to study the influence of the ratio between the diffusion coefficient of the target–ligand complex, μ_C, and μ_L on the extent of TIS. Since the size of the protein–aptamer complex is larger than that of the unbound aptamer, we can assume that $\mu_C < \mu_L$. It is unlikely though that the hydrodynamic radius of the complex is more than twice that of the aptamer. Since the diffusion coefficient depends on the radius reciprocally,[200] the diffusion coefficient of the complex is smaller than that of the aptamer by less than a factor of 2. Hence, we conducted this study for μ_C/μ_L ranging from 0.5 to 1 (with a step size of 0.05). Technically, we ran eleven species with distinguishable diffusion coefficients at identical concentrations in COMSOL.

First, we studied TIS concentration profiles under ideal condition of no noise. The concentration profiles clearly differ (**Figure 2.4a**), suggesting that species with such small differences in diffusion coefficient can be separated by TIS. The dynamic range of the signal (see **Figure A2.2** for a sample calculation of dynamic range) decreases with increasing μ_C/μ_L (**Figure 2.4b**, left axis). This decrease suggests *a priori* that the robustness of ACTIS to any noise will worsen with increasing μ_C/μ_L.

We then ran virtual ACTIS experiments to determine K_d values for the same range of variation of μ_C/μ_L from 0.5 to 1 with a step of 0.05 (see **Figure A2.3**). The dependence of the ratio of determined K_d ($K_{d,det}$) to the input value ($K_{d,inp}$) on μ_C/μ_L is shown in **Figure 2.4b**, right

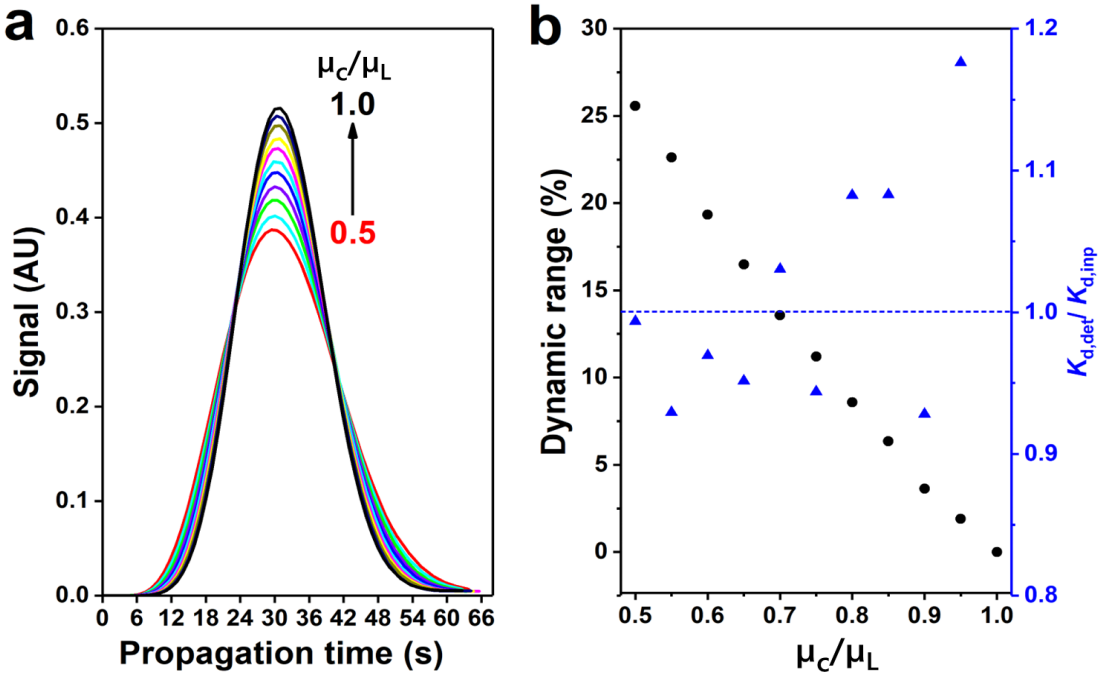


Figure 2.4. COMSOL simulation: the influence of the ratio between diffusion coefficients of the complex and the ligand (μ_C/μ_L) on **(a)** the concentration profiles of C in TIS and **(b)** the dynamic range of the signal (left axis) and accuracy of K_d determination.

axis. According to the results, the percentage difference between the determined and input K_d was within 10% (*i.e.*, $0.9 \leq K_{d,det}/K_{d,inp} \leq 1.1$) for $\mu_C/\mu_L \leq 0.9$ but rose rapidly to 18% with μ_C/μ_L increasing to 0.95. This decrease of the accuracy was associated with the noise in the simulated signal which was caused by the finite mesh size of the iterative simulation. Decreasing the mesh size expectedly resulted in decreasing this deviation (see **Figure A2.4**).

The above simulations only confirmed what was clear for us *a priori*: ACTIS could be used for K_d determination for any finite difference between μ_C and μ_L provided that the conditions are ideal. Real ACTIS experiments inevitably have three sources (or two types) of nonideality: instability in flow rate, instability in detector response (low-frequency noise), and detector's

high-frequency noise.[177] Reliable K_d measurements require that signal variations caused by these nonidealities be much smaller (*e.g.*, $\leq 10\%$) than the dynamic range of the signal. This requirement is a challenge for small differences between diffusion coefficients of C and L. For instance, the signal dynamic range is only approximately 2% for the 5% difference in diffusion coefficients ($\mu_C = 0.95\mu_L$, **Figure 2b**). In this case, reliable K_d measurements would require that all the instabilities and variations be $\leq$ 0.2%, which is hardly achievable. This example emphasizes the need for a highly-stable ACTIS instrument as well as high-quality ACTIS experimentation for target–ligand pairs with small differences between μ_C and μ_L.

Instabilities in a representative physical ACTIS setup have been recently studied in detail.[177] Instrumental low-frequency variations in the flow rate and detector response were found to cause run-to-run signal variation (RSD) of approximately 2%. Relative noise of the detector (high frequency) should be considered separately from these variations, as unlike them, it depends on the signal value, which in turn is inversely proportional to L_0. In general, accurate K_d measurements require that $L_0 < K_d$.[34] Since K_d is unknown *a priori*, it is preferable to use as low as possible L_0. However, since decreasing L_0 leads to decreasing signal-to-noise ratio (*S/N*), L_0 should be chosen such that it balances between measuring low K_d values accurately and maintaining reasonably high *S/N* that is tolerant to experimental errors. In our case, the lowest L_0 that satisfies the above criteria was 0.05 nM, which produced a *S/N* of approximately 1000, which, in turn, results in a high-frequency noise of 0.1%. Since variations from different sources are additive, the total noise caused by the ACTIS instrument was approximately 2%.

In a complete ACTIS experiment with a set of target concentrations T_0, the potentially largest source of sample-to-sample signal variation is the variation of L_0 caused by pipetting errors. This variation can be greatly minimized by using a procedure for the preparation of a series of

equilibrium mixtures described in Note A2.2. We proved experimentally that, in one titration experiment, our procedure of equilibrium-mixture preparation did not contribute detectably to signal variation from all the other sources since the overall RSD of the diffusive peak intensities in the mock titration was approximately 2% (**Figure A2.6**). In the future, the accuracy of sample preparation can be further increased by using an autosampler.[201] To test how the described sources of variations influence the accuracy of K_d, we incorporated the experimental errors that affect an ACTIS experiment performed with a single set of solutions without changing instrumental settings into our virtual ACTIS instrument.

The two types of variations that we added to the virtual instrument were: (*i*) low-frequency noise of random run-to-run variation of the flow rate with RSD of 1%, which caused 2% variation for the signals, and (*ii*) high-frequency detection noise of 0.1% of the signal. We then used this setup to run virtual ACTIS experiments with μ_C/μ_L values varying from 0.5 to 1 with a

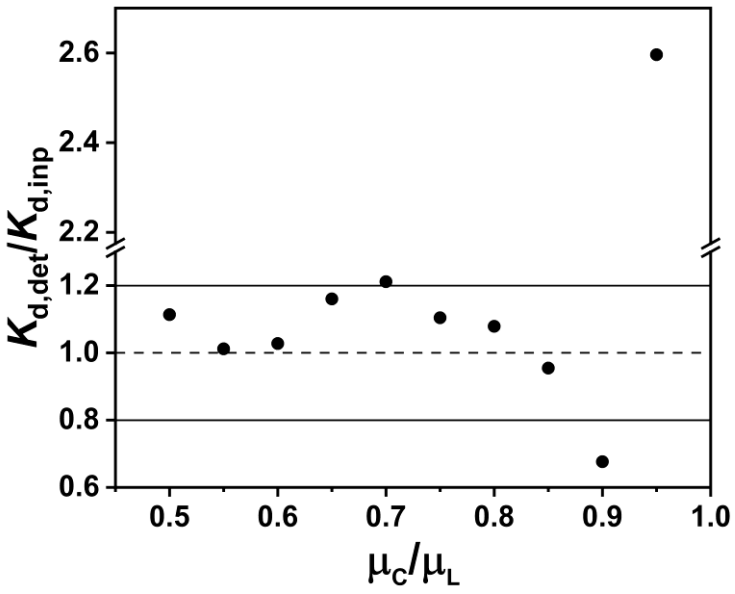


Figure 2.5. The dependence of accuracy of K_d determination in virtual ACTIS experiments (with added variations in key parameters describing the virtual instruments) on the ratio between diffusion coefficient, μ_C/μ_L. The solid lines represent a 20% deviation from the input K_d value.

step of 0.05. The separagrams and binding isotherms can be found in **Figure A2.7**, while **Figure 2.5** shows the deviation of the determined K_d value from the input one as a function of μ_C/μ_L. The accuracy of K_d determination decreases with decreasing difference between μ_L and μ_C.

If we select a 20% error in K_d (*i.e.*, $0.8 \leq K_{d,det}/K_{d,inp} \leq 1.2$) as the highest acceptable deviation, then the smallest acceptable difference between μ_L and μ_C, based on data **Figure 2.5**, is approximately 15%. We can roughly estimate the smallest size of the protein that can support such a difference between μ_L and μ_C since the 15% difference in diffusion coefficients requires 15% difference in the hydrodynamic radii of L and C. Assuming the globular geometry of all species, we can convert the 15% (factor of 1.15) difference in radii into a 50% ($1.15^3 \approx 1.5$) difference in the volume or molecular weight (also assuming the same mass density for all species). An 80-nt aptamer has a molecular weight of approximately 25 kDa, suggesting a protein of roughly 12.5 kDa as the smallest suitable target. Of course, shorter aptamers would allow working with smaller proteins. Overall, the results of our *in-silico* experiments suggest the applicability of ACTIS to K_d determination of protein–aptamer complexes, prompting us to the experimental proof of principle.

2.2.3.2. ACTIS-Based Determination of K_d for MutS–Aptamer Complex

The molecular pair to study experimentally was His-tagged MutS protein (molecular weight of approximately 93 kDa) as a target and a fluorescently-labeled 80-nt ssDNA aptamer selected for a tagless MutS protein as a ligand (the tagless MutS is no longer commercially available).[202] We chose MutS–aptamer as our model pair in this proof-of-principle work for three reasons. First, the MutS–aptamer pair has a diffusion coefficient of less than an order of magnitude and has been previously studied by our lab with a well-established method,[101] which suggested that

we could provide a reference study for this molecular pair. Second, this protein–ligand pair was challenging due to its low K_d value and presented an opportunity to prove the suitability of our sample preparation procedure and the high sensitivity of our instrumental detection system. Third, our preliminary study showed a high level of absorptivity of His-tagged MutS to surfaces (*e.g.*, the inner walls of fused silica capillaries). Working with a highly adsorbing protein gave us a chance to optimize our experimental procedure and instrument to apply ACTIS to a broad range of protein targets in the future.

MutS aptamer was shown to tightly bind the tagless MutS with a K_d = 0.12 nM.[101] Although the affinity of this aptamer to the His-tagged MutS has not been studied, we presumed that the His-tag would not affect the binding much, though we did not know the exact K_d value of the complex of this aptamer with the His-tagged MutS *a priori*. In this case, the aptamer concentration of L_0 = 0.05 nM was used as a default concentration for all experiments since this concentration was supposed to satisfy the condition of $L_0 < K_d$ provided that the tag did not improve the binding much. The molecular weight of the MutS–aptamer complex was calculated to be 118 kDa. Based on the relation between molecular weight and diffusion coefficient mentioned above, μ_C/μ_L was predicted to be approximately 0.6 ($\sqrt[3]{25}/\sqrt[3]{118}$).

Our first experimental goal was to test if TIS of the MutS–aptamer complex from the unbound aptamer could be detected. For this purpose, in the positive-control experiment, we chose the concentration of MutS of 10 nM, which is much greater than the expected K_d value of 0.12 nM and should guarantee that $C >> L$ in the equilibrium mixture. We thus used an equilibrium mixture of 10 nM MutS and 0.05 nM aptamer for this TIS assessment (**Figure 2.6a**). In a negative-control experiment, we used an equilibrium mixture of 10 nM BSA and 0.05 nM aptamer (**Figure 2.6b**). Every sample was run in triplicate. Adding MutS to the aptamer resulted

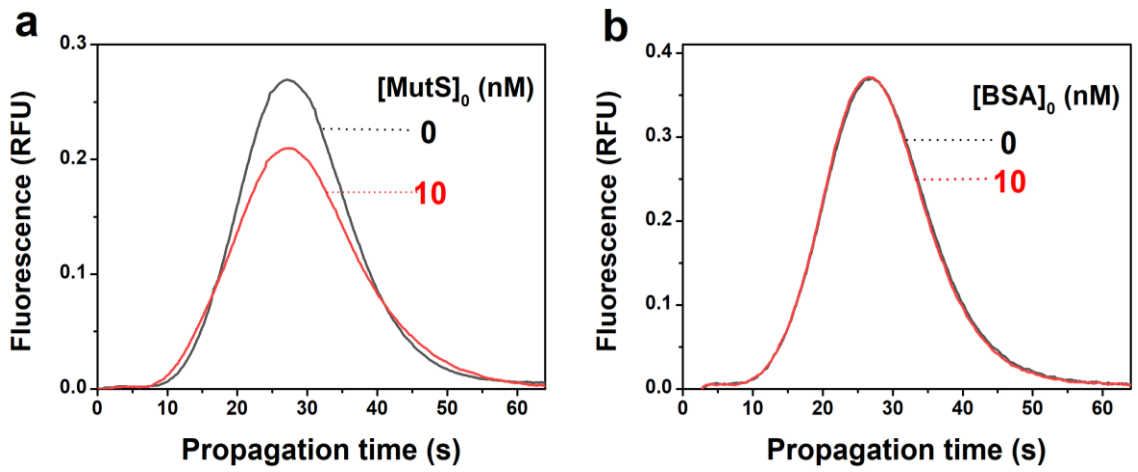


Figure 2.6. Positive control **(a)** and negative control **(b)** for the TIS of MutS–aptamer complex from the unbound aptamer. In the positive control experiment, the aptamer concentration was kept at 0.05 nM, and the concentrations of added MutS protein were 0 and 10 nM. In the negative control experiment, the aptamer concentration was kept at 0.05 nM, and the concentrations of added BSA protein were 0 and 10 nM.

in the decrease of the diffusive peak while adding BSA to the aptamer did not change the separagram. It is instructive to note that TIS resulted in a noticeable decrease of the diffusive peak without producing an apparent non-diffusive peak, which was consistent with the prediction in the virtual experiments (see **Figure A2.7**, for example). The ratio between the overall noise (*i.e.*, 2%) and the dynamic range of the signal (*i.e.*, ~ 20%) was approximately 10%, which established a good ground for ACTIS-based determination of K_d.

In general, building a binding isotherm for accurate K_d determination requires that the range of target concentrations extends from much below to much higher than the K_d value. As mentioned above, our expectation was that the K_d for our complex between the aptamer and His-tagged MutS would not differ much from K_d = 0.12 nM determined for the complex of this aptamer with the tagless MutS. Accordingly, we chose the range of target concentration in the equilibrium mixtures to be three orders of magnitude from 0.01 to 10 nM, and the constant

concentration of the aptamer was 0.05 nM in all mixtures. Every equilibrium mixture was run in triplicate. **Figure 2.7** shows representative separagrams at different protein concentrations (panel **a**) and the binding isotherm (panel **b**). The resulting value of K_d was 0.17 ± 0.02 nM. This experiment was repeated another time on a different day with new equilibrium mixtures (prepared with the same stock protein solution that was used in the first day's experiment). The resulting K_d value was 0.19 ± 0.05 nM, which was equal to the first value within the limits of experimental error.

The comparison of separagrams obtained in virtual TIS for varying μ_C/μ_L (**Figure 2.4a**) with separagrams obtained in real TIS for saturating target concentrations (*e.g.*, for $[MutS]_0$ = 10 nM in **Figure 2.7**) allows us to estimate μ_C/μ_L for the MutS–aptamer pair. The details of this comparison can be found in **Figure A2.11**. By comparing the dynamic ranges obtained from virtual and real experiments, the μ_C/μ_L value for the MutS–aptamer pair was found to be 0.6,

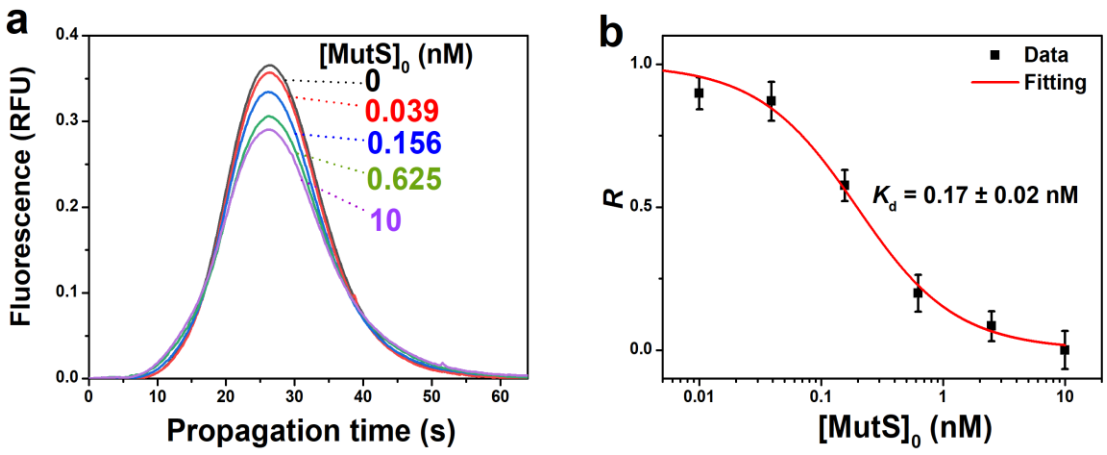


Figure 2.7. ACTIS-based K_d determination for the MutS–aptamer complex: **(a)** representative separagrams at different protein concentrations**; (b)** binding isotherm and curve fitting (using nonlinear regression) to determine K_d**.** The aptamer (ligand) concentration was kept at 0.05 nM, and the MutS (target) concentration was varied from 0 to 10 nM.

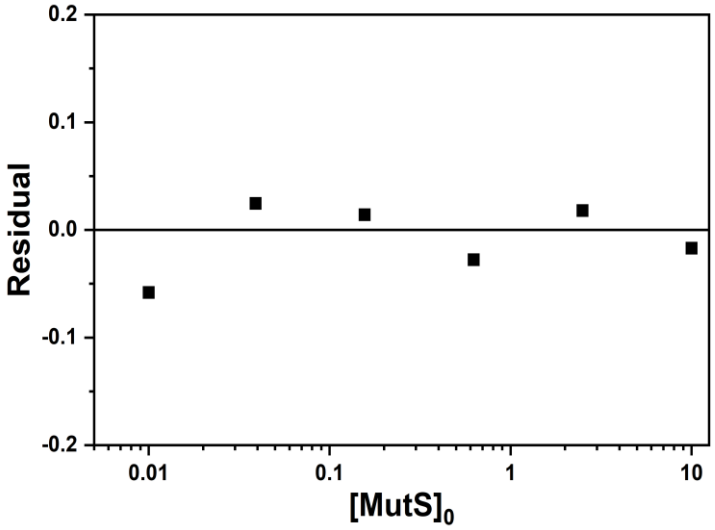


Figure 2.8. Residuals from fitting the 1:1 binding model plotted against total target concentration $[MutS]_0$. The random scatter around zero with no systematic pattern supports the validity of the model and indicates no major misspecification.

which was consistent with our prediction.

To validate the suitability of the fitting model (Eq (6) in Ref. [171]) used for the 1:1 binding process (**Eq (1.1)**), we examined the residuals by plotting them against the total target concentration $[MutS]_0$. As shown in **Figure 2.8**, the residuals are randomly distributed around zero with no apparent systematic patterns, indicating that the model accurately captures the binding behavior. This suggests that any inaccuracy in the determined K_d likely arises from experimental errors, such as systematic errors in reagent concentrations or signal measurements, rather than from model misspecification. While this validation approach is applicable to all binding pairs studied in this dissertation, we present this example for illustrative purposes, as the 1:1 binding model for each pair has already been independently confirmed using orthogonal techniques such as CE and ITC.

2.2.3.3. Comparison of ACTIS with NECEEM

While ACTIS has been proven to be intrinsically accurate, we decided to confirm its result by using another solution-based method: NECEEM.[202, 203] NECEEM is a well-established method with well understood limitations.[101] The quality of separation in NECEEM is, in general, better than in ACTIS (NECEEM can provide baseline separation of the unbound aptamer from the protein bound aptamer). On the downside, NECEEM separation is much longer (~ 10 min) than ACTIS separation (< 1 min), and accordingly much more prone to adsorption of T, L, and C onto the capillary walls. In addition, NECEEM is incompatible with electrophoresis-unfriendly buffers,[203, 204, 205] while ACTIS separation does not depend on the ionic composition of the buffer. Another limitation of NECEEM is that electrophoresis generates Joule heat which makes temperature control an issue.[206, 207, 208] ACTIS is a pressure-driven technique with no heat generation in the capillary.

To make the results of ACTIS and NECEEM experiments fully comparable, we conducted NECEEM measurements in parallel with ACTIS measurements using the same equilibrium mixtures. A new stock of MutS solution (different from the one used in the previous 2-day

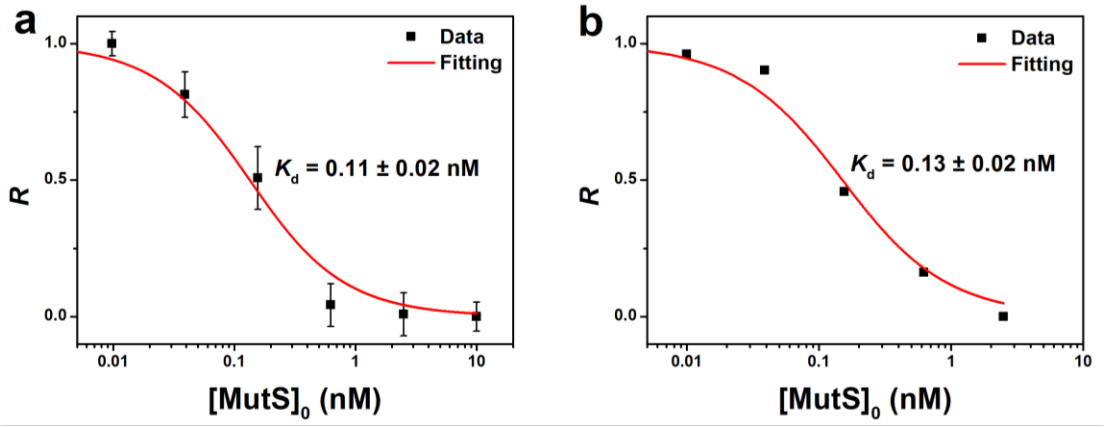


Figure 2.9. Parallel determination of K_d of MutS–aptamer complex using ACTIS **(a)** and NECEEM **(b)**. Binding isotherms and their best nonlinear-regression fits are shown.

experiment) was used in the sample preparation. As NECEEM separation is much longer than ACTIS separation, we only conducted one run of NECEEM for each equilibrium mixture in contrast to three ACTIS runs for each equilibrium mixture. The resulting binding isotherms are shown in **Figure 2.9**. The K_d values obtained were 0.11 ± 0.02 nM with ACTIS and 0.13 ± 0.02 nM with NECEEM. These two values were considered as consistent within the uncertainty range, which confirmed the accuracy of ACTIS in K_d determination for protein–aptamer complex. However, these results differed beyond experimental errors from the ACTIS results obtained in the previous 2-day experiment. We ascribe this inconsistency to hardly avoidable variations in sample preparation. For instance, compared to an older protein stock solution, the newer protein stock solution could have a higher concentration of active protein, which could lead to a lower determined K_d. In addition, at low protein concentration (*i.e.*, < 10 nM), the effects from pipetting errors and protein adsorption are always non-negligible.

According to the results obtained in ACTIS experiments performed on 3 separate days, the weighted average K_d value for His-tagged MutS–aptamer was calculated to be 0.14 ± 0.01 nM (Note A2.3), which was close to K_d = 0.12 nM determined with the tagless MutS protein. This similarity in K_d values for tagged and tagless MutS confirmed that the His tag did not affect the aptamer-binding domain of the protein.

It is important to note that the total areas under the curves in NECEEM experiments with different MutS concentrations (and thus different C/L ratios) were identical within the limits of 5% error without any apparent trend (**Table A2.1**). The constancy of the areas indicates that the fluorescence label on the aptamer does not change its quantum yield upon aptamer's binding to MutS. This, in turn, proves that the signal decrease observed in our ACTIS experiments for MutS–aptamer (**Figure 2.6a**, for example) was caused by TIS instead of quenching. Overall, the

above experimental results are the first and convincing demonstration that ACTIS can be used for K_d determination of protein–aptamer complexes.

(As one of the reviewers requested, we applied ACTIS to measure K_d of another molecular pair with single-stranded DNA binding protein (SSB, MW = 76 kDa) as the target and a fluorescently-labeled 22-nt long ssDNA as a ligand. The K_d values of SSB–22-nt ssDNA were determined to be 16 ± 4 nM and 15 ± 4 nM in two different days (see the representative separagrams and binding isotherms in **Figure A2.15**). The consistency of the determined K_d values confirmed the capability of ACTIS to determine K_d values of complexes with small μ_L/μ_C.)

2.2.4. Concluding Remarks

To conclude, we proved that, counterintuitively, a small difference between diffusion coefficients of protein–aptamer complexes and unbound aptamers can result in sufficient TIS to determine K_d accurately. This accuracy was confirmed by the well-studied solution-based method of NECEEM. Method performance in ACTIS will improve with decreasing μ_C/μ_L, *i.e.*, with decreasing size of an aptamer and/or increasing size of the protein. Rough estimates suggest a lower limit of the protein size to be 12.5 kDa for an 80-nt aptamer. In this proof-of-concept study, we used His-tagged MutS (~ 93 kDa) and 80-nt long aptamer. This proof-of-principle work demonstrates that ACTIS with fluorescence detection applies to measuring K_d for protein–aptamer pairs accurately and, due to its accuracy, is a potential candidate for a reference K_d-determination method for such binding pairs. ACTIS can potentially be used as a reference method for all molecular complexes (*e.g.*, protein–protein), though, every new application needs to be proven experimentally.

It is worth noting that in a 3-day ACTIS experiment, the weighted average K_d value for the MutS–aptamer complex was determined to be 0.14 ± 0.01 nM (Note A2.3), consistent with previous measurements.[101] However, despite using the same method (ACTIS) for the same binding pair, the K_d values varied across different days (**Figure A2.8**–**A2.10**). This day-to-day variation likely stems from systematic errors introduced during sample preparation and signal measurement. These findings highlight the limitations of relying solely on methodology improvement and underscore the importance of investigating the fundamental, method-independent determinants of K_d accuracy. They also point to the need for developing practical strategies to minimize the influence of these underlying factors on K_d determination.

Chapter 3. Fundamental Determinants of K_d Accuracy and Strategies to Mitigate Their Impact

3.1. Fundamental Determinants of K_d Accuracy

The following material was previously published and is reprinted with permission from "**Wang, T. Y.**; Ji, H.; Everton, D.; Le, A. T.; Krylova, S. M.; Fournier, R.; Krylov, S. N. Fundamental determinants of the accuracy of equilibrium constants for affinity complexes. *Analytical Chemistry* **2023**, *95* (42), 15826–15832." Copyright 2023 American Chemical Society. My contributions to the article were: (*i*) conducting the mathematical analysis, (*ii*) preparing all figures, (*iii*) interpreting the results, and (*iv*) writing the manuscript. Co-author Daniel Everton conducted the NECEEM experiments.

3.1.1. Introduction: The Significance of Understanding the Fundamental Determinants of K_d Accuracy

Equilibrium constants of chemical reactions are the cornerstone of chemical thermodynamics. We see them multiple times through the years of studies, starting with a freshmen-chemistry course. We learn how to calculate them using tabulated standard free energies of formation through **Eq (1.5)**, typically presented with three–four significant figures. Accordingly, we naturally expect that equilibrium constants are accurately determined parameters. The real world, however, is very far from this utopia.

The vast majority of equilibrium constants that are published in the scientific literature today are those of binding reactions involving large biological molecules and forming highly stable affinity complexes.[209, 210, 211] These equilibrium constants cannot be calculated through **Eq (1.5)** because standard free energies of formation of typical interactants, such as proteins, DNA, and

drug leads, *etc.*, are unknown. Therefore, equilibrium constants of affinity complexes are determined experimentally using a large scope of methods as we reviewed in **Section 1.4.1**.[212, 213, 214] Most of these methods are supported by commercial instruments in which a researcher essentially loads the solutions into the instrument, pushes a button, and reads a resulting value with multiple significant figures and a small standard deviation. While being enormously powerful and informative tools, user-friendly commercial instruments inadvertently create an illusion that the determination of equilibrium constants is a trivial technical step, which requires only following an instrument manual. Accordingly, equilibrium-constant values published in a single paper are typically determined with a single method and presumed to be accurate values. However, discrepancies between the values reported by different laboratories for the same complex may reach orders of magnitude leading to misconceptions and wrong conclusions, *e.g.*, about the potency of drug leads.[32, 35] Great variations are often observed even for the values determined with the same method, suggesting that fundamental (method-independent) sources of inaccuracy play a critical role in inconsistencies of equilibrium constants for highly stable affinity complexes.[34, 215] There is an understanding that the amount of inconsistent data grows rapidly due to the use of high-throughput methods.[105, 216] The described "crisis of inaccuracy" in the field of thermodynamics of affinity complexes calls for a systematic search for sources of inaccuracy and approaches that could improve the accuracy. This work focuses on fundamental sources of inaccuracy of equilibrium constants for affinity complexes.

In this study, we will use notations that are typically used in the field of affinity interactions. The interacting species will be called a target and a ligand. Their binding with the formation of the target–ligand complex is described by **Eq (1.1)**. The stability of the formed complex is characterized by the equilibrium dissociation constant (K_d), which can be calculated through

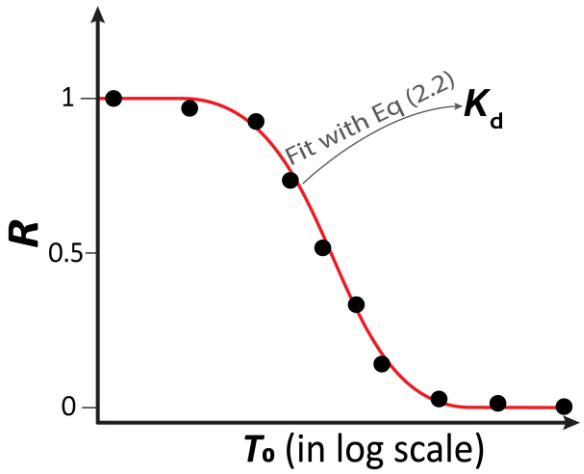


Figure 3.1. An example of determining K_d with a binding isotherm. A binding isotherm is the experimental dependence of the fraction of unbound ligand (R) on total target concentration (T_0), which is represented by black dots. The value of K_d is determined by the best fit (red line) of the isotherm with **Eq (2.2)**.

Eq (1.2). Lower K_d values correspond to greater complex stability and are typically desired.[217, 218, 219] The values of K_d define therapeutic concentrations of drugs and analytical parameters of diagnostic methods.[219, 220, 221, 222] Therefore, they are used not only for relative ranking of the ligands but also for assessment of their suitability for intended applications.[220, 221, 223] The most straightforward approach to K_d determination is an equilibrium (*i.e.*, non-kinetic) method. It involves preparing a series of equilibrium mixtures of target and ligand in which the total concentration of ligand (L_0) is the same while the total concentration of target (T_0) varies. The mixtures are used to build a binding isotherm (dots in **Figure 3.1**), *i.e.*, the experimental dependence of a fraction R of unbound ligand ($R = L/L_0$) on T_0 for constant L_0. The value of K_d is then typically found by fitting the binding isotherm with the theoretical dependence of R on T_0 (**Eq (2.2)**: $R = -(K_d + T_0 - L_0)/(2L_0) + \sqrt{((K_d + T_0 - L_0)/(2L_0))^2 + K_d/L_0}$),[101, 171, 224] while varying

K_d until the best fit (red line in **Figure 3.1**) is obtained. It is important to emphasize that **Eq (2.2)** is obtained by solving a quadratic obtained from an expression for K_d:

$$K_d = \frac{T_0 - L_0(1-R)}{(1/R - 1)} \tag{3.1}$$

which, in turn, is obtained from first principles: definitions of K_d (**Eq (1.2)**) and R (**Eq (2.1)**) as well as mass balance for the target and ligand ($T_0 = T + C$ and $L_0 = L + C$). In other words, **Eqs (2.2)** and **(3.1)** are fundamental, *i.e.*, method-independent and obtained without assumptions that would restrict the generality.

There are two fundamental requirements for K_d accuracy: (*i*) the binding reaction (**Eq (1.1)**) must approach the equilibrium and (*ii*) the study must be done in a so-called binding regime which is assured by satisfying a condition of $L_0/K_d << 1$.[34, 181, 225, 226] According to a recent review by Jarmoskaite *et al.*,[34] of 100 publications dealing with K_d determination, fewer than 10% and 5% of publications reported the satisfying of these two requirements, respectively. Failure to satisfy fundamental requirements may deem the vast majority of the published K_d values greatly inaccurate. The alarming level of researchers' ignoring the fundamental requirement is due to the lack of basic knowledge on how this can affect the accuracy of K_d. There are a few important contributions to the field, which assess K_d accuracy for specific methods,[227, 228, 229] *e.g.*, isothermal titration calorimetry[226, 230] and capillary electrophoresis.[199] However, there are no comprehensive studies that would explain what errors of K_d depend on fundamentally in the way that would further guide theorists and instruct experimenters. Accordingly, the field of experimental determination of K_d is largely an art in which researchers rely mostly on mnemonic rules and intuition in designing experiments and interpreting their results. The goal of our work was to initiate a graduate maturation of this field into a solid quantitative science through the

understanding of what fundamental errors of K_d depend on and what it means for experimentalists.

3.1.2. Materials and Methods

3.1.2.1. Chemicals, Materials, and Solutions

All chemicals were purchased from Sigma-Aldrich (Oakville, ON, Canada) unless otherwise stated. Fused-silica capillaries with an inner diameter of 75 μm and an outer diameter of 360 μm were purchased from Molex Polymicro (Phoenix, AZ, USA) and used throughout this work. His-tagged recombinant Thermus aquaticus MutS protein (MW = 92.8 kDa) was purchased from Prospec (Ness-Ziona, Israel). An Alexa488-labeled MutS-binding aptamer was synthesized by Integrated DNA Technologies (Coralville, IA, USA) with the following sequence: 5'-Alexa488-CTT CTG CCC GCC TCC TTC CTG GTA AAG TCA TTA ATA GGT GTG GGG TGC CGG GCA TTT CGG AGA CGA GAT AGG CGG ACA CT-3'.[231] The aptamer stock solution was subjected to annealing by incubating at 90°C for 2 min before cooling it to 20 °C at a rate of 0.5 °C/s, *prior to* dilution and preparation of equilibrium mixtures. One single sample buffer (SB) was used for all experimental procedures: 50 mM Tris-Acetate pH 8.2 supplemented with 0.1% Tween 20 and 0.1% Tween 80 to reduce adsorption of DNA and MutS protein to surfaces. All solutions were made using double water deionized with Milli-Q and filtered through a 0.22 μm filter (Millipore, Nepean, ON, Canada); we term it ddH_2O.

3.1.2.2. Poly (vinyl alcohol) (PVA)-Coating

To reduce sample adsorption, the inner wall of all capillaries used in capillary electrophoresis (CE) was coated with Poly(vinyl alcohol) (PVA) using a previously described procedure.[199] PVA (5%, w/v) was prepared by dissolving the polymer in boiling ddH_2O. Each time, two 130-cm-long capillaries were coated simultaneously. The two uncoated fused-silica capillaries were

sequentially flushed with 0.1 M NaOH and ddH_2O for 1 h under a pressure of 12 psi. The pretreated capillaries were then flushed with the PVA solution for 10 min at 15 psi and emptied using a 10-psi nitrogen gas flow for 10 min. PVA was immobilized on the capillary surface by drying overnight in an oven set at 140 ^oC and continuously flushed with nitrogen at a pressure of 5 psi. The detection window on the capillary was made by removing the outer polyimide layer with a fuming solution of H_2SO_4.

3.1.2.3. Experimental Procedure

All CE experiments were performed with a P/ACE MDQ apparatus (SCIEX, Concord, Ontario, Canada) equipped with a laser-induced fluorescence (LIF) detection system. Fluorescence was excited with a blue line (488 nm) of a solid-state laser and detected at 520 nm. All capillaries were PVA-coated and had a length of 50 cm (approximately 40 cm from the capillary inlet to the detector). Prior to each run, the capillary was rinsed with the running buffer at a pressure of 20 psi for 3 min.

All dilutions of sample components were made using SB. All equilibrium mixtures were prepared to final volumes of 40 µL in SB and incubated at an ambient temperature of 23 ± 1 ^oC for a minimum of 1 h prior to injection. A sample plug was injected into a capillary by pressure at 1.0 psi for 10 s (sample plug's length = 1.3 cm, volume = 59 nL) and then propagated at 0.9 psi for 30 s (propagation distance = 3.6 cm) with SB to pass the uncooled region of the capillary. The sample was then electrophoresed at 25 kV for 15 min with a negative electrode at the capillary inlet. The capillary coolant was set to 15 °C.

3.1.3. Results and Discussion

3.1.3.1. Theoretical Problem

K_d values are not measured, they are determined using **Eqs (2.2)** and **(3.1)** from known values of T_0, L_0, and R (in the case of **Eq (2.2)** we also need to use non-linear regression). Accordingly, a deviation of the determined K_d value ($K_{d,det}$) from the true K_d value (K_d) is a result of errors of T_0, L_0, and R. So, the answer to the question of what the accuracy of K_d depends on is simple: it depends on the accuracy of T_0, L_0, and R. However, there is a second important question: how does the error of K_d depend on the errors of T_0, L_0, and R? If the errors of T_0, L_0, and R are not zero, which is always the case, then the error of K_d is the result of the propagation of errors of T_0, L_0, and R. Therefore, the answer to the second question is also conceptually simple: the dependence of the accuracy of K_d on the errors of T_0, L_0, and R is defined by the rules of error propagation. Accordingly, we will apply error-propagation rules to **Eq (3.1)** to find how the deviation of $K_{d,det}$ from K_d ($\Delta K_d = K_{d,det} - K_d$) depends on deviations of T_0, L_0, and R (ΔT_0, ΔL_0, and ΔR, respectively, defined similarly to ΔK_d) from their true values.

Note that deviations ΔT_0, ΔL_0, and ΔR are not random errors with a given error distribution; they can be both positive and negative. ΔK_d is a systematic error, which defines the accuracy of $K_{d,det}$. ΔK_d may be much greater than the random error of $K_{d,det}$. Furthermore, ΔK_d is virtually impossible to determine since the true K_d value is fundamentally unknown owing to the absence of standard reference instruments for the determination of K_d or standard reference K_d values.

3.1.3.2. Propagation of Correlated Errors

Since the sample preparation processes for both target and ligand are similar and R values are determined from the signals of bound and unbound ligand, the error sources for T_0, L_0, and R are linearly correlated in K_d-determination experiments. Therefore, to investigate the error of K_d propagated from the errors of T_0, L_0, and R, we can write the general dependence of ΔK_d on ΔT_0, ΔL_0, and ΔR as:

$$\left|\Delta K_{\mathrm{d}}\right| = \left|\left(\frac{\partial K_{\mathrm{d}}}{\partial T_0}\right)\Delta T_0\right| + \left|\left(\frac{\partial K_{\mathrm{d}}}{\partial L_0}\right)\Delta L_0\right| + \left|\left(\frac{\partial K_{\mathrm{d}}}{\partial R}\right)\Delta R\right| \tag{3.2}$$

This expression corresponds to the scenario in which the input variables T_0, L_0, and R are assumed to be fully correlated; that is, any error in one variable is accompanied by proportional errors in the others due to their linear relationship. In the general form of first-order error propagation, covariance terms appear explicitly to account for partial correlations between variables. However, the actual degrees of correlation are often unknown in experimental settings. Rather than attempting to estimate these unknown correlations, we adopt a conservative approach by assuming full correlation (*i.e.*, correlation coefficient = 1). Under this assumption, the total propagated error becomes a linear sum of the individual contributions. This approach allows us to evaluate the worst-case scenario, providing an upper bound on the impact of systematic errors in T_0, L_0, and R on the calculated K_{d}.

By applying the error propagation rule (**Eq (3.2)**) to **Eq (3.1)**, and with a series of mathematical transformations (Note B1.2 in **Appendix B1**), we finally obtain:

$$\begin{aligned}
&\left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = a + b\frac{L_0}{K_{\mathrm{d}}}, \\
&a = t + \frac{r(1+r)}{0.5+r},\ b = \frac{0.5(t+l-r)}{1+r} + \frac{0.5r}{0.5+r} \ \text{for}\ 0 \le \Delta R < 0.5 \\
&a = t + \frac{r(1-r)}{0.5-r},\ b = \frac{0.5(t+l-r)}{1-r} + \frac{0.5r}{0.5-r} \ \text{for}\ -0.5 < \Delta R < 0
\end{aligned} \tag{3.3}$$

where $t = |\Delta T_0/T_0|$, $l = |\Delta L_0/L_0|$, and $r = |\Delta R/R|$, respectively. **Eq (3.3)** reveals that, when the error sources for T_0, L_0, and R are strongly correlated, the relative error of K_{d} is a linear function of L_0/K_{d} with an intercept dependent only on relative errors of T_0 and R and the slope dependent on relative errors of all three variables: T_0, L_0, and R. In the mathematical analysis (Note B1.2), we made a single simplifying assumption that the determined (from measured signals) value of R is

equal to 0.5, which leads to the least erroneous K_d,[101] thus, we are considering the best-case scenario and are finding the lower limit for the error of K_d.

Eq (3.3) explicitly shows the role of L_0/K_d in the relative error of K_d. The value of L_0/K_d is only important when the second term is greater than the first, *i.e.*, when $L_0/K_d > a/b$. In other words, in contrast to a commonly used requirement of $L_0/K_d << 1$, it is unnecessary to decrease L_0/K_d much below a/b. Rather, using $L_0/K_d << a/b$ will likely cause the opposite effect; namely, it will lead to an increase in $|\Delta K_d/K_d|$ due to the increase of r when L_0 is too low, causing an unacceptable signal to noise ratio (S/N), which will be briefly discussed later. On the other hand, when $L_0/K_d > a/b$, the increase in L_0/K_d plays a crucial role in increasing the relative error of K_d; an order of magnitude increase in L_0 will lead to an order of magnitude increase in $|\Delta K_d/K_d|$. One should appreciate that if $L_0/K_d >> a/b$, then $K_{d,det}$ may differ from true K_d by orders of magnitude.

For a large range of L_0 values, which is typically the case, it is more convenient to present the dependence of $|\Delta K_d/K_d|$ on L_0/K_d in a double-log scale. A graph of $\log(|\Delta K_d/K_d|)$ versus $\log(L_0/K_d)$ is triphasic: two asymptotically-linear ranges flank a non-linear transition range (**Figure 3.2a** and **Figure B1.1a**). For small values of L_0/K_d, *i.e.*, for $L_0/K_d << a/b$, the dependence is a linear function with no dependence on L_0/K_d: $\log(|\Delta K_d/K_d|) = \log(a) = \text{const}$. For a large value of L_0/K_d, *i.e.*, for $L_0/K_d >> a/b$, the dependence is a linear function $\log(|\Delta K_d/K_d|) = \log(b) + \log(L_0/K_d)$ with an intercept with the ordinate equal to $\log(b)$ and a slope equal to unity. The abscissa of the intersection of the two asymptotic lines is defined by $\log(L_0/K_d) = \log(a/b)$.

Although the two outer regions appear as distinct straight lines on a log-log plot, it is important to emphasize that this does not indicate distinct linear relationships in the conventional linear-linear coordinates. In this dissertation, the term "linear", when used in this context, refers specifically to apparent linearity on a log-log scale, not to mathematical linearity on standard

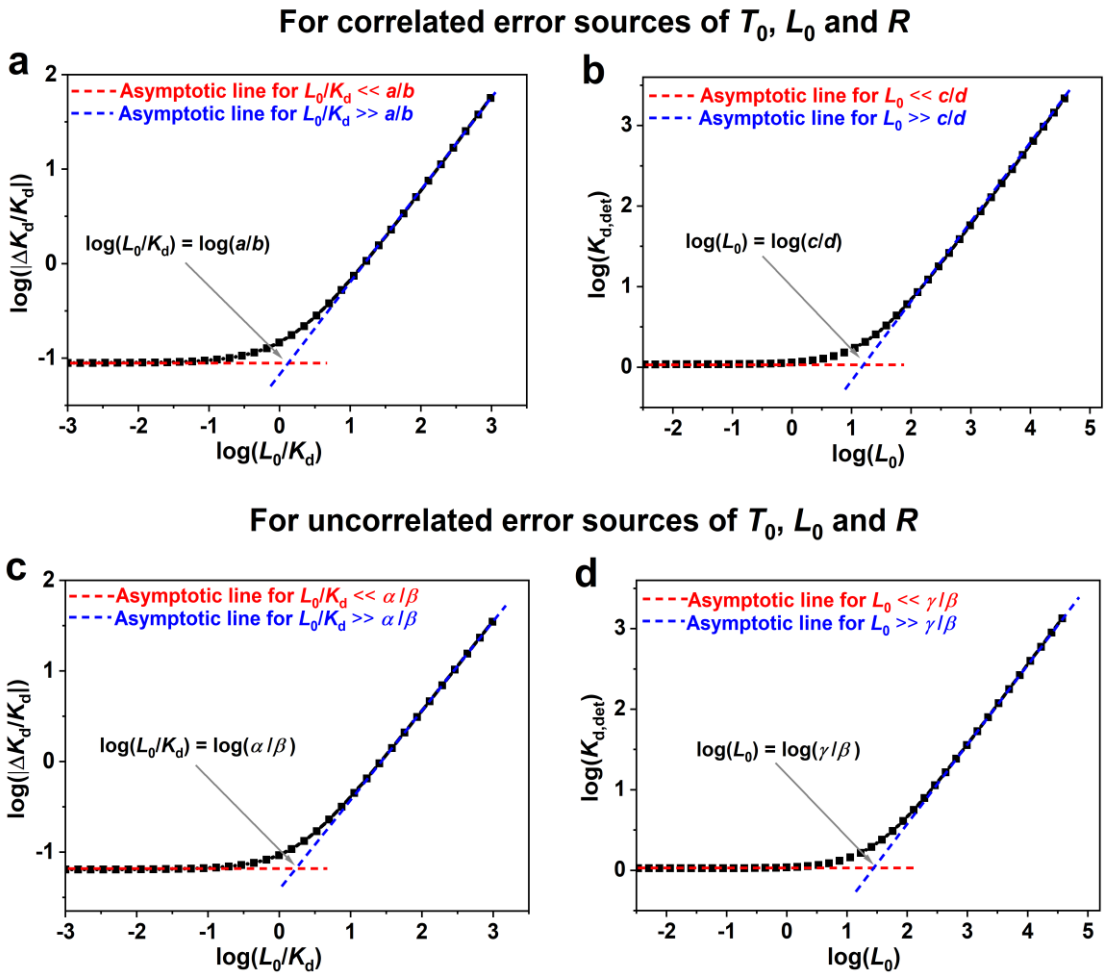


Figure 3.2. General trends in dependencies of $\log(|\Delta K_d/K_d|)$ on $\log(L_0/K_d)$ (**(a)** and **(c)**) and $\log(K_{d,det})$ on $\log(L_0)$ (**(b)** and **(d)**). In all panels, reasonable relative errors of L_0, T_0, and R were used: $|\Delta T_0/T_0| = |\Delta L_0/L_0| = 0.05$ and $\Delta R/R = 0.02$ (the results corresponding to $\Delta R/R = -0.02$ are shown in **Figures B1.1a** and **B1.1c**). In panels **(b)** and **(d)**, we set $K_d = 1$ (unitless) and $\Delta K_d \geq 0$ (the results corresponding to $\Delta K_d < 0$ are shown in **Figures B1.1b** and **B1.1d**). With these conditions, parameters a, b, c, and d were calculated to be 0.089, 0.058, 1.1, and 0.058, respectively; α, β, and γ were calculated to be 0.064, 0.036, and 1.1, respectively.

Cartesian coordinates. While descriptors such as "sensitive" and "insensitive" may more rigorously reflect the behavior of systematic error in different regimes, the terminology of "linear" regions is adopted here for consistency with the published work derived from this study.

This usage will be maintained consistently throughout the remainder of the dissertation, with this clarification provided at the outset to prevent potential misinterpretation.

Using **Eq (3.3)** and $\Delta K_d = K_{d,det} - K_d$, we can show that (for strongly correlated error sources) not only $|\Delta K_d/K_d|$ depends linearly on L_0/K_d, but also $K_{d,det}$ is a linear function of L_0 (Note B1.3):

$$\begin{aligned} K_{d,det} = c + dL_0,\ c = (1+a)K_d,\ d = b \ \text{ for } \Delta K_d \geq 0 \\ c = (1-a)K_d,\ d = -b \ \text{ for } \Delta K_d < 0 \end{aligned} \tag{3.4}$$

If presented in a double-log scale, this dependence is expectedly triphasic (**Figure 3.2b**) and has features similar to the dependence of $\log(|\Delta K_d/K_d|)$ on $\log(L_0/K_d)$ (**Figure 3.2a**).

3.1.3.3. Propagation of Uncorrelated Errors

In some experimental cases, the error sources of L_0, T_0, and R may be weakly correlated. For example, stock concentrations of the ligand and target may largely differ from the labelled values in the opposite direction. In these cases, the errors from the stock reagents are uncorrelated (independent of each other) and might be much larger than the errors caused by the instruments or methods used in sample preparation. In addition, using an uncalibrated instrument to detect ligand signals can cause the errors in determined R values to be uncorrelated with the errors of T_0 and L_0. To study the error propagation with uncorrelated (orthogonal) error sources, we should write the dependence of ΔK_d on ΔT_0, ΔL_0, and ΔR as:

$$|\Delta K_d| = \sqrt{\left(\frac{\partial K_d}{\partial T_0}\right)^2 \Delta {T_0}^2 + \left(\frac{\partial K_d}{\partial L_0}\right)^2 \Delta {L_0}^2 + \left(\frac{\partial K_d}{\partial R}\right)^2 \Delta R^2} \tag{3.5}$$

Then, we apply the error-propagation rule of **Eq (3.5)** to **Eq (3.1)**, and obtain (Note B1.4):

$$\left|\frac{\Delta K_d}{K_d}\right| = \sqrt{\alpha^2 + \lambda\frac{L_0}{K_d} + \beta^2\left(\frac{L_0}{K_d}\right)^2}$$

For $0 \le \Delta R < 0.5$:

$$\alpha = \sqrt{t^2 + \left(\frac{r + r^2}{0.5 + r}\right)^2}, \beta = \frac{0.5}{1+r}\sqrt{t^2 + l^2 + \left(\frac{0.5r}{0.5+r}\right)^2}, \lambda = \frac{t^2}{1+r} + 2\left(\frac{0.5r}{0.5+r}\right)^2 \quad (3.6)$$

For $-0.5 < \Delta R < 0$:

$$\alpha = \sqrt{t^2 + \left(\frac{r - r^2}{r - 0.5}\right)^2}, \beta = \frac{0.5}{1-r}\sqrt{t^2 + l^2 + \left(\frac{0.5r}{r-0.5}\right)^2}, \lambda = \frac{t^2}{1-r} + 2\left(\frac{0.5r}{r-0.5}\right)^2$$

Here, as explained in the previous section, $t = |\Delta T_0/T_0|$, $l = |\Delta L_0/L_0|$, and $r = |\Delta R/R|$, respectively, and we only investigate the case with the determined R being equal to 0.5 to simplify the study.

Although the forms of **Eq (3.3)** and **Eq (3.6)** are very different, **Eq (3.6)** shows a significant similarity to **Eq (3.3)** regarding the triphasic dependence of $|\Delta K_d/K_d|$ on L_0/K_d. For small L_0/K_d values, **Eq (3.6)** approaches a linear function of $|\Delta K_d/K_d| = \alpha = \text{const}$ that shows no dependence on L_0/K_d. For large L_0/K_d values, **Eq (3.6)** approaches a linear function of $|\Delta K_d/K_d| = \beta(L_0/K_d)$ that shows a high sensitivity of $|\Delta K_d/K_d|$ to L_0/K_d. There is a non-linear transition range between these two linear lines. Moreover, both **Eq (3.3)** and **Eq (3.6)** suggest that the minimum $|\Delta K_d/K_d|$ value depends on a parameter (a in **Eq (3.3)** or α in **Eq (3.6)**) that is defined only by t and r, while the sensitivity of $|\Delta K_d/K_d|$ to L_0/K_d mainly depends on a parameter (b in **Eq (3.3)** or β in **Eq (3.6)**) that is defined by t, l, and r.

To show the dependence of $|\Delta K_d/K_d|$ on L_0/K_d in a large range of L_0/K_d for uncorrelated error sources, we demonstrate "$\log(|\Delta K_d/K_d|)$ versus $\log(L_0/K_d)$" in **Figure 3.2c**. In the double-log scale (**Figure 3.2c**), the asymptotic lines for small and large L_0/K_d values are approximated with functions of $\log(|\Delta K_d/K_d|) = \log(\alpha)$ and $\log(|\Delta K_d/K_d|) = \log(\beta) + \log(L_0/K_d)$, respectively.

Figure 3.2c suggests that, to avoid significant increase of r caused by unacceptably low signal to noise ratio, it is unnecessary to decrease L_0/K_d much below a certain value α/β that is defined by the abscissa of intersection of the two asymptotic lines (*i.e.*, $\log(L_0/K_d) = \log(\alpha/\beta)$). According to the results in **Figure 3.2** and **Figure B1.1**, with the same t, l, and r, the value of α/β is slightly greater than a/b because $|\Delta K_d|$ obtained with **Eq (3.5)** is smaller than that determined with **Eq (3.3)**.

Now, let's discuss how $K_{d,det}$ depends on L_0 when the error sources are uncorrelated. With **Eq (3.6)** and the definition of ΔK_d (*i.e.*, $\Delta K_d = K_{d,det} - K_d$), we obtain (Note B1.5):

$$\begin{aligned} K_{d,det} &= K_d + \sqrt{\alpha^2 K_d^2 + \lambda K_d L_0 + \beta^2 L_0^2} \text{ when } \Delta K_d \geq 0 \\ &= K_d - \sqrt{\alpha^2 K_d^2 + \lambda K_d L_0 + \beta^2 L_0^2} \text{ when } \Delta K_d < 0 \end{aligned} \quad (3.7)$$

Eq (3.7) shows that the dependence of $K_{d,det}$ on L_0 is also triphasic. For small L_0 values, $K_{d,det}$ is insensitive to the change of L_0 and **Eq (3.7)** approaches a linear function of $K_{d,det} = \gamma$ ($\gamma = (1 + \alpha)K_d$ when $\Delta K_d \geq 0$ and $\gamma = (1 - \alpha)K_d$ when $\Delta K_d < 0$) (Note B1.5). For large L_0 values, **Eq (3.7)** approaches another linear function of $K_{d,det} = \beta L_0$ that shows high sensitivity of $K_{d,det}$ to L_0 (only applicable for $\Delta K_d \geq 0$ since $K_{d,det}$ has to be greater than zero) (Note B1.5). A non-linear transition range exists between the two linear lines. **Figure 3.2d** shows an example of the dependence of $K_{d,det}$ on L_0 for uncorrelated error sources of T_0, L_0, and R in a double-log scale, which has the expected triphasic feature.

3.1.3.4. Experimental Validation

The dependence shown in **Figures 3.2b** and **3.2d** has two known variables — L_0 and $K_{d,det}$ — and therefore, it can be tested experimentally, which we have undertaken in this work. Experimentally, we used a well-established solution-based method of nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) as the K_d-determination method and MutS–

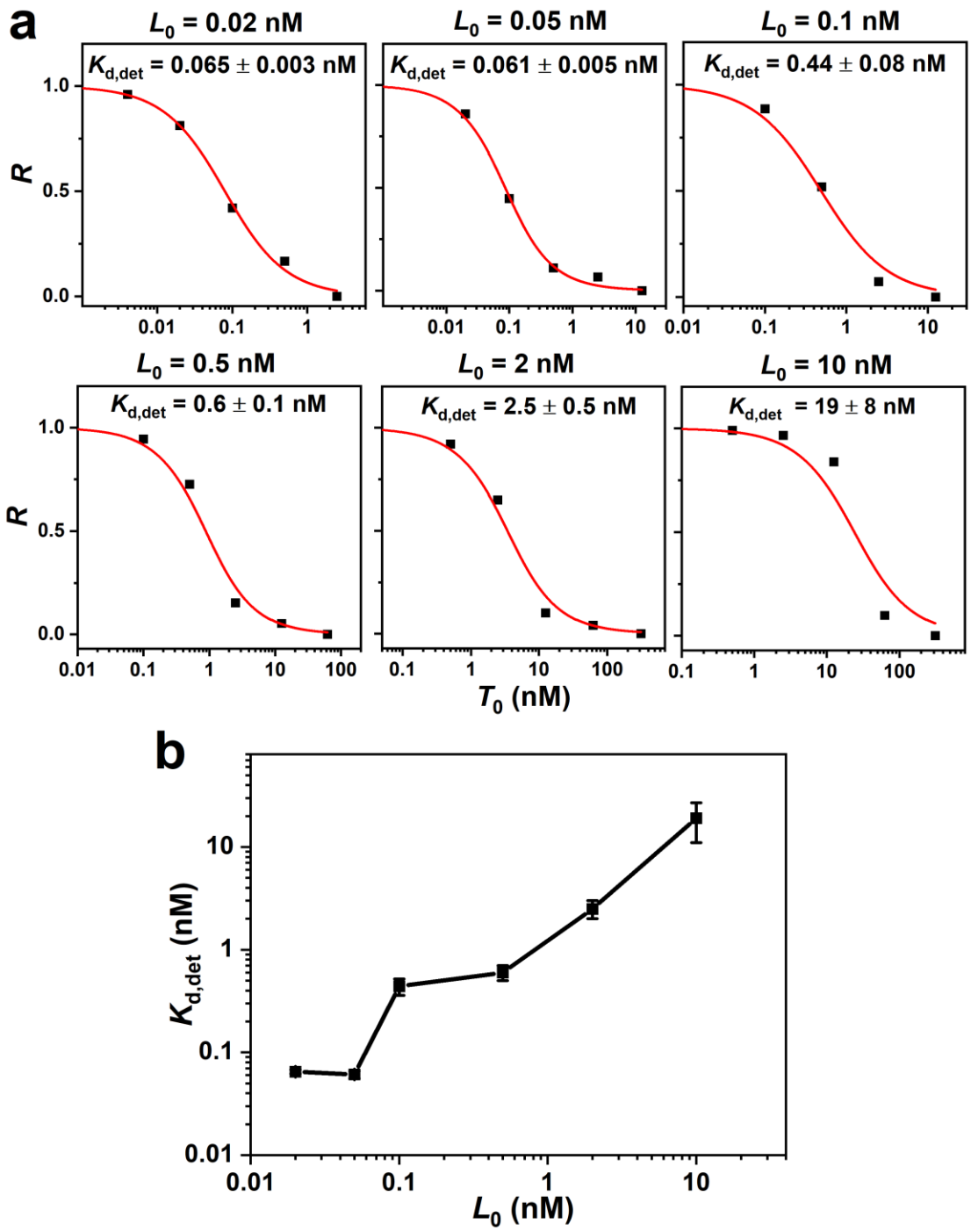


Figure 3.3. The influence of L_0 on binding isotherms **(a)** and the dependence of $K_{d,det}$ on L_0 **(b)**. In panel **(a)**, all nonlinear regressions were conducted using OriginPro software with the iteration algorithm of Levenberg Marquardt. See **Appendix B1** for other details.

aptamer as the binding pair.[97, 101, 191, 232] In NECEEM, unbound ligand and complex (bound ligand) are separated by an electric field in a capillary due to their different charge-to-size ratios.[97, 101, 191, 232] In six NECEEM experiments, the aptamer concentration (L_0) was kept constant at 0.02, 0.05, 0.1, 0.5, 2, and 10 nM, respectively; MutS concentration (T_0) varied from 0 to 312.5 nM (all concentrations are given as nominal values with no significance to the number of digits). The lowest value of L_0 (*i.e.*, 0.02 nM) was chosen to be equal to the limit of quantitation (LOQ) (**Figure B1.2**) linked to *S*/*N* as discussed below. The influence of L_0 on binding isotherms and the dependence of $K_{d,det}$ on L_0 are shown in **Figure 3.3**. **Figure 3.3b** shows that (in a double-log scale) $K_{d,det}$ is insensitive to the change of L_0 when L_0 is at low concentrations (*e.g.*, << ~0.1 nM), and $K_{d,det}$ is sensitively dependent on L_0 when L_0 is at high concentrations. The results in **Figure 3.3b** are consistent with the theoretical predictions shown in **Figures 3.2b** and **3.2d**, which confirms the validity of our theoretical analysis.

3.1.3.5. Instructive Guidance for Experimentalists

Although **Figure 3.2** suggests that, with any type of experimental error sources (correlated or uncorrelated), we can always use an as-low-as-possible L_0 to avoid the error of $K_{d,det}$ being greatly magnified, the choice of the lowest experimentally suitable L_0 is dictated by the limit of quantitation (LOQ) of an instrument used to measure the signal. The value of LOQ is, by definition, the analyte concentration (ligand concentration L_0 in our case) for which *S*/*N* is equal to a certain value $X >> 1$, which guarantees that the noise does not affect quantitation significantly. If the noise is independent of the signal and the value of LOQ is known, *S*/*N* can be calculated for any given concentration of the ligand as $S/N = XL_0/\text{LOQ}$ (Note B1.6). The minimum acceptable ligand concentration that satisfies this equality is $L_0 = \text{LOQ}$. Since *R* values are determined based on the measured signals *S*, using $L_0 < \text{LOQ}$ is counterproductive as it will

cause a significant increase in the relative error of R (r), and, accordingly, in the relative error of $K_{d,det}$.

Based on above analysis, to achieve good accuracy of $K_{d,det}$ in an experiment, if LOQ/K_d < a/b (or α/β for uncorrelated error sources), L_0/K_d should satisfy:

$$\frac{\text{LOQ}}{K_d} \leq \frac{L_0}{K_d} < \frac{a}{b} \left(\text{or } \frac{\alpha}{\beta} \right) \tag{3.8}$$

where a, b, α, and β are defined by the absolute values of relative systematic errors of T_0, L_0, and R (*i.e.*, t, l, and r) as shown in **Eqs (3.3)** and **(3.6)**. Because t, l, and r are relatively difficult to determine, practically, LOQ can be determined and used as a sole determinant of ligand concentration to minimize the error of $K_{d,det}$ in a K_d-determination experiment. It is important to emphasize that using L_0 = LOQ does not imply that $L_0/K_d < a/b$ (or α/β) and thus does not guarantee the accuracy of K_d. Therefore, with the current state of affairs, the only way to confirm (indirectly) that $L_0/K_d < a/b$ (or α/β) is to conduct experiments not only at L_0 = LOQ but also at L_0 > LOQ (*e.g.*, the experiments corresponding to L_0 = 0.02 and 0.05 nM shown in **Figure 3.3**). If K_d is not affected by the value of L_0 then one can assume that $L_0/K_d < a/b$ (or α/β) and deem K_d accurate.

3.1.4. Concluding Remarks

To conclude, in this study, we investigated the theoretical dependence of the relative deviation of $K_{d,det}$ from K_d ($|\Delta K_d/K_d|$) on the ratio of L_0/K_d with error propagation. Our analysis shows that, with fixed correlated or uncorrelated errors of T_0, L_0, and R, $|\Delta K_d/K_d|$ always presents a triphasic dependence on L_0/K_d in a double-log scale: When L_0/K_d is small, $|\Delta K_d/K_d|$ is insensitive to the change of L_0/K_d and approaches a constant that is defined by the errors of T_0 and R. When L_0/K_d is large, $|\Delta K_d/K_d|$ is sensitive to the change of L_0/K_d, *i.e.*, an order of

magnitude increase in L_0 will lead to an order of magnitude increase in $|\Delta K_d/K_d|$. There is a non-linear transition range between the two linear phases. We also theoretically demonstrated that the dependence of $K_{d,det}$ on L_0 should also show the triphasic features in a double-log scale, which was experimentally confirmed with NECEEM experiments using MutS–aptamer as the binding pair. The results of this work suggest that, without knowing the errors of T_0, L_0, and R, experimentalists should use $L_0 = \text{LOQ}$ (*i.e.*, the smallest ligand concentration that does not sacrifice the accuracy of R) in an equilibrium K_d-determination method and conduct an extra experiment with $L_0 > \text{LOQ}$ to confirm the accuracy of determined K_d. To emphasize, we foresee that our findings (*e.g.*, **Eqs (3.3)** and **(3.6)** and the triphasic dependence of $|\Delta K_d/K_d|$ on L_0/K_d in **Figure 3.2**) can help create an approach of assessing K_d accuracy from a single binding isotherm if the ranges of relative errors in T_0, L_0, and R can be estimated.

3.2. Maximizing K_d Accuracy and Defining the Confidence Intervals of Systematic Errors in Experimental Variables

The following material was previously published and is reprinted with permission from "**Wang, T. Y.**; Rukundo, J.-L.; Mao, Z.; Krylov, S. N. Maximizing the Accuracy of Equilibrium Dissociation Constants for Affinity Complexes: From Theory to Practical Recommendations. *ACS Chemical Biology* **2024**, *19* (9), 1852–1867." Copyright 2024 American Chemical Society. My contributions to the article were: (*i*) conducting mathematical analysis and numerical simulations, (*ii*) preparing all figures, (*iii*) interpreting the results, and (*iv*) writing the manuscript. Co-author Jean Luc Rukundo conducted the numerical simulations and mathematical analysis for "Diffusivity-Based Methods".

3.2.1. Introduction: The Critical Need for a Systematic Study on Strategies to Maximize K_d Accuracy

Despite the critical importance of accurate K_d values across a wide range of applications, K_d values determined for the same complex under similar conditions often differ from each other by folds or even orders of magnitude while having small relative standard deviations (*i.e.*, random errors).[32, 33] Such substantial differences suggest that K_d determination is subject to large systematic errors, or in other words, is inherently inaccurate.

A determined K_d value ($K_{d,det}$) always differs from the true value of K_d, and systematic error in $K_{d,det}$ (ΔK_d) is defined as:

$$\Delta K_d = K_{d,det} - K_d \tag{3.9}$$

Since the true value of K_d is fundamentally unknown, the systematic error of $K_{d,det}$ cannot be found using this definition. Therefore, the question of "how accurate is the determined value of K_d?" lacks a definite quantitative answer. Our inability to quantitatively characterize the accuracy of $K_{d,det}$ underscores the imperative to implement all available measures aimed at improving this accuracy, or in other words, minimizing the systematic error in $K_{d,det}$.

As we mentioned in the previous sections, the ramifications of inaccuracies in K_d determination are profound. Inaccurate K_d values can distort the ranking of drug leads, potentially prioritizing low-potency compounds over high-potency ones, thereby compromising the efficacy of R&D efforts in drug discovery.[33, 233, 234, 235, 236] Moreover, inaccurate K_d values can lead to misinterpretation of biological data and perpetuate fundamental misconceptions in biological research for extended periods.[33, 237, 238] It is crucial to emphasize the necessity of accurate K_d determination, which unfortunately remains a challenge. Compounding this issue is the absence of a method to quantitatively assess K_d accuracy.

Over the past decade, notable works and practical guidelines have been developed to help researchers improve the accuracy of experimentally determined K_d values.[34, 180, 181] Despite this, only a small portion of studies adhere to these established guidelines.[34] This disregard is primarily due to two reasons. First, many experimentalists are unaware of how significantly systematic errors in variables (*e.g.*, concentrations) can affect the accuracy of K_d determination. Second, implementing the recommended measures, such as constructing multiple binding isotherms with varying incubation times and concentrations of the limiting component, is both time-consuming and resource-intensive. The widespread neglect of established guidelines has motivated us to conduct a systematic theoretical study aimed at (*i*) identifying fundamental determinants of K_d accuracy (**Section 3.1**).[179] (*ii*) quantitatively demonstrating the importance of eliminating common sources of systematic errors and providing guidance for doing so, (*iii*) developing an approach for assessing the accuracy of experimentally determined K_d from a single binding isotherm (will be discussed in **Chapters 4** and **5**),[10] and, ultimately, (*iv*) establishing consensus standard requirements for binding experiments by the broader research community.[11] Here we focus on the second aspect.

A classic way to determine K_d includes two major steps (**Figure 3.4a**).[35, 67, 101, 171] Step 1 is experimentally building a binding isotherm using a set of equilibrium mixtures containing two components: a limiting component, which we term ligand (L), and an excess component, which we term target (T). The ligand is kept at a constant total concentration L_0 across all equilibrium mixtures while the total concentration of the target T_0 is varied from mixture to mixture in a wide range. A binding isotherm is the dependence of a fraction R of unbound (or bound) ligand on T_0 at a constant L_0. Step 2 in K_d determination is nonlinear regression of the binding isotherm with a known regression model, $R = F(L_0, T_0, K_d)$, derived from the definitions of R and K_d as

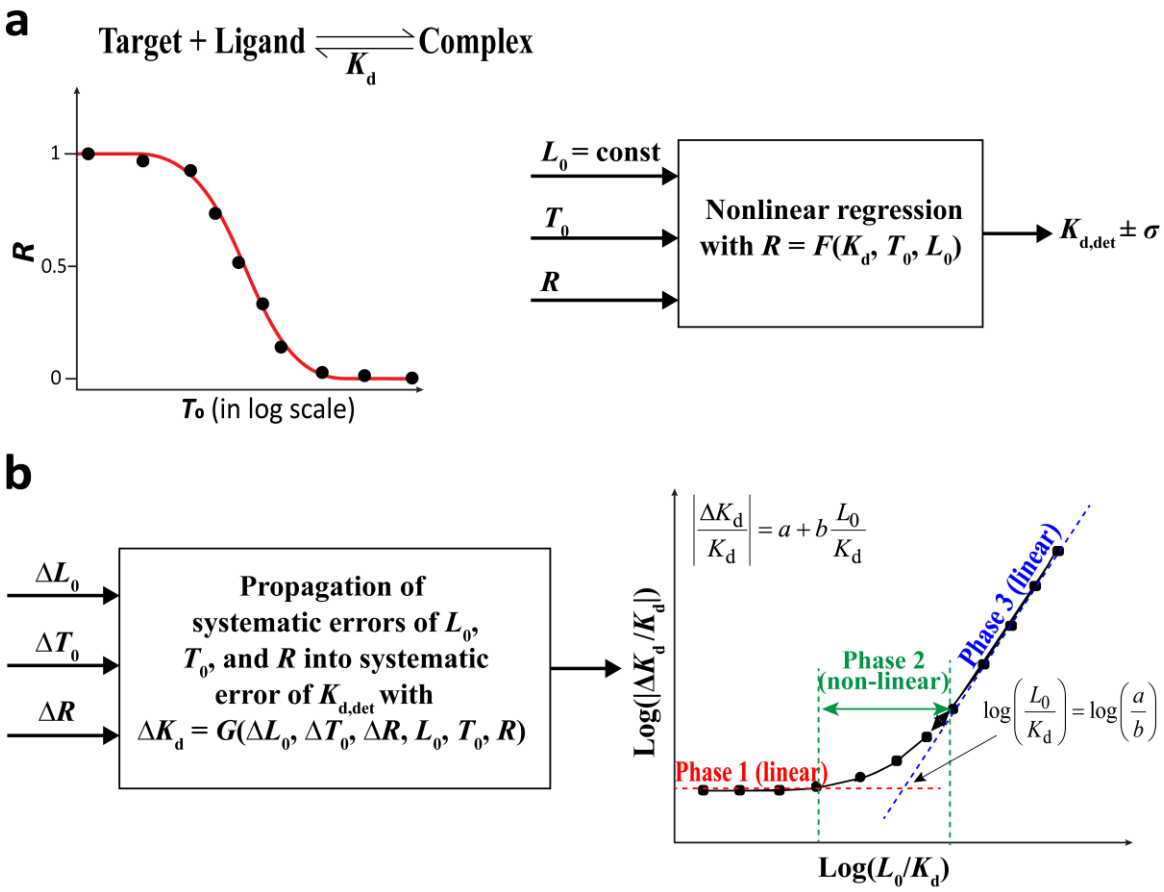


Figure 3.4. Schematic representations of **(a)** determination of K_d and its standard deviation (σ) by nonlinear regression with three variables from the binding isotherm (L_0, T_0, and R) and **(b)** propagation of systematic errors in these variables leading to a triphasic dependence of the relative systematic error in $K_{d,det}$ on the ratio L_0/K_d.

well as mass balance for ligand, target, and complex. In this regression procedure, K_d is an unknown sought parameter, L_0 and T_0 are known independent variables, and R is a known dependent variable. The result of nonlinear regression is a determined K_d value ($K_{d,det}$) and its standard deviation (random error) σ obtained under the assumption that all three variables (L_0, T_0, and R) are subject to random errors only. The value of σ describes the precision of $K_{d,det}$, but

suggests nothing about the accuracy of $K_{d,det}$.[239, 240] In other words, the probability of the K_d to be found inside the $K_{d,det} \pm \sigma$ range is unknown and can be very low.

As K_d is not measured directly but determined using nonlinear regression with a known regression model, the accuracy of $K_{d,det}$ is defined by the accuracy of the three variables (L_0, T_0, and R) (**Figure 3.4b**). Note, the use of an incorrect regression model will lead to inaccurate $K_{d,det}$ even if L_0, T_0, and R are free of systematic errors. Using error propagation rules, we can link the systematic error in $K_{d,det}$ to systematic errors in L_0, T_0, and R: $\Delta K_{d,det} = G(\Delta L_0, \Delta T_0, \Delta R, L_0, T_0, R)$. The dependence of $\Delta K_{d,det}$ on L_0 is such that when L_0 is much greater than the unknown K_d value, then the relative systematic error in $K_{d,det}$ is equal to a combination of relative systematic errors in the variable multiplied by L_0/K_d.[179] Decreasing systematic errors in $K_{d,det}$ thus requires decreasing both L_0 and systematic errors in the variables. As decreasing L_0 is restricted by limits of quantitation (LOQ) of instruments,[179] decreasing systematic errors in variables is viewed as an effective way of improving the accuracy of K_d determination.

Minimizing the systematic error in $K_{d,det}$ is an important task that can be translated into minimizing systematic errors in the variables. The drawback of such minimization is that it is blind; we minimize the error without knowing its value. The limitation of blind minimization is that even if all known precautions are taken to minimize systematic errors in variables (and, thus, in $K_{d,det}$), the resulting $K_{d,det}$ may still be insufficiently accurate for an intended application, but we will not know that. The major problem, which is not directly addressed in this work, is the lack of an approach for the quantitative assessment of the accuracy of $K_{d,det}$. We view the understanding of error propagation (**Figure 3.4b**) as an important starting point in the development of such an approach. The prospective approach for quantitative assessment of $K_{d,det}$ accuracy requires more than just the minimization of systematic errors in variables; it requires

their quantitative assessment. Thus, we started the presented work with two goals in mind: (*i*) minimization of systematic errors in variables and (*ii*) quantitative assessment of these errors for the prospective approach of quantitative assessment of the accuracy of $K_{d,det}$. Accordingly, after providing the required theoretical background on K_d determination and error propagation, we analyze major sources of systematic errors in variables R, L_0, and T_0, and we also identify means of minimizing these errors and study how they can be quantitatively assessed.

This work represents the first rigorous examination of the universal sources of systematic errors in K_d determination using equilibrium methods. As such, it serves as an indispensable resource for researchers conducting binding studies by offering detailed insights and practical methodologies to significantly enhance the accuracy of K_d determination. Furthermore, we anticipate that this work will lay the groundwork for incorporating K_d accuracy considerations into undergraduate and graduate curricula focused on binding methods. Additionally, it will support reviewers and editors in assessing the rigor and adequacy of experiments reported in scientific publications that present K_d values. By providing a comprehensive analysis and practical guidelines, this study aims to elevate research standards in this field, ensuring more reliable and reproducible results, ultimately benefiting the broader scientific community.

3.2.2. Review of the Theoretical Background

3.2.2.1. Binding Process

Since the considerations of stoichiometry and cooperativity of binding would greatly complicate this study, here we only consider the simple 1:1 binding process between a target and a ligand that was described in **Eq (1.1)**, which is the most abundant type of interactions in biological and technological contexts:[180, 241, 242] As shown in **Eq (1.2)** ($K_d = k_{off}/k_{on} = TL/C$), K_d

can be defined through either the rate constants (k_{on} and k_{off}) or equilibrium concentrations of the target (T), ligand (L), and complex (C).

Lower K_d values correspond to higher thermodynamic stability of the complex. The equilibrium concentrations are linked to each other through the formal total concentrations of the target (T_0) and ligand (L_0) according to the rules of mass balance:

$$T_0 = T + C, \quad L_0 = L + C \tag{3.10}$$

T_0 and L_0 are concentrations of target and ligand, respectively, after mixing solutions of these two reactants but before any complex is formed. Experimentally, T, L, and C cannot be directly measured to determine K_d. However, K_d can be calculated with known T_0, L_0, and the fraction of unbound ligand R; the latter links K_d to T_0 and L_0.

3.2.2.2. Fraction R of Unbound Ligand

The determination of K_d requires experimentally finding a fraction R of target-unbound ligand (or a fraction of target-bound ligand, $1 - R$), in the equilibrium mixture of the target and ligand. The definition of R is demonstrated in **Eq (2.1)**. The general procedure for finding R starts with preparing the equilibrium mixture of the ligand and target. The procedure is facilitated by the physical or spectral separation of the unbound ligands from the complex. The separation is complete if the peaks or spectra of the unbound ligand and the complex do not overlap; otherwise, it is incomplete. The choice of a signal-processing approach depends on whether or not the separation is complete.

Complete separation of the unbound ligand from the target-ligand complex allows one to measure two signals from the equilibrium mixture — one is a signal from the unbound ligand (S_L) and the other is a signal from the target-ligand complex (S_C) — and express R using the two signals:[101, 243, 244]

$$R = \frac{L}{L_0} = \frac{S_L}{S_L + S_C/\varphi} \tag{3.11}$$

where φ is a factor by which the signal of the ligand changes when it binds the target, *e.g.*, in the case of fluorescence detection, φ is the quantum yield of the target-bound ligand relative to that of the unbound ligand.

Only a cumulative signal S from the unbound ligand and the target-bound ligand in the equilibrium mixture can be measured when their separation is incomplete. Typically, R is determined using additive signals (signal additivity is explained in detail in Note B2.1 in **Appendix B**), which satisfy the principle of superposition:[67, 76, 171, 245]

$$S = S_L{}^* \frac{L}{L_0} + S_C{}^* \frac{C}{L_0} \tag{3.12}$$

where $S_L{}^*$ and $S_C{}^*$ are signals from equimolar concentrations of pure ligand and pure complex and S is a signal from a mixture containing both unbound ligand and complex of a total concentration equal to that used to measure $S_L{}^*$ and $S_C{}^*$. Note that, signal being linearly dependent on concentration is a sufficient condition for the establishment of **Eq (3.12)** (Note B2.2). In the case of incomplete separation of signals from the unbound and target-bound signals, the value of R can still be determined, but three signals, S, $S_C{}^*$, and $S_L{}^*$, are required instead of two signals in **Eq (3.12)**:[67, 76, 171, 246]

$$R = \frac{S - S_C{}^*}{S_L{}^* - S_C{}^*} \tag{3.13}$$

Measuring signal S for the mixture is trivial. Measuring signal $S_L{}^*$ for pure ligand is also trivial as $S_L{}^*$ is the signal from the ligand in the absence of target. Measuring signal $S_C{}^*$ from the pure complex in the case of incomplete separation requires that $C >> L$, which is achieved *via* using a saturating total concentration of the target in the equilibrium mixture:

$$T_0 > L_0, \quad T_0 >> K_d \tag{3.14}$$

It is hard to ensure the second inequality is satisfied as (true) K_d is fundamentally unknown. Moreover, target solubility and a significant increase in the viscosity of the sample due to excessively high T_0 may limit the ability to find S_C^*. Accordingly, complete separation of the free and target-bound ligands is preferable to incomplete separation.

3.2.2.3. General Approach to K_d Determination

There are two major ways to find K_d when R has been determined. In the first way, **Eq (3.1)**, the formula that explicitly relates K_d to R is used.[247] This formula is obtained by solving **Eq (1.2)** using mass balance equations for both ligand and target (**Eq (3.10)**), and the definition of R (**Eq (2.1)**). This approach requires a single value of R. Calculation of K_d with this formula is most accurate for R approximately equal to 0.5.[101] However, K_d values determined this way are very sensitive to random errors of R, T_0, and L_0, especially when R is closer to zero or unity than to 0.5 (Note B2.3). Therefore, an alternative approach, which relies on nonlinear regression, is typically used.

In the nonlinear-regression approach, K_d is determined by using multiple values of R found for a constant value of L_0 but different values of T_0. The dependence of R on T_0 is called a binding isotherm, which exhibits a characteristic sigmoidal shape in a semilogarithmic presentation of the isotherm (dots in **Figure 3.4a**). K_d is then computed by fitting the binding isotherm with a theoretical dependence of R on T_0 (aka, nonlinear-regression model) while using K_d as a fitting parameter. A correct model (**Eq (2.2)**, $R = -\left(K_d + T_0 - L_0\right)/2L_0 + \sqrt{\left[\left(K_d + T_0 - L_0\right)/2L_0\right]^2 + K_d/L_0}$) can be obtained by solving **Eq (3.1)** for R without any simplifying assumption.[19, 20, 32] An example of the best fit of a binding isotherm

obtained with **Eq (2.2)** is shown as a red line in **Figure 3.4a**. In the nonlinear-regression procedure, K_d is a determined parameter, L_0 and T_0 are independent variables, and R is a dependent variable.

The regression model presented by **Eq (2.2)** (as well as other regression models) assumes that only random errors are present in the variables, *i.e.*, R, T_0, and L_0. Those errors result in a random error of $K_{d,det}$ which is reported at the output of nonlinear regression as a standard deviation σ. The standard deviation (σ) of $K_{d,det}$ along with the mean value of $K_{d,det}$ indicates the precision of $K_{d,det}$. Importantly, it is not necessary to repeat the determination of R for the same value of T_0 to assess the random error of $K_{d,det}$.

If any of R, T_0, or L_0 has a systematic error, then K_d cannot be determined accurately with a model presented by **Eq (2.2)**. Furthermore, if systematic errors exist in R, T_0, or L_0 and their magnitudes are unknown, then not only is $K_{d,det}$ inaccurate, but also the systematic error in $K_{d,det}$ remains unknown. On the other hand, if some information about the accuracy of R, T_0, or L_0 is known then the correct regression model presented by **Eq (2.2)** can in principle help assess the accuracy of $K_{d,det}$.

3.2.2.4. Propagation of Systematic Errors

The accuracy of $K_{d,det}$ depends on the accuracies of L_0, T_0, and R. If the values of L_0, T_0, and R were known accurately (*i.e.*, had no systematic errors), then K_d determination would be accurate (*i.e.*, subject to random errors only). However, this is not the case, and systematic errors in L_0, T_0, and R (designated as ΔL_0, ΔT_0, and ΔR) always exist and translate to the systematic error in $K_{d,det}$ (designated as ΔK_d) as explained below assuming (for now) that ΔL_0, ΔT_0, and ΔR are known.

The manner in which ΔL_0, ΔT_0, and ΔR are translated into ΔK_d is governed by the error-propagation rules. Such rules, in turn, depend (though not critically) on whether ΔL_0, ΔT_0, and

ΔR are strongly or weakly correlated. If they are strongly correlated (which will be the case if similar procedures are used for the preparation of solutions of ligand and target, and ΔR is solely the consequence of ΔL_0 and ΔT_0), then we can approximate the absolute value of ΔK_d by **Eq (3.2)**. Recall that we can apply **Eq (3.2)** to the dependence of K_d on L_0, T_0, and R (**Eq (3.1)**) and obtain $|\Delta K_d/K_d| = a + b(L_0/K_d)$ (**Eq (3.3)**), where a is a constant depending only on $|\Delta T_0/T_0|$ and $|\Delta R/R|$ while b is a constant depending on all three relative errors: $|\Delta T_0/T_0|$, $|\Delta L_0/L_0|$, and $|\Delta R/R|$.[179]

If ΔL_0, ΔT_0, and ΔR are weakly correlated (which will be the case if different procedures are used for the preparation of solutions of ligand and target, and ΔR is independent of ΔL_0 and ΔT_0), then we can approximate the absolute value of ΔK_d by applying **Eq (3.5)** (error-propagation rule for fully uncorrelated errors) to **Eq (3.1)** and obtain $\left|\Delta K_d/K_d\right| = \sqrt{\alpha^2 + \lambda\left(L_0/K_d\right) + \beta^2\left(L_0/K_d\right)^2}$ (**Eq (3.6)**), where α and λ are constants depending only on $|\Delta T_0/T_0|$ and $|\Delta R/R|$ while β is a constant depending on all three relative errors: $|\Delta T_0/T_0|$, $|\Delta L_0/L_0|$, and $|\Delta R/R|$. In propagating the errors to obtain **Eqs (3.3)** and **(3.6)**, a single simplifying assumption was made, that the determined (from measured signals) value of R is equal to 0.5, which leads to the least erroneous $K_{d,det}$.[101, 179] Thus, we are considering the best-case scenario, and **Eqs (3.3)** and **(3.6)** represent the lower limit for relative systematic errors in $K_{d,det}$.

Although **Eqs (3.3)** and **(3.6)** appear different, they are similar in the triphasic shape of $|\Delta K_d/K_d|$ dependence on L_0/K_d (**Figure 3.4b**). The first phase corresponds to low L_0/K_d values and is linear with negligible dependence of $|\Delta K_d/K_d|$ on L_0/K_d. The reason for the first phase to be virtually parallel to the x-axis is that the first term prevails over the second in **Eq (3.3)** and the first two terms prevail over the third term in **Eq (3.6)** for low values of L_0/K_d. For phase 3, which

corresponds to large L_0/K_d values, **Eqs (3.3)** and **(3.6)** approach another linear phase: $|\Delta K_d/K_d| = b(L_0/K_d)$ and $|\Delta K_d/K_d| = \beta(L_0/K_d)$, respectively; that phase shows a high sensitivity of $|\Delta K_d/K_d|$ to L_0/K_d. There is a non-linear transition range (phase 2) between the two linear phases. Moreover, both **Eqs (3.3)** and **(3.6)** suggest that the minimum $|\Delta K_d/K_d|$ value (in phase 1) depends on a single parameter (a in **Eq (3.3)** or α in **Eqs (3.6)**) that is defined only by $|\Delta T_0/T_0|$ and $|\Delta R/R|$, while the sensitivity of $|\Delta K_d/K_d|$ to L_0/K_d in phase 3 mainly depends on a parameter (b in **Eq (3.3)** or β in **Eq (3.6)**) that is defined by relative errors of all three variables: $|\Delta T_0/T_0|$, $|\Delta L_0/L_0|$, and $|\Delta R/R|$.

As suggested by **Eqs (3.3)** and **(3.6)**, to increase the accuracy of $K_{d,det}$ (*i.e.*, to decrease $|\Delta K_d/K_d|$), one can decrease the ligand concentration L_0 (to reduce the L_0/K_d ratio) and/or minimize the relative systematic errors in L_0, T_0, and R. Since L_0 cannot be lower than the LOQ of an instrument utilized for measuring signals (and, thus, finding the values of R),[179] it is not practical to decrease L_0/K_d to a very low value in many cases, especially in the studies of very stable complexes (small values of true K_d). Thus, understanding the sources of systematic errors in L_0, T_0, and R and how these errors influence the accuracy of $K_{d,det}$ is crucial for minimizing such systematic errors and further improving the accuracy of $K_{d,det}$.

3.2.3. Results and Discussion

3.2.3.1. The Effect of Incorrect Regression Model on K_d Accuracy

Before considering sources of systematic errors in L_0, T_0, and R, we would like to demonstrate the importance of using a correct regression model (*e.g.*, **Eq (2.2)**). It is common in the research community to fit binding isotherms with a simplified hyperbolic equation:[34, 180, 181]

$$R = \frac{K_d}{K_d + T_0} \tag{3.15}$$

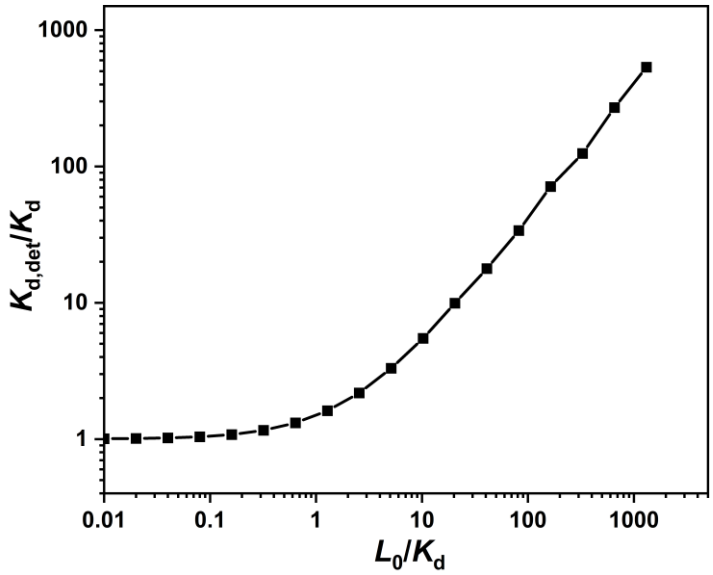


Figure 3.5. The dependence of $K_{d,det}/K_d$ on L_0/K_d when using the simplified hyperbolic equation (**Eq (3.15)**) to fit the theoretical binding isotherms that were simulated with **Eq (2.2)**. There was no error added to L_0, T_0, and R.

This equation is derived from the definitions of K_d and R (**Eqs (1.2)** and **(2.1)**) and mass balance for ligand and target (**Eq (3.10)**) with assuming $T \approx T_0$.[34, 180, 181] **Eq (3.15)** is only applicable to K_d determination under the condition of $L_0 << K_d$;[34, 180, 181] however, since true K_d is fundamentally unknown and L_0 cannot be below the LOQ of the instrument, it is hard to ensure that the condition of $L_0 << K_d$ is achieved. In essence, despite the simplicity of **Eq (3.15)**, it is an incorrect (or inaccurate) regression model, and using it will necessarily result in systematic errors in $K_{d,det}$ even if L_0, T_0, and R are free of systematic errors. It is, therefore, instructive to demonstrate the extent of inaccuracy of $K_{d,det}$ caused by the use of **Eq (3.15)** instead of **Eq (2.2)** in nonlinear regression (**Figure 3.5**).

Importantly, using **Eq (3.15)** leads to an overestimation of $K_{d,det}$ and can result in $L_0 < K_{d,det}$ even when L_0 is greater than the true K_d. This emphasizes the importance of not comparing L_0 to $K_{d,det}$, as it is not only incorrect but also misleading. Our results strongly suggest that the correct

model (**Eq (2.2)**) must be used for nonlinear regression if both T_0 and L_0 are known. Exceptions include cases where determining what L_0 represents is challenging, making the use of **Eq (2.2)** impossible.[34] Classic examples are methods with surface-immobilized ligand, such as BLI and SPR. Therefore, using the correct regression model is crucial for the accurate determination of K_d. Below, we will only use **Eq (2.2)**.

3.2.3.2. Sources of Systematic Errors in R

Systematic errors in $K_{d,det}$ can be attributed to systematic errors in R, which can originate from various sources. Here, we analyze four common sources of ΔR, which are: (*i*) using non-additive signals to calculate R, (*ii*) using an uncalibrated signal-detection instrument or incorrectly determined calibration parameters in calculating R, (*iii*) using insufficient incubation time without allowing mixtures to reach equilibrium, and (*iv*) measuring S_C^* (signal for pure complex) without satisfying saturation condition described in **Eq (3.14)**. Most of these error sources result from theoretical or experimental mistakes and should be addressed during the experimental design and preliminary experiments. Some other less common but well-studied error sources, such as disturbing equilibrium by (*i*) physical separation of the target-ligand complexes from the unreacted ligand, (*ii*) change of conditions (*e.g.*, temperature) caused by detection, *etc*.,[96, 180, 248, 249, 250, 251] are not presented and discussed in this work.

In this manuscript, we consider only the universal origins of inaccuracies in R. However, each method has its own specific sources of inaccuracies. While we do not delve into these method-specific sources, we strongly recommend that practitioners thoroughly study the limitations and potential inaccuracies associated with their chosen methods. Comprehensive analyses of these limitations and sources of inaccuracy are available in the literature for all widely used methods of K_d determination. Below, we provide references to some useful materials for several methods

that we subjectively view as most widely used in the research community of Chemical Biologists: Enzyme-Linked Immuno-Sorbent Assays (ELISA),[252, 253, 254] Surface Plasmon Resonance (SPR),[40, 255, 256] Bio-Layer Interferometry (BLI),[50, 257, 258] Microscale Thermophoresis (MST), [259, 260, 261] Isothermal Titration Calorimetry (ITC), [60, 262 263] and Nuclear Magnetic Resonance (NMR). [264, 265, 266]

3.2.3.2.1. Non-Additive Signals

Systematic errors in R can arise from the use of a non-additive signal, which does not satisfy **Eq (3.12)** and, accordingly, **Eq (3.13)** (derived from **Eq (3.12)**) cannot be used to calculate R. Examples of K_d-determination approaches using signals that have been proven to be additive/non-additive are shown in **Table 3.1**. Note that the additivity status of some signals listed in **Table 3.1** is proven in this work while we do not have such information for Back-Scattering Interferometry (BSI). **Table 3.1** serves as a starting point for reviewing the additive characters of the signals used in various K_d-determination approaches. Researchers in the field of K_d determination are invited to comment on and/or correct this information on https://www.researchgate.net/post/Additivity_of_signals_used_in_Kd-determination_approaches, where we will upload the updated table regularly. Here, considering examples of non-additive signals within the context of specific K_d-determination approaches is instructive. Our results indicate that a signal must be proven to be additive before it can be used to calculate R. Otherwise, large systematic errors in $K_{d,det}$ might be caused by mistakenly using non-additive signals in R calculations.

Mobility-Based Methods:

A method commonly used for K_d determination for relatively unstable complexes is based on mobilities of unbound ligand and target-bound ligand in capillary electrophoresis (CE).[88] In this

Table 3.1. Examples of K_d-determination approaches with information of signals used to calculate R and their additivity.

Approach	Signal	Additivity*	References with proof†	Supplementary references
Accurate Constant *via* Transient Incomplete Separation (ACTIS)	Fluorescence intensity	✓	[1]	[2], [3]
	Ion-signal intensity	✓		
Affinity Capillary Electrophoresis (ACE)	Electrophoretic mobility	✓	[4]	[5], [6]
	Migration velocity (with constant mobility of electroosmotic flow)	✓		
	Migration time	✗	"Mobility-Based Methods" in *Section 3.2.3.2.1*	N/A
Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures (NECEEM)	Peak area	✓	[7]	[8], [9]
Back-Scattering Interferometry (BSI)	Refractive index in solution	?	N/A	[10], [11]
Surface Plasmon Resonance (SPR)	Refractive index on surface	✓	[12]	[13], [14], [15]
Bio-Layer Interferometry (BLI)	Wavelength shift	✓	[16]	[17], [18]
Enzyme-Linked Immuno-Sorbent Assays (ELISA)	Absorbance	✓	[19]	[20], [21]
Solid Phase Radio-Immuno-Assay (SPRIA)	Radioactivity	✓	[22]	[23], [24]
Flow Induced Dispersion Analysis (FIDA)	Apparent diffusion coefficient	✗	"Diffusivity-Based Methods" in *Section 3.2.3.2.1*	[25], [26]
Taylor Dispersion Analysis (TDA)	Reciprocal of apparent diffusion coefficient	✗	"Diffusivity-Based Methods" in *Section 3.2.3.2.1*	[27]
Rotational Depolarization of Fluorescence	Anisotropy	✓	[28]	[29], [30]
	Polarization	✗	[31], Note B2.4	[32], [33], [34]
Fluorescence Resonance Energy Transfer (FRET)	Fluorescence intensity	✓	[35]	[36]
High-Performance Liquid Chromatography (HPLC)	Peak area	✓	[37]	[38]
Isothermal Titration Calorimetry (ITC)	Binding enthalpy	✓	[39]	[40], [41]
Microscale Thermophoresis (MST)	Fluorescence intensity	✓	[42]	[43]
Native Mass Spectrometry (MS)	Ion signal intensity	✓	[44]	[45], [46]
Nuclear Magnetic Resonance (NMR)	Chemical shift	✓	[47]	[48], [49], [50]
	Longitudinal and transverse relaxation rates	✗	[51]	[52]
Electrophoretic Mobility Shift Assay (EMSA)	Radioactivity or light/fluorescence intensity	✓	[53]	[54], [55]
Gel Filtration	Absorbance	✓	[56]	[57]

*Symbol "✓" indicates that a proof of additivity exists, symbol "✗" indicates that the proof of non-additivity exists, and symbol "?" indicates that the authors did not find a proof of additivity or non-additivity.

† In some references, it is not directly shown that whether the measured signals satisfy **Eq (3.12)** or not. In these cases, if the measured signal is proven to be proportional to the quantities (*e.g.*, concentrations, layer thicknesses, *etc.*) of target-unbound and target-bound ligands, we consider that the additivity of the signal is proven (refer to discussion in Notes B2.1

References for Table 3.1:

[1] Sisavath, N.; Rukundo, J. L.; Le Blanc, J. Y.; Galievsky, V. A.; Bao, J.; Kochmann, S.; Stasheuski, A. S.; Krylov, S. N. *Angew. Chem.* **2019**, *131* (20), 6707-6711.

[2] Wang, T. Y.; Rukundo, J.-L.; Le, A. T.; Ivanov, N. A.; Le Blanc, J. Y.; Gorin, B. I.; Krylov, S. N. *Anal. Chem.* **2022**, *94* (44), 15415-15422.

[3] Rukundo, J.-L.; Kochmann, S.; Wang, T. Y.; Ivanov, N. A.; Le Blanc, J. Y.; Gorin, B. I.; Krylov, S. N. *Anal. Chem.* **2021**, *93* (34), 11654-11659.

[4] Gomez, F. A.; Avila, L. Z.; Chu, Y. H.; Whitesides, G. M. *Anal. Chem.* **1994**, *66* (11), 1785-1791.

[5] Dubský, P.; Dvořák, M.; Ansorge, M. *Anal. Bioanal. Chem.* **2016**, *408*, 8623-8641.

[6] Chu, Y. H.; Whitesides, G. M. *J. Org. Chem.* **1992**, *57* (13), 3524-3525.

[7] Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2002**, *124* (46), 13674-13675.

[8] Kanoatov, M.; Galievsky, V. A.; Krylova, S. M.; Cherney, L. T.; Jankowski, H. K.; Krylov, S. N. *Anal. Chem.* **2015**, *87* (5), 3099-3106.

[9] Berezovski, M.; Drabovich, A.; Krylova, S. M.; Musheev, M.; Okhonin, V.; Petrov, A.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127* (9), 3165-3171.

[10] Bornhop, D. J.; Latham, J. C.; Kussrow, A.; Markov, D. A.; Jones, R. D.; Sørensen, H. S. *Science* **2007**, *317* (5845), 1732-1736.

[11] Latham, J. C.; Stein, R. A.; Bornhop, D. J.; Mchaourab, H. S. *Anal. Chem.* **2009**, *81* (5), 1865-1871.

[12] Stenberg, E.; Persson, B.; Roos, H.; Urbaniczky, C. *J. Colloid Interface Sci.* **1991**, *143* (2), 513-526.

[13] Blackburn, C.; Sullivan, M. V.; Wild, M. I.; O' Connor, A. J.; Turner, N. W. *Anal. Chim. Acta* **2024**, *1285*, 342004.

[14] Stahelin, R. V. *Mol. Biol. Cell* **2013**, *24* (7), 883-886.

[15] Gasse, C.; Zaarour, M.; Noppen, S.; Abramov, M.; Marlière, P.; Liekens, S.; De Strooper, B.; Herdewijn, P. Modulation of BACE1 activity by chemically modified aptamers. *ChemBioChem* **2018**, *19* (7), 754-763.

[16] Apiyo, D.O. Biolayer Interferometry (Octet) for Label-free Biomolecular Interaction Sensing. In *Handbook of Surface Plasmon Resonance*, 2nd ed.; Schasfoort, R.B.M., Ed.; Royal Society of Chemistry: London, UK, 2017; pp. 356–397.

[17] Sultana, A.; Lee, J. E. *Curr. Protoc. Protein Sci.* **2015**, *79* (1), 19-25.

[18] Petersen, R. L. *Biosensors* **2017**, *7* (4), 49.

[19] Friguet, B.; Chaffotte, A. F.; Djavadi-Ohaniance, L.; Goldberg, M. E. *J. Immunol. Methods* **1985**, *77* (2), 305-319.

[20] Orosz, F.; Ovádi, J. *J. Immunol. Methods* **2002**, *270* (2), 155-162.

[21] Kim, B.; Dikova, E.; Sheller, U.; Dikov, M.; Gavrilova, E.; Egorov, A. Evaluation of dissociation constants of antigen-antibody complexes by ELISA. *J. Immunol. Methods* **1990**, *131* (2), 213-222.

[22] Ekins, R. P.; Newman, G. B.; Piyasena, R.; Banks, P.; Slater, J. D. H. *J. Steroid Biochem.* **1972**, *3* (3), 289-304.

[23] Murthy, G.; Venkatesh, N. *J. Biosci.* **1996**, *21*, 641-651.

[24] Lessard, J. L.; Carlton, D.; Rein, D. C.; Akeson, R. *Anal. Biochem.* **1979**, *94* (1), 140-149.

[25] Jensen, H.; Østergaard, J. *J. Am. Chem. Soc.* **2010**, *132* (12), 4070-4071.

[26] Clark, S. M.; Leaist, D. G.; Konermann, L. *Rapid Commun. Mass Spectrom.* **2002**, *16* (15), 1454-1462.

[27] Cottet, H.; Biron, J.-P.; Martin, M. *Anal. Chem.* **2007**, *79* (23), 9066-9073.

[28] Gradinaru, C. C.; Marushchak, D. O.; Samim, M.; Krull, U. J. *Analyst* **2010**, *135* (3), 452-459.

[29] Jameson, D. M.; Ross, J. A. *Chem. Rev.* **2010**, *110* (5), 2685-2708.

[30] Lakowicz, J. R.; Lakowicz, J. R. Fluorescence anisotropy. In *Principles of fluorescence spectroscopy*; Springer, 1999; pp 291-319. DOI: 10.1007/978-1-4757-3061-6_10

[31] Weber, G. *Biochem J* **1952**, *51* (2), 145-155.

[32] Park, S.-H.; Raines, R. T. Fluorescence polarization assay to quantify protein-protein interactions. In *Protein–Protein Interactions: Methods and Applications*, Fu, H. Ed.; Humana Press, 2004; pp 161-165. DOI: 10.1385/1-59259-762-9:161

[33] Nosjean, O.; Souchaud, S.; Deniau, C.; Geneste, O.; Cauquil, N.; Boutin, J. A. *SLAS Discovery* 2006, *11* (8), 949-958.

[34] Yan, K.; Hunt, E.; Berge, J.; May, E.; Copeland, R. A.; Gontarek, R. R. *Antimicrob. Agents Chemother.* 2005, *49* (8), 3367-3372.

[35] Liao, J.-y.; Song, Y.; Liu, Y. *Acta Pharmacol. Sin.* 2015, *36* (12), 1408-1415.

[36] Liu, J.; Lu, Y. *Journal of the American Chemical Society* 2002, *124* (51), 15208-15216.

[37] Sebille, B.; Zini, R.; Madjar, C.-V.; Thuaud, N.; Tillement, J.-P. *J. Chromatogr. B: Biomed. Sci. Appl.* 1990, *531*, 51-77.
[38] Lowe, E. A.; Lu, M.; Wang, A.; Cortez, H.; Ellis, D.; Liu, X. C. *J. Sep. Sci.* 2006, *29* (7), 959-965.
[39] Equations Used for Fitting ITC Data. *ITC Data Analysis in Origin – Tutorial Guide*; MicroCal: Northampton, MA, 2004; pp 104–108.
[40] Neves, M. A. D.; Slavkovic, S.; Churcher, Z. R.; Johnson, P. E. *Nucleic Acids Res.* 2017, *45* (3), 1041-1048.
[41] Neves, M. A. D.; Shoara, A. A.; Reinstein, O.; Abbasi Borhani, O.; Martin, T. R.; Johnson, P. E. *ACS Sens.* 2017, *2* (10), 1539-1545.
[42] Baaske, P.; Wienken, C. J.; Reineck, P.; Duhr, S.; Braun, D. *Angew. Chem., Int. Ed.* 2010, *49* (12), 2238-2241.
[43] Jerabek-Willemsen, M.; Wienken, C. J.; Braun, D.; Baaske, P.; Duhr, S. *Assay Drug Dev. Technol.* 2011, *9* (4), 342-353.
[44] El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S. *Anal. Chem.* 2012, *84* (9), 3867-3870.
[45] Jafari, M. R.; Deng, L.; Kitov, P. I.; Ng, S.; Matochko, W. L.; Tjhung, K. F.; Zeberoff, A.; Elias, A.; Klassen, J. S.; Derda, R. *ACS Chem. Biol.* 2014, *9* (2), 443-450.
[46] Tjernberg, A.; Carnö, S.; Oliv, F.; Benkestock, K.; Edlund, P.-O.; Griffiths, W. J.; Hallén, D. *Anal. Chem.* 2004, *76* (15), 4325-4331.
[47] Rabenstein, D. L.; Sayer, T. L. *Anal. Chem.* 1976, *48* (8), 1141-1146.
[48] Fielding, L. *Prog. Nucl. Magn. Reson. Spectrosc.* 2007, *51* (4), 219-242.
[49] Shoara, A. A.; Churcher, Z. R.; Slavkovic, S.; Johnson, P. E. *ACS Omega* 2021, *6* (37), 24209-24217.
[50] Johnson, P. E.; Tomme, P.; Joshi, M. D.; McIntosh, L. P. *Biochemistry* 1996, *35* (44), 13895-13906.
[51] Fushman, D.; Cahill, S.; Cowburn, D. *J. Mol. Biol.* 1997, *266* (1), 173-194.
[52] Monasterio, O. *Methods* 2001, *24* (2), 97-103.
[53] Fried, M. G. *Electrophoresis* 1989, *10* (5-6), 366-376.
[54] Hellman, L. M.; Fried, M. G. *Nat. Protoc.* 2007, *2* (8), 1849-1861.
[55] Hsieh, Y.-W.; Alqadah, A.; Chuang, C.-F. *J. Vis. Exp.* 2016, (117), e54863.
[56] Hummel, J.; Ver Ploeg, D. A.; Nelson, C. A. *J. Biol. Chem.* 1961, *236* (12), 3168-3172.
[57] Hummel, J. P.; Dreyer, W. *Biochim. Biophys. Acta* 1962, *63* (3), 530-532.

method, a short plug of the ligand of concentration L_0 is injected into a capillary prefilled with the running buffer containing the target at a concentration of T_0. This plug of the ligand moves through the capillary by an electric field under the condition of pseudo-equilibrium in binding process (**Eq (1.1)**) which is equivalent to a condition that the characteristic equilibration time (t_{eq}) is much smaller than the characteristic separation time (t_{sep}):

$$t_{eq} << t_{sep} \tag{3.16}$$

The characteristic equilibration time for a case of $T_0 >> L_0$ can be approximated by:[248]

$$t_{eq} \approx \left(k_{off} + k_{on}T_0\right)^{-1} \tag{3.17}$$

The characteristic separation time is:

$$t_{sep} = \frac{l}{\min\left(\left|v_T - v_L\right|, \left|v_C - v_L\right|\right)} \tag{3.18}$$

where l is the length of the plug of ligand, and v_T, v_L, and v_C are velocities of the target, ligand, and complex, respectively. Usually, the complex moves with the velocity intermediate to those of the unbound ligand and target, that is $|v_C - v_L| < |v_T - v_L|$, and **Eq (3.18)** can be written as:

$$t_{sep} = \frac{l}{\left|v_C - v_L\right|} \tag{3.19}$$

The migration time of the peak of ligand to the detector t is a common measurable signal in capillary electrophoresis (CE),[267] but time is not additive, while if the mobility of electroosmotic flow is independent of T_0, the velocity and inverse migration time are both additive: [88, 268, 269]

$$\begin{aligned} v &= v_L \frac{L}{L_0} + v_C \frac{C}{L_0} \\ \frac{1}{t} &= \frac{1}{t_L}\frac{L}{L_0} + \frac{1}{t_C}\frac{C}{L_0} \end{aligned} \tag{3.20}$$

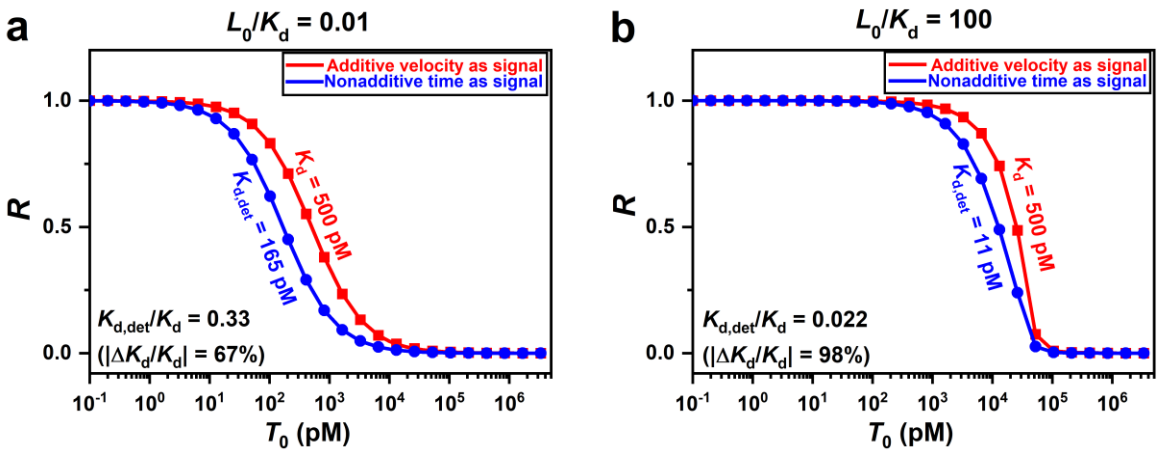


Figure 3.6. Comparison of the results produced by using the additive velocity (v) and non-additive time (t) as signals: binding isotherms and $K_{d,det}$ obtained with signals of velocity (red) and time (blue) for L_0/K_d = 0.01 **(a)** and L_0/K_d = 100 **(b)**. In the simulation, the theoretical/input K_d = 500 pM.

Binding isotherms with both time and velocity as a signal have characteristic and undistinguishable sigmoidal shapes; however, they are shifted with respect to each other (**Figure 3.6**). If one wrongly uses time instead of inverse time to calculate R with **Eq (3.13)**, then the resulting $K_{d,det}$ will have significant systematic errors at both low and high L_0/K_d values (**Figure 3.6**).

Methods Based on Fluorescence Anisotropy:

Both fluorescence anisotropy (r) and polarization (P) are parameters that quantitatively describe the degree of light polarization in different contexts.[270] Fluorescence anisotropy was introduced specifically to characterize light emitted by a fluorophore in three dimensions while polarization was introduced to characterize polarization of a beam of collimated light from a light source in a plane perpendicular to the beam. Thus, fluorescence anisotropy and beam-light

polarization are applied to different dimensionalities which are reflected in formulas expressing r and P through the two component light intensities, factoring in the total light intensity.

Fluorescence anisotropy is defined as follows:[270]

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \tag{3.21}$$

where $I_{\parallel}$ is the intensity of light emitted by a fluorophore with a polarization orientation parallel to that of the excitation light; $I_{\perp}$ is the intensity of emitted light with a polarization orientation perpendicular to that of the excitation light. The denominator of **Eq (3.21)** contains a factor of 2 because there are two symmetrical dimensions perpendicular to the direction of excitation-light polarization, but only the intensity associated with one of them (*i.e.*, $I_{\perp}$) is typically measured. The denominator represents the total intensity of light in the three-dimensional space.

Beam-light polarization is defined as follows:[271]

$$P = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{3.22}$$

where $I_{\max}$ is the intensity of light polarized in the direction which corresponds to maximum intensity; $I_{\min}$ is the intensity of light in the direction perpendicular to the first one. The denominator also represents the total intensity of light but in two-dimensional space. **Eq (3.22)** is sometimes written using $I_{\parallel}$ instead of $I_{\max}$ and $I_{\perp}$ instead of $I_{\min}$:[270]

$$P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}} \tag{3.23}$$

which is legitimate provided that we understand that P describes a two-dimensional case due to the missing factor of 2 in the denominator.

Formally, anisotropy and polarization calculated with **Eq (3.21)** and **Eq (3.23)**, respectively, are closely related and interconvertible with:[270]

$$r = \frac{2P}{3-P} \tag{3.24}$$

Anisotropy has been proven to be an additive signal:[272]

$$r = r_L \frac{L}{L_0} + r_C \frac{C}{L_0} \tag{3.25}$$

and thus can be used to correctly calculate R with **Eq (3.13)** and, then, build an accurate binding isotherm and accurately determine K_d with **Eq (2.2)**.[273] However, by combining the interconversion formula (**Eq (3.24)**) and the proof of anisotropy additivity (**Eq (3.25)**), we can easily conclude that P is non-additive (Note B2.4) and, thus, cannot be directly used to calculate R. The non-additivity of P (without transformations) was explained decades ago,[273, 274] but it is still mistakenly used instead of r in **Eq (3.13)** for finding R and K_d determination.

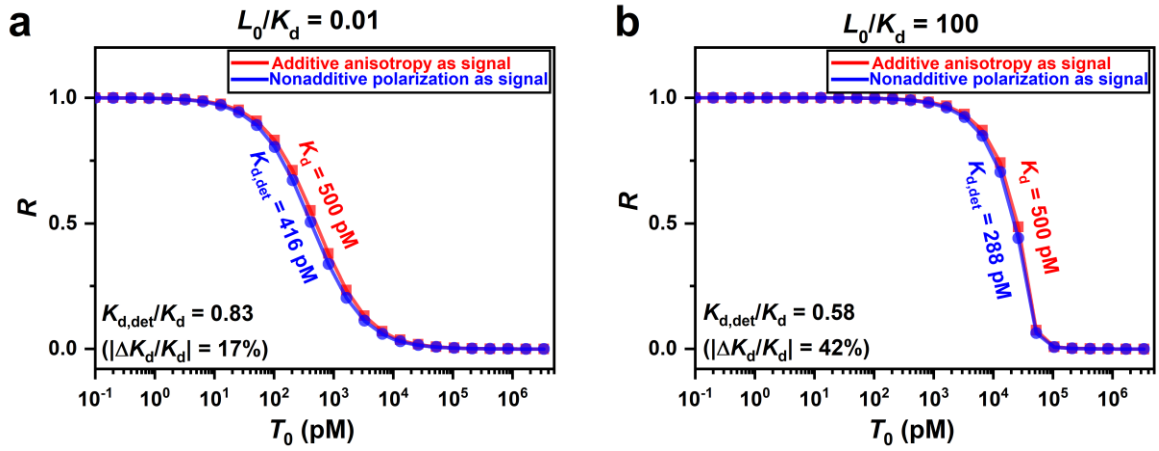


Figure 3.7. Comparison of the results produced by additive anisotropy (r) and non-additive polarization (P): binding isotherms and $K_{d,det}$ obtained with anisotropy (red) and polarization (blue) at L_0/K_d = 0.01 **(a)** and L_0/K_d = 100 **(b)**. In the simulation, the theoretical/input K_d = 500 pM.

Additionally, some commercial instruments provide polarization (P) as a default output,[275, 276] which can mislead users into calculating R based on P. Such a mistake can only be explained by a widely spread ignorance, as replacing P with r is not only correct but also does not require any changes in instrument hardware.

When one uses the non-additive P instead of additive r to calculate R with **Eq (3.13)**, the binding isotherm will slightly shift due to the systematic errors introduced to R (**Figure 3.7**). Without including any other source of systematic errors, these small shifts of binding isotherm cause < 20% relative systematic error in $K_{d,det}$ for small L_0/K_d. However, the effect of this shift on the accuracy of $K_{d,det}$ grows with increasing L_0/K_d and may result in > 40% relative systematic errors for large L_0/K_d (**Figure 3.7b**, Note B2.5 and **Figure B2.1**). The large discrepancy of $K_{d,det}$ from K_d at high L_0/K_d (Note B2.5, **Figure B2.1**) indicates that non-additive P must not be used to calculate R in K_d determination since the true value of L_0/K_d is unknown *a priori*. If P is the default output of a commercial instrument,[24, 25] it must be converted to r with interconversion formula (**Eq (3.24)**) before calculating R and conducting the standard downstream procedures of K_d determination, such as building binding isotherm and nonlinear regression.

Diffusivity-Based Methods:

A signal from the ligand (used in **Eq (3.13)**) must change upon complex formation. The target–ligand complex is larger than unbound ligand and this size difference creates a foundation for size-dependent signals. A larger size results in slower translational movement leading to velocity as a signal (as discussed in the "*Methods Based on Fluorescence Anisotropy*" section above). Furthermore, a larger size leads to slower diffusion and, accordingly, the apparent diffusion coefficient D_{app} of the ligand was suggested and used as a signal in K_d determination. One of the methods utilizing D_{app} is based on monitoring Taylor dispersion of the ligand

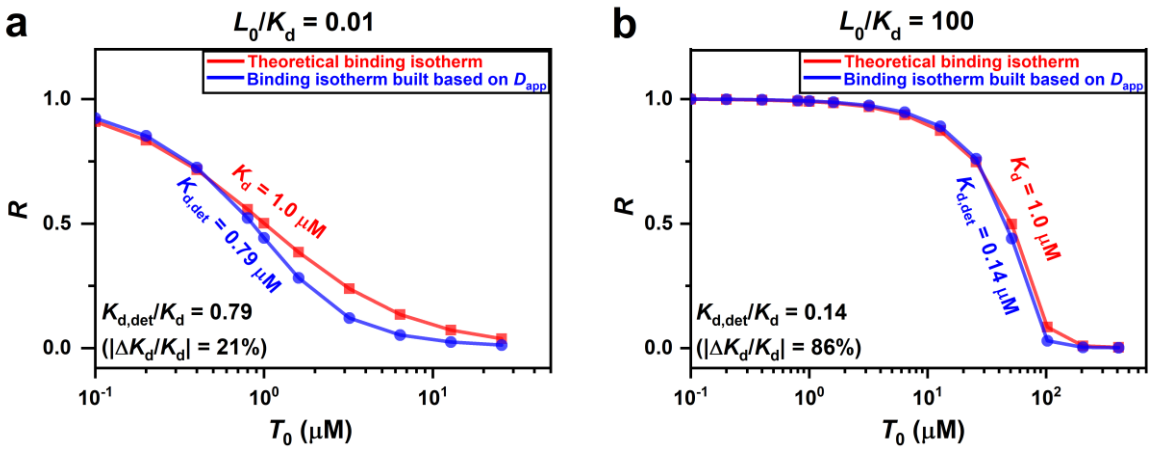


Figure 3.8. Comparison of the results produced by using apparent diffusion coefficient as signal (blue) and theoretical results (red) for $L_0/K_d = 0.01$ **(a)** and $L_0/K_d = 100$ **(b)**.

molecule in a laminar pipe flow within a capillary filled with the target.[277, 278, 279] No inherent sources of inaccuracy have been reported for this approach so far. Here, we analyze D_{app} for additivity, required for its use as signal in **Eq (3.13)** for calculation of R, and prove that D_{app} it is non-additive.

Two geometrically different implementations of finding K_d *via* determination of the apparent diffusion coefficient (by monitoring Taylor dispersion) have been suggested:[277, 278, 279] one method utilizes a short plug of the equilibrium mixture of T and L, while the other uses a continuous flow of the equilibrium mixture. In each method, the apparent diffusion coefficient, D_{app}, is found by fitting the signal proportional to the corresponding function with D_{app} as an unknown parameter.[278, 280] In this work, we will not compare these two implementations of determining D_{app}, but only analyze if D_{app} is additive.

There are four components in the equilibrium mixture: L (unbound ligand), T (unbound target), C (target–ligand complex), and one solvent. If D_{app} is additive, it must satisfy:

$$D_{app} = D_{app,L}\frac{L}{L_0} + D_{app,C}\frac{C}{L_0} \tag{3.26}$$

where $D_{app,L}$ and $D_{app,C}$ are the apparent diffusion coefficients for pure ligand L and pure complex C, respectively, and D_{app} is the apparent diffusion coefficient of L (resulting from both the unbound and target-bound ligand) in the equilibrium mixture.

In **Appendix B2** (Note B2.6 and **Figure B2.2**), we proved the non-additivity of D_{app} with mathematical derivations and numerical simulations. As shown in the results of simulations in **Figure 3.8**, employing the additivity assumption of **Eq (3.26)** results in large systematic errors in $K_{d,det}$, especially for high L_0/K_d (**Figure 3.8b**). Consequently, we consider the diffusion coefficient, along with Taylor-dispersion methods, unsuitable for accurately determining K_d.

Although the apparent diffusion coefficient cannot be directly used to calculate R due to its non-additivity, R can be determined by analyzing an additive signal (*e.g.*, fluorescence) that varies with the fraction of unbound ligand due to the differing diffusivities of L and C. [171] Species with different diffusion coefficients can undergo incomplete separation in laminar flow, a phenomenon known as TIS (transient incomplete separation). [281, 282, 283] Based on the theory of TIS, we developed the method of "accurate constant via transient incomplete separation" (ACTIS) using laser-induced fluorescence (LIF) or mass spectroscopy (MS) as detection methods. [171] ACTIS has been validated to be accurate, robust, and rugged both in computer simulations and with physical instruments. [172, 177, 178]

3.2.3.2.2. Mis-calibration of Signal-detecting Instrument and Inaccurate Calibration Parameters

K_d determination relies on measuring signal changes due to complex formation. For accurate R values, the instrument detecting the ligand must be properly calibrated, ensuring a linear relationship between the detected signal and concentration. Mis-calibrated instruments can introduce systematic errors in R, affecting the accuracy of $K_{d,det}$. The necessity of appropriate calibration for signal-detecting instruments is well-known, and calibration procedures vary from

instrument to instrument.[284, 285, 286] In **Appendix B2** (Note B2.7 and **Figure B2.3**), we demonstrate the effect of a mis-calibrated fluorescence detector on $K_{d,det}$.

Calibration parameters are often used in data analysis to standardize detected signals.[101, 270, 287] Accurate calibration parameters need to be determined *prior to* K_d-determination experiments. Errors in calibration parameters lead to errors in calculated R, impacting the accuracy of $K_{d,det}$. As an example, here we show the effect of errors in one common calibration parameter — φ (the quantum yield of target-bound ligand relative to that of unbound ligand in fluorescence detection) on the accuracy of $K_{d,det}$. Additionally, we discuss methods to enhance the accuracy of the calibration parameter φ.

In K_d-determination approaches detecting changes in fluorescence intensity due to the target's binding to the ligand (where the ligand denotes the component labelled with a fluorophore and of smaller size), fluorescence quenching (*i.e.*, quantum yield change of a fluorophore) often

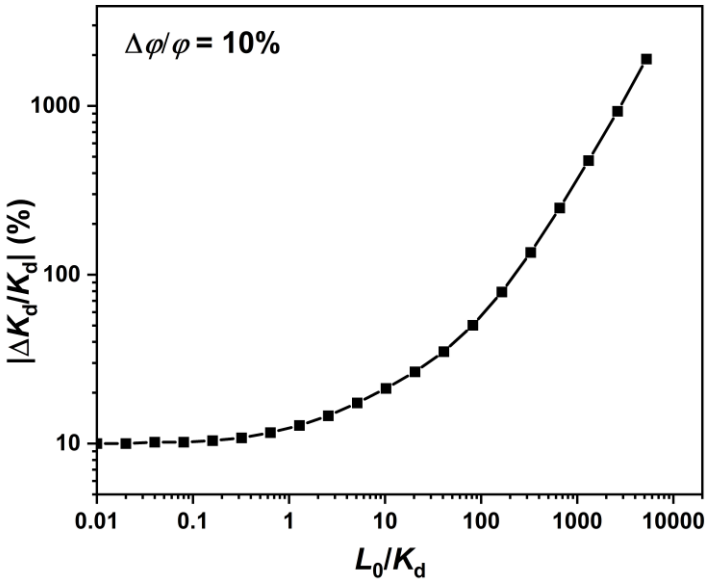


Figure 3.9. The dependence of relative systematic error of $K_{d,det}$ (*i.e.*,$|\Delta K_d/K_d|$) on L_0/K_d with adding 10% relative systematic error in φ (*i.e.*, $\Delta\varphi/\varphi$ = 10%) for a complete-separation approach with fluorescence as the signal.

occurs when the ligand binds to the target.[288] In complete-separation approaches, such as nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM), R is calculated with **Eq (3.11)**, which involves a calibration parameter, φ — the quantum yield of the target-bound ligand relative to that of the unbound ligand.[101, 243, 244] Here, we compare $K_{d,det}$ to K_d by introducing a 10% relative systematic error to φ (*i.e.*, $\Delta\varphi/\varphi = 0.1$) across a wide range of L_0/K_d (**Figure 3.9**). The large relative systematic error in $K_{d,det}$ ($|\Delta K_d/K_d|$) at high L_0/K_d (as depicted in **Figure 3.9**) underscores the importance of accurately determining φ in complete-separation approaches.

In complete-separation approaches, φ is determined by comparing the signal area under the peak for pure complex (S_C^*) to the area for pure ligand (S_L^*):[101, 243, 244]

$$\varphi = \frac{S_C^*}{S_L^*} \tag{3.27}$$

Determining accurate φ necessitates finding S_L^* and S_C^* from pure ligand and pure complex, respectively, with identical concentrations, and accurately measuring the corresponding areas. To meet these requirements, we first need to decide on the concentration of ligand L_0 to use in order to produce S_L^* **Eq (3.27)**. While using the ligand concentration of L_0 = LOQ minimizes systematic errors in $K_{d,det}$ derived from systematic errors in L_0, T_0, and R,[179] we recommend employing a higher ligand concentration of $L_0 \geq 10 \times$ LOQ in the preliminary experiment to determine S_L*. A sample with higher ligand concentration (*e.g.*, $L_0 = 10 \times$ LOQ) can yield a larger and more easily defined area of S_L^* (and S_C^*), leading to more accurate measured signal areas and reducing errors in determined φ.

To measure the area of S_C^* from pure complex with the same concentration, *i.e.*, L_0, the conditions outlined in **Eq (3.14)** must be met to bind all ligands with a total concentration of L_0. To increase the likelihood of meeting the condition of $T_0 >> K_d$ (unknown) (**Eq (3.14)**), the

equilibrium mixture used for measuring S_C^*, should contain the highest possible target concentration T_0, which is the solubility of the target in the sample buffer. Note that the increase in sample viscosity caused by high T_0 would unlikely affect the accuracy of measured S_C^* in complete-separation approaches because signals (peak areas) are independent of viscosities. Furthermore, to ensure the condition of $T_0 > L_0$ (in **Eq (3.14)**) is met, L_0 used in determining S_L^* and S_C^* cannot be excessively high. Therefore, we recommend using $L_0 = 10 \times$ LOQ and $T_0 =$ "solubility of target" as the component concentrations in the equilibrium mixture for measuring S_C^*. It is important to emphasize that all conditions of preliminary experiments, *e.g.*, temperature, sample buffer, detection system, *etc*., must be identical to those in the subsequent K_d-determination experiments to ensure the determined φ is a correct calibration parameter in downstream data analysis.

Unlike complete-separation approaches, incomplete-separation approaches, *e.g.*, ACTIS, do not require the parameter φ to correct signal S_C^* **Eq (3.13)** for calculating R, since the detected fluorescence intensity multiplied with a constant quenching coefficient satisfies the requirement of signal superposition (**Eq (3.12)**).[171] Thus, although fluorescence quenching affects the detected signal S, its presence does not affect the accuracy of $K_{d,det}$ obtained in approaches with incomplete separation of free ligand from the complex.

In **Appendix B2** (Note B2.8), we also investigated the effect of the errors in another method-specific calibration parameter, G (grating factor in fluorescence anisotropy), on the accuracy of $K_{d,det}$ and provided suggestions on improving the accuracy of the experimentally determined G.

3.2.3.2.3. Not Reaching Equilibrium or Saturation

Careful and correct pre-experimental calculations and experimental design are crucial for the accurate determination of K_d. In these preparative steps, researchers should decide on: (*i*) the

most suitable K_d-determination approach for the studied binding pair, (*ii*) experimental conditions, such as temperature and wavelength of the light source, (*iii*) the concentration of each binding partner in each equilibrium mixture, (*iv*) the incubation time for the samples, among other factors. Among these experimental parameters, the total concentrations of binding partners in equilibrium mixtures (*i.e.*, L_0 and T_0) and the incubation time for the samples are usually difficult to choose due to the lack of reliable input information. Here we demonstrate the effect of insufficient incubation time and failure to reach saturation in the binding isotherm (caused by the mis-selection of concentrations) on the accuracy of $K_{d,det}$. We provide suggestions on how to ensure the incubation time is sufficient and how to ensure that the equilibrium mixture with the highest T_0 satisfies the saturation condition.

Not Reaching Equilibrium in the Binding Process:

The first experimental step of the classic methodology of K_d determination is to prepare a set of equilibrium mixtures containing a constant concentration of the limiting component L_0 and varied concentrations T_0 of the target. To minimize the systematic error in $K_{d,det}$ propagated from ΔL_0, ΔT_0, and ΔR, we should choose L_0 = LOQ.[179] In practice, the lowest nonzero T_0 is usually chosen as low as possible (often much lower than L_0),[181] for which the pseudo-first order conditions of $T_0 << L_0$ and $L \approx L_0$ are satisfied. Therefore, the characteristic time t_{eq} (**Eq (3.17)**) of the reversible binding process (**Eq (1.1)**) with the lowest nonzero T_0 is expressed as:

$$t_{eq} \approx \left(k_{off} + k_{on} L_0\right)^{-1} \quad \text{or} \quad t_{eq} \approx \left(k_{off} + k_{on}\mathrm{LOQ}\right)^{-1} \text{ when } L_0 = \mathrm{LOQ} \tag{3.28}$$

which limits (represents the longest) characteristic time for all the binding processes with different T_0 in a single K_d-determination experiment. To reach ≥ 95% of the equilibrium concentration of the complex with the lowest nonzero T_0, the incubation time (t_{inc}) must reach at

least 3 times t_{eq} shown in **Eq (3.28)** (Note B2.9). In studies of molecular pairs with high affinity (*i.e.*, low K_d with high k_{on} and/or low k_{off}) and using low L_0 in all samples, the sufficient incubation time (*i.e.*, $t_{inc} \geq 3t_{eq}$) can be extremely long (*e.g.*, a few days) (Note B2.9). Thus, it is common to overestimate $K_{d,det}$ due to using insufficient incubation time.[34]

To demonstrate the effect of insufficient incubation time on the accuracy of $K_{d,det}$, we simulated the binding isotherms using the spreadsheet Data S1 stored on the Figshare website (https://doi.org/10.6084/m9.figshare.25464685). We utilized the incubation time within a range of $0.1t_{eq}$ to $19t_{eq}$ (t_{eq} was calculated using **Eq (3.28)**) with a 1.5-fold step size, using $L_0/K_d = 0.1$ (**Figure 3.10a**), and summarized the dependence of $K_{d,det}/K_d$ on the incubation time (**Figure 3.10b**). The results of the simulations (**Figure 3.10b**) show that, even with low $L_0/K_d = 0.1$, insufficient incubation time can cause $K_{d,det}$ to be several times higher than the true K_d. As expected, both the binding isotherm and $K_{d,det}$ stabilize when $t_{inc} \geq 3t_{eq}$.

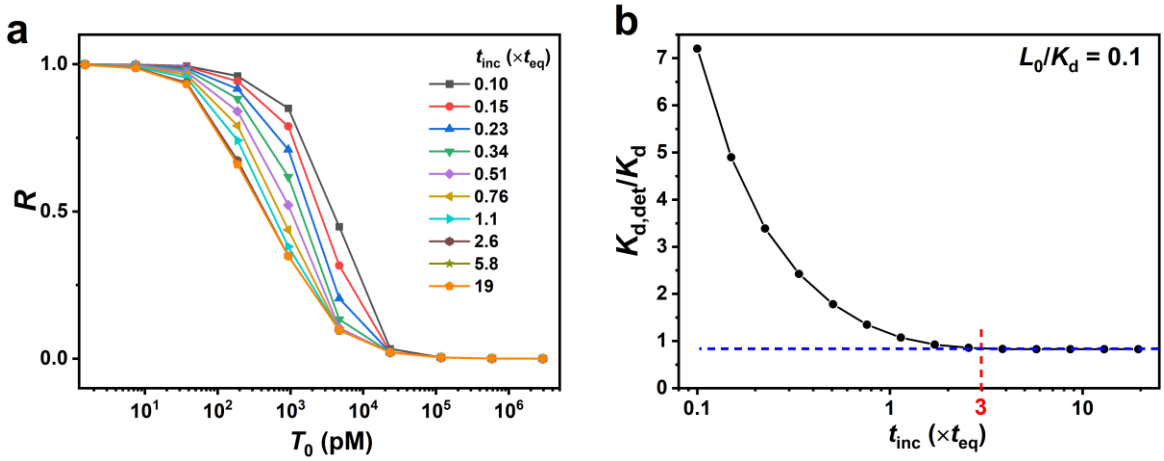


Figure 3.10. The dependencies of representative binding isotherms **(a)** and $K_{d,det}/K_d$ **(b)** on incubation time t_{inc}. In the simulations, L_0/K_d was chosen to be 0.1. Due to the assumption of pseudo-first order conditions (either $T_0 << L_0$ and $L \approx L_0$ or $L_0 << T_0$ and $T \approx T_0$) satisfied by each EM, the stabilized $K_{d,det}$ was slightly smaller than the input K_d, *i.e.*, stabilized $K_{d,det}/K_d$ was slightly smaller than unit, which does not affect estimating the sufficient incubation time.

Although there are some experimental and mathematical approaches to estimate sufficient incubation time (Note B2.10), at this stage, the only reliable method of eliminating the effect of incubation time on the accuracy of $K_{d,det}$ is conducting multiple K_d-determination experiments with increased incubation times until $K_{d,det}$ stabilizes.[34] Note that, as equilibration time is longer at lower T_0, one can plan a sequence of experiments so that equilibrium mixtures with higher T_0 are analyzed before equilibrium mixtures with lower T_0 allowing the latter more time for equilibration.

Not Reaching Saturation in Binding Isotherm:

For K_d-determination approaches with complete separation of unbound ligand from target-bound ligand, although a complete binding isotherm with R ranging from 0 to 1 is preferable, a partial binding isotherm can be used to determine K_d if the quantum yield ratio φ is accurately predetermined (**Eq (3.27)**). However, in most K_d-determination approaches, free ligand and complex cannot be completely separated, and R is calculated with **Eq (3.13)**, in which the signals from pure ligand and pure complex (*i.e.*, S_L^* and S_C^*) play crucial roles. As we mentioned above, the measurement of S_L^* is trivial, while the determination of S_C^* requires that the binding process reach saturation, *i.e.*, $C >> L$. Since all the R values are calculated based on S_L^* and S_C^* (**Eq (3.13)**), an inaccurate S_C^*, resulting from a binding isotherm not reaching saturation, can distort the whole binding isotherm and significantly reduce the accuracy of $K_{d,det}$.

Here, we simulated binding isotherms (with Data S1, https://doi.org/10.6084/m9.figshare.25464685) by adding 10% relative systematic error in S_C^* (*i.e.*, binding isotherm does not reach saturation and $\Delta S_C^*/S_C^* = 0.1$) for a large range of L_0/K_d (**Figure 3.11a**), and summarized the dependence of the $K_{d,det}/K_d$ ratio on L_0/K_d (**Figure 3.11b**). **Figure 3.11b** indicates that, for binding isotherms not reaching saturation conditions, $K_{d,det}$ can

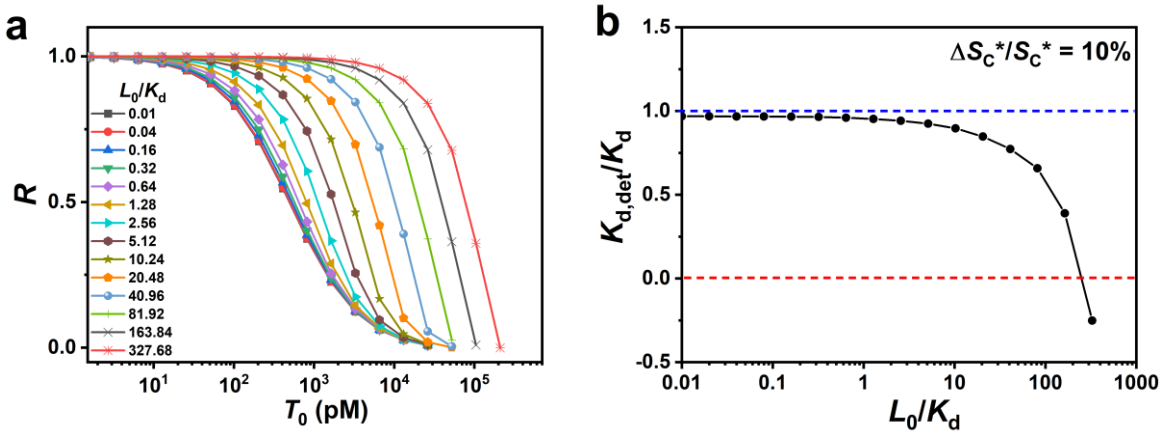


Figure 3.11. The dependencies of representative binding isotherms **(a)** and $K_{d,det}/K_d$ **(b)** on L_0/K_d with adding 10% relative systematic error in S_C^*, *i.e.*, the measured S_C^* is 10% higher than true S_C^* determined at saturation. In these simulations, we assume the cumulative signal S decreases with the increase of T_0.

be much lower than the true K_d (even reaching impossible negative values) at high L_0/K_d, emphasizing the importance of using T_0 that allows the binding isotherm to reach saturation condition.

To ensure that the binding isotherm reaches saturation, the conditions of the equilibrium mixture with the highest T_0 should satisfy **Eq (3.14)**, *i.e.*, $T_0 > L_0$ and $T_0 >> K_d$. The condition of $T_0 > L_0$ is easy to satisfy if L_0 is chosen to be close to the LOQ of the instrument (usually in pico- to nano-molar range). However, it is difficult to conclude if the condition $T_0 >> K_d$ is satisfied without knowing the true K_d.

A potential solution to this problem is to use the largest possible target concentration T_0 (which represents the solubility of the target in the sample buffer), in the equilibrium mixture (as discussed in *Section 3.2.3.2.2* for determining φ in complete-separation approaches). However, this solution has two main limitations for incomplete-separation approaches. First, using excessively high T_0 itself may affect the accuracy of the measured S_C^*. For instance, the

increased viscosity of the sample caused by high T_0 (*e.g.*, protein concentration) can affect the detected signals for diffusivity-based (or -related) approaches such as ACTIS and MST. Second, if the binding isotherm reaches saturation with T_0 much smaller than solubility, reaching the highest possible T_0 would waste a large amount of the target, which is often a precious protein.

Therefore, to ensure that a binding isotherm reaches saturation and to avoid the drawbacks associated with using excessively high T_0, we suggest a criterion for saturation in a binding isotherm. The criterion entails ensuring that zero lies within the uncertainty range of the slope for the linear fitting of the "cumulative signal S versus T_0" obtained from the equilibrium mixtures with the three highest T_0. Additionally, the highest T_0 should be at least twofold greater than the lowest T_0 in the three equilibrium mixtures.

3.2.3.2.4. Estimating Systematic Error in R Quantitatively

To provide researchers a comprehensive guidance to minimize the systematic error in R, we summarize the measures discussed above into a checklist, which is an expanded version of the checklist proposed by Jarmoskaite *et al*. (see **Table B2.1**).[34] By using the checklist in **Table B2.1**, researchers can reduce the likelihood of common theoretical or experimental mistakes in determining R and thus significantly reduce ΔR. Based on the experimentalist's confidence in the determined calibration parameters, the accuracy of the instruments used to measure signals, and other factors, the range of minimized $\Delta R/R$ (relative systematic error of R) with a certain confidence level might be estimated. However, at this stage, we have not identified an approach for quantitatively determining the interval of minimized systematic error of R that can be used as a reliable input to calculate the systematic error range of $K_{d,det}$. Note that, the random error of R (δR) is independent of ΔR and is translated into the random error of $K_{d,det}$ (σ) through nonlinear regression of a binding isotherm.

3.2.3.3. Sources and Confidence Intervals of Systematic Errors in Concentrations

Sources of systematic errors in T_0 and L_0 arise from various factors. If concentrations are calculated based on the weight of solid material and volume of solvent, imperfectly calibrated mass- and volume-measuring equipment as well as errors in product purity, can lead to systematic errors in concentrations. If concentrations are calculated using spectrophotometry (Lambert-Beer law), systematic errors in molar extinction coefficients will result in systematic errors in concentrations. These common sources can induce systematic errors in target and ligand concentrations of stock solutions, which, in turn, can propagate into systematic errors in concentrations of other diluted solutions used in K_d-determination experiments. Another common source of systematic errors in T_0 and L_0 occurring at any step of an experiment is solute adsorption onto pipette tips, vials, channels, *etc*. In **Appendix B2**, we illustrate the effect of systematic errors in T_0 and L_0 on the accuracy of $K_{d,det}$ for a large range of L_0/K_d (**Figure B2.5**) and delve into the common sources of systematic errors in concentrations along with strategies to mitigate them (Note B2.11). To assist researchers in minimizing systematic error in concentrations L_0 and T_0, the measures to minimize ΔL_0 and ΔT_0 are summarized in a checklist (**Table B2.1**) in **Appendix B**.

When the systematic errors of variables are minimized, it is important to have an approach that can quantitatively assess the minimized systematic errors since estimating the systematic error in $K_{d,det}$ — using the logic described in **Figure 3.4b** — requires input values for ΔL_0, ΔT_0, and ΔR. If exact values of systematic errors in variables can be determined, then the case is trivial as the variable can be simply corrected for this systematic error. Conversely, if nothing is known about systematic errors in variables, then the problem is ill-posed as completely unknown systematic errors cannot be propagated. However, there is a third case where the exact values of

systematic errors in variables are unknown, but the confidence intervals of the systematic errors with desired confidence levels can be quantitatively assessed. Confidence intervals of systematic errors are not commonly discussed in statistics, but they likely play a crucial role in the determination of physicochemical parameters with nonlinear regression. Here, we discuss two cases of assessing confidence intervals of systematic errors in concentrations.

The first case involves a rough estimation based on experiences, akin to the estimation of $\Delta R/R$ range mentioned in *Section 3.2.3.2.4*. If techniques listed in **Table B2.1** (*e.g.*, SDS-PAGE) for determining large systematic errors or approaches of method optimization (*e.g.*, surface-modification) are inaccessible, an experienced experimentalist should be able to estimate the maximum ranges of relative systematic errors in concentrations with some confidence, *e.g.*, ± 20% for ligand concentration and ± 30% for target concentration with confidence.

The second case involves a quantitative assessment based on the relative random errors (*i.e.*, relative standard deviations, RSD) in concentrations of stock solutions. If all necessary procedures for minimizing systematic errors (*e.g.*, device calibration, purity measurement, and method optimization) have been properly conducted, we can reasonably assume that the systematic errors in L_0 and T_0 of equilibrium mixtures are from the random errors in the concentrations of ligand and target stock solutions, *i.e.*, the normally distributed concentrations of stock solutions in infinite preparations (**Figure 3.12**). This assumption is valid because all the diluted solutions used in K_d determination with nonlinear regression are prepared from a single condensed ligand stock solution and a single condensed target stock solution. The random errors in the concentrations of stock solutions (present in an infinite number of sample preparations) are usually much greater than those of equilibrium mixtures due to the more complicated preparation procedure with more error sources. Thus, the confidence intervals of systematic

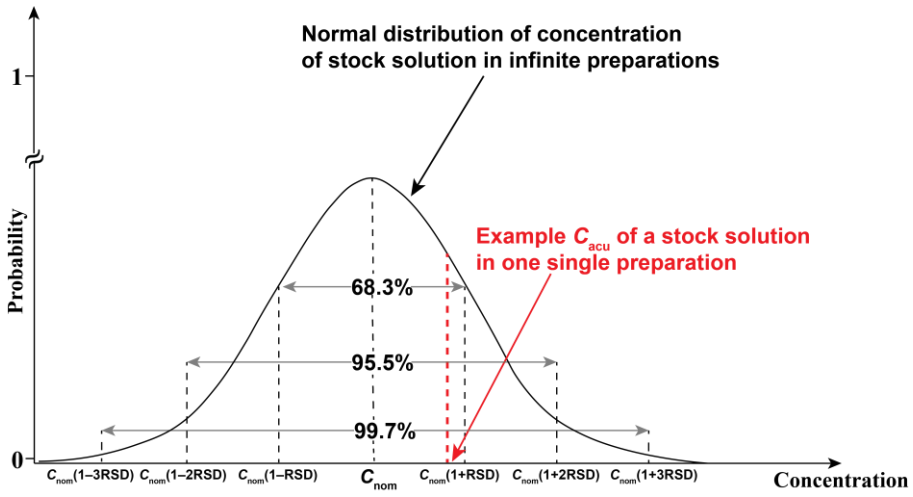


Figure 3.12. Demonstration of the relationship between the accurate concentration (C_{acu}) of a stock solution in one single preparation and the concentration distribution in an infinite number of preparations. With minimized systematic error in concentration, we assume the average concentration is equal to the nominal concentration C_{nom}.

errors in L_0 and T_0 (in equilibrium mixtures) can be estimated by measuring the RSD of the concentrations of their stock solutions in multiple preparations. Since there is no reliable measure to determine concentration directly, RSD of concentrations should be determined by experimentally measuring the spectroscopic signals (*e.g.*, light absorption or fluorescence intensity) from stock solutions in multiple preparations (see an example in **Table B2.2**).

For a single preparation of stock solution with a nominal (desired) concentration C_{nom}, the confidence interval (*e.g.*, with 68.3% confidence level) of systematic error can be calculated as $\Delta C \in [-C_{nom} \times \text{RSD}, C_{nom} \times \text{RSD}]$, for which the confidence interval of accurate concentration C_{acu} is $C_{acu} \in [C_{nom}(1 - \text{RSD}), C_{nom}(1 + \text{RSD})]$ (**Figure 3.12**). Note that, the confidence level of systematic error or accurate concentration can be varied by using different numbers of RSD in the calculation of confidence intervals. As a result, due to equilibrium mixtures being prepared

by diluting stock solutions, the confidence intervals (*e.g.*, with 68.3% confidence level) of systematic errors in nominal L_0 and T_0 can be assessed as $\Delta L_0 \in [-\text{RSD} \times L_0, \text{RSD} \times L_0]$ and $\Delta T_0 \in [-\text{RSD} \times T_0, \text{RSD} \times T_0]$, respectively. It is important to emphasize that the confidence intervals of ΔL_0 and ΔT_0 derived from the RSD of concentrations of stock solutions are the lower limits of systematic error ranges assuming that all other systematic errors in the concentrations of stock solutions have been eliminated. Despite lacking an approach for quantitatively assessing the confidence interval of ΔR, the lower limits of confidence intervals of ΔL_0 and ΔT_0 explained here can be potentially used to assess the lower limit of confidence interval of ΔK_d with error propagation.[10]

3.2.4. Concluding Remarks

Since the systematic error in a determined K_d (ΔK_d) is translated from the systematic errors in the variables used to determine K_d, *i.e.*, ΔL_0, ΔT_0, and ΔR (**Figure 3.4b**), understanding the sources of systematic errors in all variables and minimizing these errors are pivotal to minimizing ΔK_d. It is crucial to emphasize that systematic errors in the variables can result in an incorrect ranking of ligands during high-throughput screening. This challenges the common belief that such errors systematically shift all $K_{d,det}$ values in the same direction without affecting the ranking order. For instance, ligands with low true K_d values can have high L_0/K_d values; without minimizing the systematic errors in variables, $K_{d,det}$ can be much greater than K_d, wrongly categorizing these ligands as low-potency.

To determine an accurate K_d, the first step (often ignored by researchers) is to ensure there is no theoretical or experimental mistake in the K_d-determination approach. Such mistakes can introduce systematic errors in R, and eventually translate into systematic errors in $K_{d,det}$, which can be extremely large at unfavorable (large) values of L_0/K_d (see, for example, **Figure 3.6**). In

this work, we discussed the most common sources of systematic errors in R, such as non-additive signals, mis-calibrated instruments, inaccurate calibration parameters, *etc*., and quantitatively studied the effect of errors from these sources on the accuracy of $K_{d,det}$. We also provided suggestions on how to reduce or eliminate the systematic error in R caused by the common error sources. With careful theoretical and experimental design and proper instrument calibration, ΔR can be minimized, while the confidence interval of the minimized ΔR cannot be quantitatively determined.

The sources of systematic errors in concentrations (*i.e*., ΔL_0, ΔT_0), such as mis-calibrated measuring equipment, product impurities, and reagent adsorption onto surfaces, have been well-studied, and many approaches to eliminate these error sources have been developed (Note B2.11).[135, 289, 290, 291] In this work, we summarized the measures of eliminating or minimizing ΔR, ΔL_0, and ΔT_0 caused by the common error sources into a checklist (**Table B2.1**), which is a convenient tool for researchers in the field of K_d determination. For any specific K_d-determination approach, researchers should establish a standard operating procedure (SOP) for minimizing systematic errors in concentrations based on the chemical properties of the ligand and target used. The SOP should be established prior to any experiment, and it should be strictly followed during the K_d-determination experiments.

Although the systematic errors in concentrations cannot be fully eliminated, and their true values can hardly be determined accurately, the confidence intervals of the minimized systematic errors in concentrations can be quantitatively assessed with the quantitative study of random errors in concentrations of stock solutions. We foresee that, by combining the quantitatively assessed confidence intervals of systematic errors in concentrations and numerical

error propagation for nonlinear regression (will be discussed in **Chapter 4**), the lower limit of confidence interval of systematic error in $K_{d,det}$ can be determined.[10]

In summary, this investigation represents a pioneering effort to theoretically analyze the common sources of unknown systematic errors in K_d determination. Our findings offer valuable insights and practical strategies that are crucial for researchers aiming to maximize the accuracy of K_d measurements in binding studies. By thoroughly understanding and addressing these error sources, this work provides a robust framework that can be integrated into both academic curricula and professional practice, thereby enhancing the reliability of experimental outcomes. Additionally, our study serves as a critical reference for peer reviewers and journal editors in assessing the methodological soundness of research articles reporting K_d values. Ultimately, the adoption of these guidelines will foster higher standards of reliability in the field, benefiting the wider scientific community and advancing the quality of research in molecular interactions.

Chapter 4. The Accuracy Confidence Interval (ACI): Concept and Applications in Assessing Physicochemical Parameter Accuracy

4.1. ACI: A Practical Approach to Quantitatively Assessing Equilibrium-Constant Accuracy from a Single Binding Isotherm

The following material was previously published and is reprinted with permission from "**Wang, T. Y.**; Latimer, J.; Rukundo, J.-L.; Kogan, I.; Krylova, S. M.; Schreiber, S.; Kohlmann, P.; Jose, J.; Krylov, S. N. A Practical Approach to Quantitatively Assessing Equilibrium-Constant Accuracy from a Single Binding Isotherm. *Precision Chemistry* **2025**, *3* (2), 89–104." Copyright 2025 American Chemical Society. My contributions to the article were: (*i*) conducting the mathematical analysis and equation derivations, (*ii*) performing all virtual experiments, (*iii*) conducting the ACTIS experiments, (*iv*) preparing all figures, (*v*) interpreting the results, and (*vi*) writing the manuscript. Co-author Jessica Latimer developed the Python code for the ACI-Kd program. Co-author Isaac Kogan converted the Python code to a web application. Co-author Philip Kohlmann conducted the FRET experiments.

4.1.1. Introduction: The Critical Need for Assessing K_d Accuracy

Adequate quantitative description of natural phenomena requires the knowledge of accurate (*i.e.*, free of unknown systematic errors) values of physicochemical parameters. Finding accurate values is only straightforward for parameters that can be directly measured against reference standards of length, mass, time, temperature, electric charge, *etc*.[26, 27, 28, 29] The majority of physicochemical parameters, however, do not have reference standards and, therefore, must be computed using formulas linking them to sets of dependent and independent variables.[3, 15, 292, 293, 294] Nonlinear regression of experimental data is required for computing some of such parameters,

e.g., various equilibrium and pseudo-equilibrium constants (dissociation constant of biomolecular complexes, stability constant of metal–ligand complexes, Michaelis constant of enzymatic reactions, and Langmuir constant of adsorbent–adsorbate interactions).[3, 4, 5, 6, 7, 8] The regression models used to compute physicochemical parameters are usually correct, *i.e.*, derived from "first principles" rather than "guessed" to approximate empirical data.

Nonlinear regression accounts only for random errors of variables and, accordingly, assesses only the random error of the computed parameter. This random error describes the precision of the parameter determined with varied independent variables.[239, 240] This error defines a range which we call a "Precision Confidence Interval" (PCI) of the sought parameter. It is often presented as a mean value of the parameter plus/minus one or more standard deviations.[295] Being an easily calculated descriptor of precision, PCI does not directly relate to the accurate value. In other words, the probability with which the accurate value lies within PCI is not only unknown

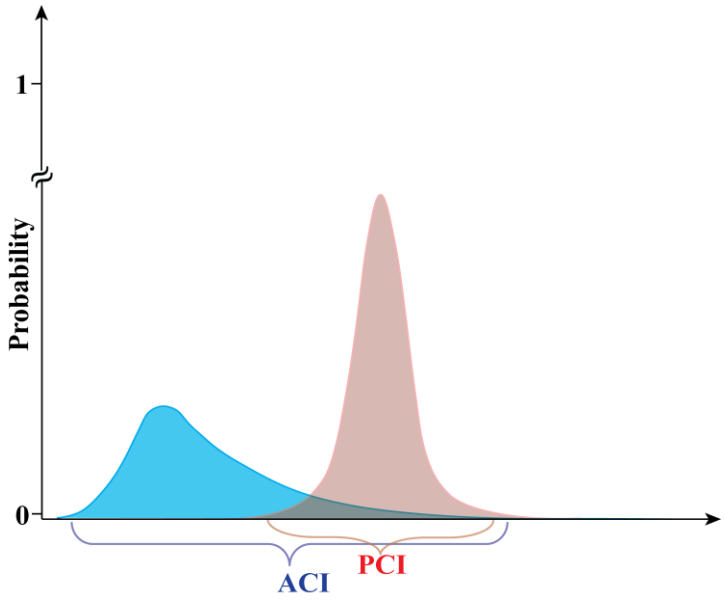


Figure 4.1. A schematic depiction of potentially large differences between the precision confidence interval (PCI) and accuracy confidence interval (ACI).

but also can be very small. Basing conclusions solely on PCI may lead to serious misconceptions.[32, 35]

When finding an unknown parameter with nonlinear regression, we should aim to determine not only its PCI but also a range in which the accurate value of the sought parameter lies with a desired probability; we call this range an "accuracy confidence interval" (ACI). Accuracy and precision are unlinked (orthogonal) descriptors of the exactness (or trueness) of the determined value. Accordingly, ACI and PCI may differ significantly in position, width, and shape of probability distribution as schematically depicted in **Figure 4.1.** This difference makes it necessary to determine not only the PCI but also the ACI, raising a question of how the ACI of a parameter computed with nonlinear regression can be determined.

With the advent of computers, nonlinear regression has become a widely used tool for determining unknown parameters using mathematical models that connect these parameters to known variables via explicit formulae. Statistical methods based on Monte Carlo simulation have been developed to propagate the confidence intervals (uncertainties) of variables to the confidence intervals (uncertainties) of determined parameters.[140, 141, 142] These methods were adopted for the determination of equilibrium constants, and both commercial statistical software and open-access codes are available for their implementation.[141, 296, 297, 298]

Despite these advancements, the quantitative assessment of the accuracy of equilibrium constants is grossly neglected by molecular scientists who routinely determine such constants using nonlinear regression. There are several reasons for this alarming oversight, with the major cause arguably being that the existing methods are not practical for researchers in this field. These methods may be difficult to understand for non-mathematicians, and they can require programming skills and/or intensive computational resources for implementation.[141, 296, 297, 298]

Additionally, there is a significant conceptual barrier: the statistical methods do not clearly explain what should be used as input parameters for the confidence intervals of the variables, which is a very complex issue on its own.[142, 182]

In this study, we aim to address these knowledge gaps by (*i*) developing a practical approach to finding the ACI using language and concepts comprehensible to molecular scientists, (*ii*) creating a computationally efficient tool that is immediately accessible to all users regardless of their level of preparation, and (*iii*) explaining to experimentalists what confidence intervals of variables (concentrations) should be used for the determination of ACI of the determined parameter (equilibrium constant).

Here, we introduce "ACI-Concept", a methodology for finding ACI of parameters computed with correct nonlinear-regression models. The need for a correct model is essential for our approach which combines the analysis of regression sensitivity to errors of independent variables that are fixed in regression (*e.g.*, concentration of a limiting component in the determination of equilibrium constants) with the standard propagation of errors of all (independent and dependent) variables or only independent variables (*e.g.*, concentrations in the determination of equilibrium constants). We also propose a strategy to translate the theory underlying ACI-Concept into practical workflows for the determination of ACI.

Our ACI-determination strategy is designed to directly identify the boundaries of the ACI at a specified confidence level, focusing on determining the width of the ACI, which is arguably the most critical factor in assessing physicochemical parameters. This approach offers significantly greater computational efficiency compared to traditional Monte Carlo simulation-based methods, which require extensive resampling.[140, 141, 142]

We chose the equilibrium dissociation constant (K_d) of affinity complexes as the first parameter for which a practical ACI-determination workflow was to be developed. K_d is determined through nonlinear regression of binding isotherms and is arguably the most frequently computed physicochemical parameter, reported in over 1,000 peer-reviewed papers per month. Many more K_d values are likely produced in high-throughput screening by drug-discovery companies solely for internal use. In the absence of a practical procedure for molecular scientists to find ACI of K_d, conclusions from binding experiments are based solely on PCI, which may drastically differ from ACI. The inability to find ACI of equilibrium constants can lead to serious misinterpretations and misconceptions.[15, 32, 35, 299]

We used ACI-Concept to develop a practical workflow for the determination of ACI of K_d; we term this workflow "ACI-Kd". ACI-Kd was verified with computer-simulated binding isotherms, resulting in a wide range of accuracy of determined values of K_d. The ACI-Kd workflow was initially exemplified by determining the ACI of K_d values in virtual binding experiments. It was then applied to real experiments involving two biomolecular complexes, where binding isotherms were obtained using methods based on different physical principles. To assist the scientific community with adopting ACI-Kd, we implement it as a user-friendly webapp (https://aci.sci.yorku.ca) which computes both ACI and PCI and, thus, facilitates the assessment of both accuracy and precision of the determined K_d values. No math or programming skills are required from the users and the results, including graphs suitable for publication, are produced almost instantly. The consideration of both ACI and PCI while interpreting binding experiments will greatly improve the reliability of drawn conclusions.

It is important to emphasize that ACI-Kd is not a "panacea for all ills" as it considers only one potential source of inaccuracy: the transformation of small systematic errors of

concentrations into large systematic errors of K_d. This source is arguably the most important as it is fundamental, but there are unique sources of inaccuracy for each experimental method and other imperfections that can bias ACI of K_d making it less reliable.[169, 300, 301] We emphasize the need for thoughtful experimentation, which should include (*i*) understanding and minimizing all method-dependent biases, (*ii*) excluding all known systematic errors, (*iii*) optimizing experimental conditions, (*iv*) ensuring equilibrium and saturation in binding reactions, *etc*.[182] At the end of the day, ACI of K_d is only as good as the data used to compute it.

We anticipate that ACI-Concept will be used to develop workflows for finding ACI of other physicochemical parameters computed through nonlinear regression. A key condition for the applicability of ACI-Concept is that the regression model is correct, which is typically not the case in areas other than natural sciences, and that the error is analytically propagated, which is always possible to do.[302, 303, 304]

4.1.2. Results and discussion

4.1.2.1. Accuracy and ACI

We start with a theoretical consideration of the accuracy of parameters computed with nonlinear regression with a correct model. Accuracy indicates the closeness of the determined value to the true value of the parameter. A true value is never known but can be used here as a theoretical abstraction. The input values used in numerical simulations can be considered as true. When discussing accuracy in the experimental context, we will use “accurate” instead of “true” as a descriptor of value exactness. The accuracy can be expressed quantitatively in different ways, but its measures are always defined via the determined and true values. For example, a relative systematic error is a difference between the determined and true values divided by one of them, whereas an accuracy ratio is a ratio between them.[305, 306] Because a true value is

fundamentally unknown, neither measure of accuracy can be calculated using its definition. Moreover, since a determined value is always uncertain due to experimental systematic errors, accuracy can only be assessed with the corresponding uncertainty. This uncertainty defines ACI of the determined value.

4.1.2.2. ACI-Concept: Combining Regression-Stability and Error-Propagation Analyses

In the determination of equilibrium constants with nonlinear regression, researchers usually assume that parameter p is a known function of a measurable independent variable x_c that is kept as a constant in a single nonlinear regression, a measurable independent variable x_v that is varied in a large range, and a measurable dependent variable, signal S:[3, 4, 5, 6, 7, 8]

$$p = F(x_c, x_v, S) \tag{4.1}$$

The case can be generalized, for example, to multiple independent variables that are kept constant (Note C1.1 in **Appendix C**). The requirement that this function is known is equivalent to the requirement that the regression model is correct. In turn, the correct regression model is needed as it is a key reference in accuracy assessment with the accurate value being unknown. Assessing the accuracy of a determined value of p (p_{det}) requires the knowledge of systematic errors of all the variables, which are Δx_c, Δx_v, and ΔS, and we assume that they are also measurable (known). Here, we consider only the determination of p via a commonly used regression approach which assumes Δx_c, Δx_v, and ΔS are all null and allows only random errors in S. Our goal is to integrate the non-zero systematic errors of variables in this approach. In the simplest implementation of this approach, a theoretical dependence of S on a single variable x_v is fitted into a corresponding experimental dependence with x_c kept constant:

$$S = f(x_c, x_v, p) \tag{4.2}$$

using p as a fitting parameter. A regression procedure can be done with various computer algorithms, and it typically reports p_{det} along with its standard deviation (*i.e.*, random error), an indicator of the precision of p_{det} determined with varied x_v.[240] The random error does not show how close p_{det} is to the true value of p (note, we use p without a subscript for both naming parameter p and indicating the true value of p). Here, we are developing an approach to assessing the accuracy of p_{det} assuming that the systematic errors of the variables are analytically propagated into the systematic error of the sought parameter.

The initiating point in this development is our insight that this approach can be built upon a link between the accuracy of p_{det} and the sensitivity of p_{det} to a change in x_c across different nonlinear-regression processes. As p is a constant, we can discuss instead a link between the accuracy of p_{det} and the sensitivity of p_{det} to a change in the x_c/p ratio. This link can be proven mathematically, and it implies that the accuracy of p_{det} can be improved by choosing x_c at which p_{det} is less sensitive to the change of x_c/p across different regressions.[101, 179] The underlying reason for this relationship between p_{det} accuracy and its sensitivity to changes in x_c/p is that when p_{det} is less sensitive to variations in x_c/p, the error in p_{det} — propagated from the systematic errors in x_c, x_v, and S (represented by the vector $\mathbf{\Delta} = \{\Delta x_c, \Delta x_v, \Delta S\}$) — is minimized.[179]

Experimentalists are aware of this relation between the accuracy and sensitivity and sometimes vary x_c across different p-determination experiments to identify the x_c (with an unknown value of x_c/p) where p_{det} is not sensitive to a change in x_c.[34, 307, 308] However, experimentally varying x_c greatly increases the amount of required experimentation, which restricts the use of this approach. More importantly, this approach can only suggest if the error of p_{det} has been minimized for a fixed $\mathbf{\Delta}$ but cannot produce a quantitative assessment of the accuracy of p_{det}. These two limitations led us to consider the known qualitative sensitivity of p_{det}

to x_c (or x_c/p) for a fixed $\mathbf{\Delta}$ from a different angle: the x_c/p ratio can be assessed by measuring the sensitivity of p_{det} to the changes in the vector of systematic errors $\mathbf{\Delta} = \{\Delta x_c, \Delta x_v, \Delta S\}$. To simplify the analysis, we reduce the change of the vector $\mathbf{\Delta} = \{\Delta x_c, \Delta x_v, \Delta S\}$ to a change in a single element, Δx_c.

The sensitivity of the determined value p_{det} to changes in Δx_c can be quantified by the ratio p_{det}'/p_{det}, where p_{det} and p_{det}' represent the values of parameter p determined at a specific value of x_c/p for two different vectors of systematic errors in variables, such as $\mathbf{\Delta} = \{\Delta x_c, \Delta x_v, \Delta S\}$ and $\mathbf{\Delta'} = \{\Delta x_c', \Delta x_v, \Delta S\}$, respectively. We refer to this sensitivity of the determined parameter to systematic errors of variables in nonlinear regression as "regression stability". We postulate that if the theoretical relation between p_{det}'/p_{det} and x_c/p can be solved analytically or numerically when expressed explicitly, the accurate value of the studied parameter can be computed by evaluating the regression stability and assessing the value of x_c. The following is the logic developed upon this postulate.

Theoretically, when the range of x_v (and corresponding dependent variable S) adequately covers the critical region for determining p, and data is densely sampled, we can assume the accuracy of p is weakly dependent on the specific choices of x_v.[309] Under this readily achievable condition, the systematic error of p_{det} (Δp) will primarily depend on the chosen value of the variable kept constant (x_c) and the errors of all variables (**Eq (4.3)**):

$$\Delta p = G(x_c, \mathbf{\Delta}) \tag{4.3}$$

where $\mathbf{\Delta} = \{\Delta x_c, \Delta x_v, \Delta S\}$. Also, since a systematic error is *per se* the deviation of the determined value from the accurate value, **Eq (4.3)** can be rewritten to:

$$\Delta p = p_{det} - p = G(x_c, \mathbf{\Delta}) \Rightarrow p_{det} = G(x_c, \mathbf{\Delta}) + p \Rightarrow p_{det} = g(x_c, \mathbf{\Delta}) \tag{4.4}$$

A transformation of **Eq (4.4)** into an expression linking a unitless accuracy ratio p_{det}/p to unitless ratio x_c/p and unitless relative errors of variables has been previously demonstrated for a specific type of **Eq (4.1)**.[179] A similar transformation strategy can be applied to other types of **Eq (4.1)**. The transformed dimensionless expression can be presented as:

$$\frac{p_{det}}{p} = H\left(\frac{x_c}{p}, \mathbf{\Delta}^*\right), \text{ where } \mathbf{\Delta}^* = \left\{\frac{\Delta x_c}{x_c}, \frac{\Delta x_v}{x_v}, \frac{\Delta S}{S}\right\} \tag{4.5}$$

To study the sensitivity of p_{det}/p to the unitless relative error of the variable kept constant ($\Delta x_c/x_c$), we need to have at least two points, corresponding to errors $\mathbf{\Delta}^* = \{\Delta x_c/x_c, \Delta x_v/x_v, \Delta S/S\}$ and $\mathbf{\Delta}^{*\prime} = \{\Delta x_c/x_c + \delta, \Delta x_v/x_v, \Delta S/S\}$, where δ should be large to ensure that the values of p_{det} obtained with these two errors are distinguishable. **Eq (4.5)** can be written for $\mathbf{\Delta}^{*\prime}$ as:

$$\frac{p_{det}'}{p} = H\left(\frac{x_c}{p}, \mathbf{\Delta}^{*\prime}\right) \tag{4.6}$$

To investigate the sensitivity of the determined parameter to the systematic errors of variables at different x_c/p, we divide **Eq (4.6)** over **Eq (4.5)** to obtain:

$$\frac{p_{det}'}{p_{det}} = H\left(\frac{x_c}{p}, \mathbf{\Delta}^{*\prime}\right) \Big/ H\left(\frac{x_c}{p}, \mathbf{\Delta}^*\right) = h\left(\frac{x_c}{p}, \mathbf{\Delta}^*, \mathbf{\Delta}^{*\prime}\right) \tag{4.7}$$

The theoretical relation shown in **Eq (4.7)** can be numerically tabulated by fitting the simulated datasets (involving relative systematic errors of $\Delta x_v/x_v$ and $\Delta S/S$) with **Eq (4.2)**, using $(1 + \Delta x_c/x_c)x_c/p$ and $(1 + \Delta x_c/x_c + \delta)x_c/p$ as the "nominal" values of the variable kept constant. The relationship between these two nominal values is given by $(1 + \Delta x_c/x_c + \delta)x_c/p = \alpha[(1 + \Delta x_c/x_c)x_c/p]$, where $\alpha = 1+\delta/(1 + \Delta x_c/x_c)$. To determine the x_c/p ratio for fixed $\mathbf{\Delta}^*$ and $\mathbf{\Delta}^{*\prime}$, one can consider the inverse relationship between p_{det}'/p_{det} and x_c/p, as follows:

$$\frac{x_c}{p} = h^{-1}\left(\frac{p_{det}'}{p_{det}}\right), \text{ at } \boldsymbol{\Delta}^* \text{ and } \boldsymbol{\Delta}^{*\prime} \tag{4.8}$$

In practice, p_{det} and p_{det}' can be determined by using the nominal value of x_c (involving relative systematic error of $\Delta x_c/x_c$) and αx_c, respectively, in a regression procedure with a correct regression model (fitting **Eq (4.2)** into experimental data that includes experimental $\Delta x_v/x_v$ and $\Delta S/S$). Note that, the value of α is controllable in the data fitting process, and the confidence interval of $\boldsymbol{\Delta}^*$, *i.e.*, the combination of $\Delta x_c/x_c$, $\Delta x_v/x_v$, and $\Delta S/S$, with a desired confidence level, can be potentially determined.

We presume that the ratio p_{det}'/p_{det} for a given value of α can be experimentally determined. By substituting this ratio into the numerically established **Eq (4.8)**, which involves the confidence interval of $\boldsymbol{\Delta}^*$ and utilizes the same value of α, the confidence interval of x_c/p can be determined. Finally, since the nominal x_c is known, we can, in principle, determine the confidence interval of p, *i.e.*, ACI of p. To determine ACI, it is essential to first establish the confidence interval of relative systematic errors of the variables (*i.e.*, the components of vector $\boldsymbol{\Delta}^*$). This aspect is studied in the next section.

4.1.2.3. Systematic Errors of Variables Arising from Random Errors

The above theoretical consideration assumes that the confidence interval of systematic errors of variable $\boldsymbol{\Delta}^*$ is known. Let's define what kind of systematic errors ACI-Concept is applicable to. As we explained in our previous work,[182] only if the systematic errors arise from normally-distributed random errors of variables, the confidence intervals of systematic errors can be assessed.[182]

Although normally-distributed systematic errors are not commonly discussed in statistics, they likely play an important role in the determination of physicochemical parameters with

nonlinear regression.[182] We would like to use an example of concentrations as variables. If we assume that the concentration is determined without a true systematic error, it always has a normally distributed random error. Such an error is necessarily present when preparing a stock solution and if the same stock solution is used to prepare all other solutions in the experiment, then the random error translates into a normally-distributed systematic error.[182] Such normally distributed systematic errors establish the confidence interval for the vector of systematic errors in variables, **Δ***, which, in turn, defines the ACI within the ACI-Concept framework through error propagation (**Figure 4.2**). Thus, ACI-Concept is meant to be used for normally-distributed

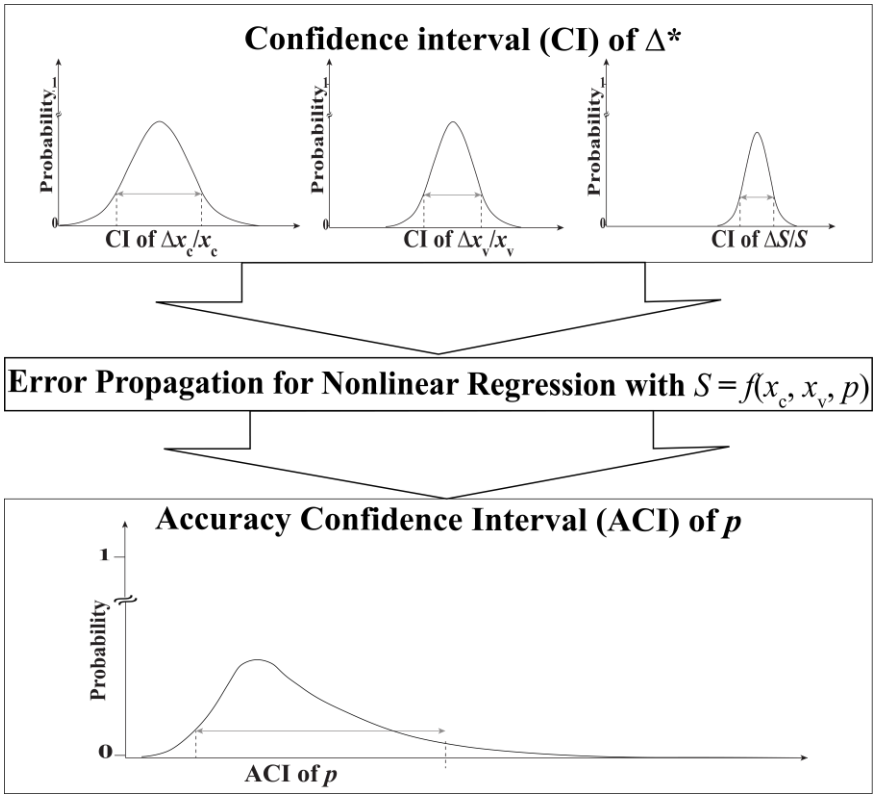


Figure 4.2. A schematic depiction of how the ACI of parameter p is defined by the confidence interval (CI) of **Δ***, which is the combination of CIs for the relative systematic errors of the variables.

systematic errors of variables. It is important to emphasize that normally-distributed systematic errors represent a lower bound of systematic errors as we assume that all identifiable and correctable sources of systematic errors have already been addressed and eliminated.[182] Therefore, ACI is a narrower limit of the range in which the accurate value of the sought parameter lies.

4.1.2.4. Graphical Strategy for Development of ACI-Determination Workflows

To be instructive, the above general theory needs to be converted into a specific strategy. There may be more than one strategy for assessing the accuracy of computed values based on the described combination of regression-stability analysis and error-propagation analysis. Here we propose a strategy that we nicknamed "graphical". It aims at finding ACI, *i.e.*, a range in which a true value of p lies with a desired probability.

The essence of the graphical strategy is the following. The required experimental data include (*i*) the dependence of S on x_v for a given x_c and (*ii*) the confidence intervals of (experimentally determinable) relative systematic errors of all variables, which are used to calculate the confidence interval of $\mathbf{\Delta^*}$ (*i.e.*, combinations of $\Delta x_c/x_c$, $\Delta x_v/x_v$, and $\Delta S/S$). To find the confidence interval of $\mathbf{\Delta^*}$, we can first analytically solve **Eq (4.1)** for a specific case using the rules of error propagation to obtain a specific form of **Eq (4.5)**. Then, the experimentally determined confidence intervals of relative systematic errors of variables are plugged into **Eq (4.5)** to find the confidence interval of $\mathbf{\Delta^*}$ that is bordered with two boundary conditions: $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$, which cause the smallest and largest accuracy ratio p_{det}/p, respectively. **Eq (4.7)** is used to compute p_{det}'/p_{det} as functions of x_c/p for these two values of $\mathbf{\Delta^*}$ (*i.e.*, $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$). The two lines along with the grey shadow between them in a graph schematically depicted in **Figure 4.3**, which we call an ACI graph, border the ranges within which x_c/p is positioned with a probability

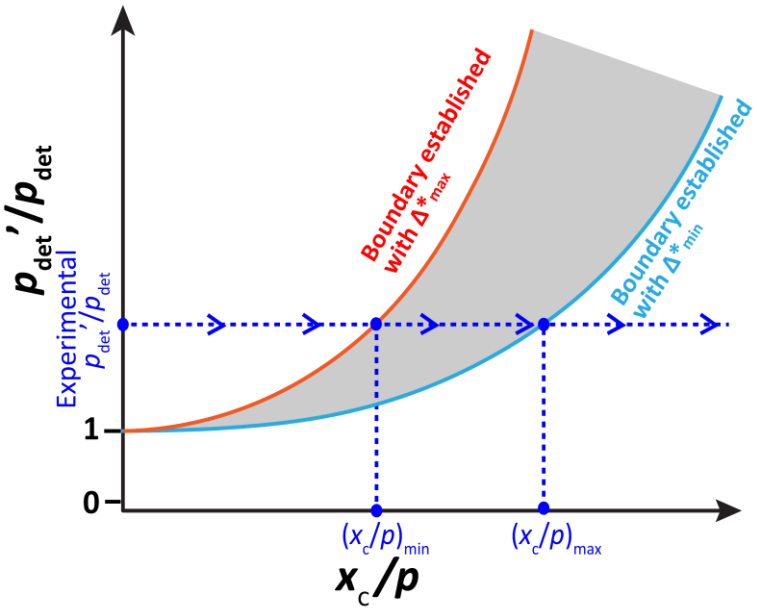


Figure 4.3. A schematic of an ACI graph.

defined by those at which **Δ*** was obtained. The numerically derived ACI graph (**Figure 4.3**) is equivalent to the explicit expression presented in **Eq (4.7)**. A known value of p_{det}'/p_{det} is placed on the graph as a horizontal line, and the intersections of this line/surface with the two border lines define the boundary values of $(x_c/p)_{min}$ and $(x_c/p)_{max}$. This step corresponds to solving **Eq (4.8)** for x_c/p. Finally, the nominal value of x_c is used to calculate the confidence interval in which p lies: between $p_{min} = x_c/(x_c/p)_{max}$ and $p_{max} = x_c/(x_c/p)_{min}$. The presented graphical strategy can presumably be applied to any parameter computed with the regression approach, but proving this is not a subject of our current work. Note that our graphical strategy is designed to directly identify the boundaries of the ACI at a specified confidence level without relying on random resampling, a core element of traditional Monte Carlo simulation-based methods.[140-142] Below, we use the graphical strategy to develop a workflow for finding ACI of K_d.

4.1.2.5. Focus on K_d

Drugs and diagnostic probes (collectively called ligands in this work) are developed to form stable affinity complexes with their molecular targets (usually proteins).[210, 310-312] The stability of a target–ligand complex (denoted as TL or C) is characterized by the equilibrium dissociation constant (K_d) of the binding reaction between a target (T) and a ligand (L) (**Eq (1.1)**). K_d can be calculated through **Eq (1.2)**, where we used T, L, and C to represent equilibrium concentrations of the target, ligand, and complex, respectively, in the reversible one-to-one binding process described by **Eq (1.1)**. Lower K_d values correspond to the higher thermodynamic stability of the complex and are usually desired in the development of drugs and diagnostic probes.

The values of K_d define therapeutic concentrations of drugs and analytical parameters of diagnostic methods; they must be known accurately for important therapeutic and diagnostic applications of ligands.[220, 221, 222] K_d values are a key decision-influencing parameter in early-stage drug discovery where the most potent candidates are chosen for further development. However, the discrepancies between K_d values obtained even within the same laboratory for the same complex and with the same method often reach many folds or even orders of magnitude.[35, 179, 215] Such discrepancies may lead to misconceptions and wrong conclusions about ligand potency with far-going negative consequences.[32, 35, 215]

There are no reference standards for K_d values; they are classically computed using the following procedure based on nonlinear regression (or a variation of this procedure). A series of equilibrium mixtures of the target and ligand are prepared in which the total concentration of ligand $L_0 = L + C$ is kept constant (limiting component) while the total concentration of target $T_0 = T + C$ is varied in a wide range (2 or 3 orders of magnitude). A fraction of unbound ligand R ($R = L/L_0$) is measured for each equilibrium mixture and plotted as a function of T_0; the dependence of R on T_0 is called a binding isotherm (**Figure 4.4a**).[69, 76, 171, 181] Finally, K_d is

determined by nonlinear regression of the binding isotherm with the theoretical dependence of R on T_0 of **Eq (2.2)** ($R = -(K_d + T_0 - L_0)/2L_0 + \sqrt{[(K_d + T_0 - L_0)/2L_0]^2 + K_d/L_0}$) using K_d as the only unknown parameter (red line in **Figure 4.4a**).[101, 171, 179, 224]

It is important to note that **Eq (2.2)** is obtained without simplifying assumptions from the definition of K_d (**Eq (1.2)**), the definition of R (**Eq (2.1)**: $R = L/L_0$) and mass balance for ligand and target (**Eq (3.10)**: $L_0 = L + C$ and $T_0 = T + C$).

The computation of K_d would be straightforward and error-free if R, L_0, and T_0 were known precisely and accurately. However, the values of R, L_0, and T_0 always have errors making

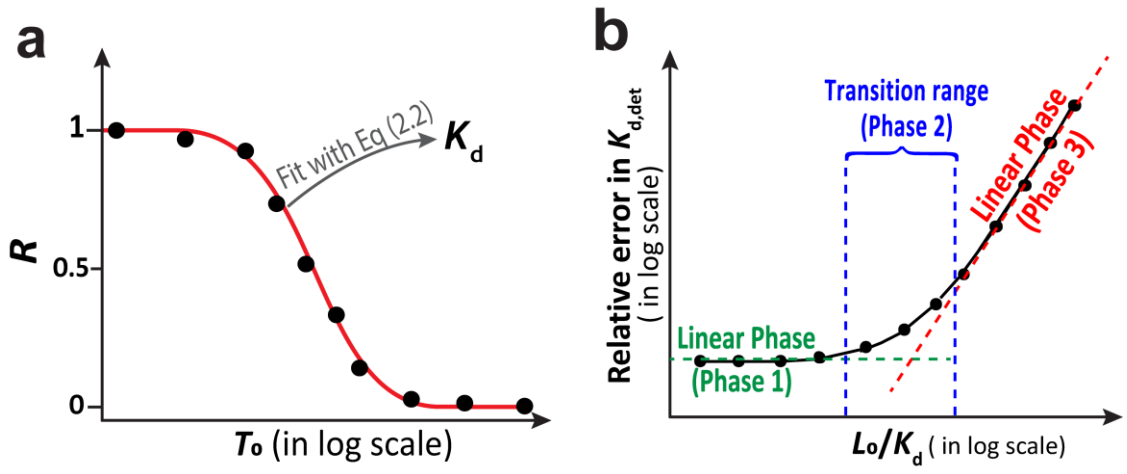


Figure 4.4. A schematic illustration of K_d determination. Panel **(a)** shows a binding isotherm (dots), *i.e.*, the experimental dependence of the fraction of unbound ligand (R) on the total target concentration (T_0), and the result of fitting the isotherm with **Eq (2.2)** (red line), which is the theoretical dependence of R on T_0. Panel **(b)** illustrates the triphasic dependence of relative error in the determined value of $K_{d,det}$ on L_0/K_d with constant errors in T_0, L_0, and R, where L_0 and K_d represent true values of the total ligand concentration and the equilibrium constant, respectively. In Phase 1, the relative error of $K_{d,det}$ is virtually insensitive to change in L_0/K_d, while in Phase 3 it is sensitive. (This figure is modified from **Figures 3.1** and **3.4b**).

experimental determination of K_d error-prone, especially for highly stable complexes that need to be studied at low concentrations of both ligand and target.[34, 101, 105, 179, 313] It is important to stress that nonlinear regression of a binding isotherm produces only PCI.

It has been demonstrated numerically and proven analytically that the translation of the systematic errors of R, L_0, and T_0 into the systematic error of the determined value of K_d ($K_{d,det}$) is triphasic when presented as a function of the ratio between true values of L_0 and K_d: L_0/K_d (**Figure 4.4b**).[101, 179] Note, we use "K_d" to name accurate/true K_d values which correspond to an ideal case of no errors in L_0 and T_0 (accurate K_d is fundamentally unknown; we can only aim to find a confidence interval in which the accurate value of K_d lies with a known probability). "$K_{d,det}$" is used to name determined K_d values which are subject to systematic errors. These errors are due to the nonlinear regression of binding isotherms under the assumption of no systematic errors in L_0 and T_0, while such errors are always present. This triphasic dependence suggests that with constant systematic errors of R, L_0, and T_0, the sensitivity of $K_{d,det}$ to the change in L_0 increases with increasing L_0/K_d (**Figure 4.4b**). We can also interpret **Figure 4.4b** in a different way: the sensitivity of $K_{d,det}$ to the systematic errors of R, L_0, and T_0 increases with the increase of L_0/K_d. This relationship can serve as a foundation for determining L_0/K_d, which subsequently aids in evaluating the accuracy of the determined K_d.

The triphasic dependence in **Figure 4.4b** suggests the use of the lowest possible L_0 for decreasing errors of $K_{d,det}$. However, L_0 cannot be lower than the limit of quantitation (LOQ) of an instrument utilized for measuring the values of R.[179, 314, 315] Thus, for very stable complexes (small K_d values), it is likely that the lowest L_0 that can be used for a given LOQ is much greater than K_d. In such cases, L_0/K_d is large (Phase 3 in **Figure 4.4b**) and unknown. Accordingly, small errors of R, L_0, and T_0 are greatly magnified when propagated into the error of $K_{d,det}$. This type of

error in K_d is systematic and fundamental, *i.e.*, it cannot be reduced by repeating experiments with the same solutions (or even with different solutions but the same mass-measuring and volumetric equipment) and does not depend on the instrumental platform used to measure R.

We are aware of only one way to confirm that the error of $K_{d,det}$ is not greatly magnified by using too high L_0. It requires experimentally obtaining multiple binding isotherms for different values of L_0 and, essentially, plotting $K_{d,det}$ as a function of L_0.[34, 307, 308] If $K_{d,det}$ is found to not depend significantly on L_0, *i.e.*, $K_{d,det}$ is insensitive to changing L_0, then it is assumed that the value of $K_{d,det}$ is accurate, *i.e.*, $K_{d,det} \approx K_d$. This approach has two major limitations. First, building multiple binding isotherms is much more resource-intensive than building one. This limitation is likely the main reason for this approach not being used in the vast majority of peer-reviewed works reporting K_d determination. The second limitation is that this approach cannot provide any quantitative information on the accuracy of $K_{d,det}$, *i.e.*, on ACI. These limitations motivated us to apply ACI-Concept to the development of ACI-Kd, a practical workflow for finding ACI of K_d.

4.1.2.6. ACI-Kd — a Workflow for Finding ACI of K_d

The relation between the terms used in the theory section and the specific case of K_d determination is explicit: p corresponds to K_d, S corresponds to R, x_c corresponds to L_0, and x_v corresponds to T_0. As a result, **Eq (4.1)** corresponds to **Eq (3.1)** ($K_d = \left[T_0 - L_0(1-R)\right] / (1/R - 1)$).[247] To determine ACI of K_d, we should apply error propagation rules to **Eq (3.1)** and thus obtain an equation similar to **Eq (4.5)** that suggests how the confidence interval of **Δ*** can be defined based on the confidence intervals of variables that have normally-distributed systematic errors.

To obtain an analytical solution of error propagation for **Eq (3.1)**, it's necessary to focus on a single data point by fixing R. In a representative case of $R = 0.5$, which corresponds to the least

error in K_d determination,[101] the general **Eq (4.5)** can be specified as (Note C1.2):

$$\frac{K_{d,det}}{K_d} = \left[\frac{0.5(\Delta T_0/T_0 - \Delta L_0/L_0 - \Delta R/R)}{1+\Delta R/R} + \frac{0.5\Delta R/R}{0.5+\Delta R/R}\right]\frac{L_0}{K_d} + \left[\Delta T_0/T_0 + \frac{\Delta R/R(1+\Delta R/R)}{0.5+\Delta R/R} + 1\right] \quad (4.9)$$

In the experimental determination of K_d with nonlinear regression, careful experimental design and method optimization typically result in normally-distributed systematic errors in the concentrations of stock solutions. These solutions are used to prepare other solutions, thus, making T_0 and L_0 the independent variables with normally-distributed systematic errors.[182] In contrast, at this stage, we have not identified a quantitative approach for assessing the confidence interval of ΔR.[182] Therefore, due to the lack of inputting information for ΔR, this study will only consider the systematic errors of T_0 and L_0. Consequently, the determined ACI of K_d will be the narrower limit of the confidence interval for K_d with the specified confidence levels (probabilities).

Without considering ΔR, **Eq (4.9)** can be simplified to **Eq (4.10)** (Note C1.2):

$$\frac{K_{d,det}}{K_d} = 0.5\left(\frac{\Delta T_0}{T_0} - \frac{\Delta L_0}{L_0}\right)\frac{L_0}{K_d} + \frac{\Delta T_0}{T_0} + 1 \quad (4.10)$$

With this simplification, the vector of relative systematic errors **Δ*** corresponds to $\{\Delta L_0/L_0, \Delta T_0/T_0\}$, *i.e.*, the combination of relative systematic errors of L_0 and T_0. In the derivations of **Eq (4.9)** and **Eq (4.10)**, we assumed the error sources of the concentrations are strongly correlated, which is a valid assumption for the vast majority of experimental cases.[179, 316] The case for weakly correlated errors is conceptually and quantitatively similar but not considered here.[179]

Based on the above specifications, the dependence of p_{det}'/p_{det} on x_c/p shown in **Eq (4.7)** becomes the dependence of $K_{d,det}'/K_{d,det}$ on L_0/K_d. Here, $K_{d,det}$ and $K_{d,det}'$ are the K_d values determined by fitting **Eq (2.2)** into a single experimental binding isotherm twice, with two

different ligand concentrations (used in fitting): $L_{0,fit}$ and $L_{0,fit}'$. $K_{d,det}$ is found with $L_{0,fit} = L_0$ (nominal value), and $K_{d,det}'$ is found with $L_{0,fit}' = \alpha L_0$. As discussed in the theory section, multiplying the nominal L_0 by a factor α introduces a large "artificial" systematic error to L_0 in order to test the sensitivity of $K_{d,det}$ to changes in $\Delta L_0/L_0$. To ensure that $K_{d,det}'$ is distinguishable from $K_{d,det}$ — especially in the range of low L_0/K_d values — α must be significantly greater or smaller than unity. We choose α to be much less than 1 (*e.g.*, $\alpha = 0.005$) to create a substantial difference between $L_{0,fit}$ and $L_{0,fit}'$, while also ensuring that $K_{d,det}'$ remains positive (Note C1.3, **Table C1.1**).

According to the general theory and graphical strategy schematized in **Figure 4.3**, if the confidence intervals of $\Delta L_0/L_0$ and $\Delta T_0/T_0$ can be experimentally estimated and the boundary conditions of combinations of $\Delta L_0/L_0$ and $\Delta T_0/T_0$ (*i.e.*, $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$) can be determined, then we can numerically establish the ACI graph of "$K_{d,det}'/K_{d,det}$ vs L_0/K_d" and use it to find ACI of K_d. Based on these considerations, we propose a step-by-step ACI-Kd workflow for finding ACI of K_d value (**Figure 4.5**). ACI-Kd requires two procedures to be specified/developed: (*i*) a method for estimating the confidence intervals of the relative systematic errors of L_0 and T_0 and (*ii*) a method of establishing the ACI graph of "$K_{d,det}'/K_{d,det}$ vs L_0/K_d".

In K_d-determination experiments, all predictable systematic errors of variables must be eliminated with method optimizations including calibration of instruments, elimination of material adsorption to surface, the necessary mathematical corrections based on the purities of reagents, *etc*.[182, 317, 318, 319] When the systematic errors of concentrations are minimized, the unpredictable systematic errors of L_0 and T_0, which inevitably exist in each experiment, are from the normally-distributed systematic errors in the concentrations of their stock solutions.[182] Specifically, for each single preparation of a stock solution in the ACI-Kd workflow, the normal

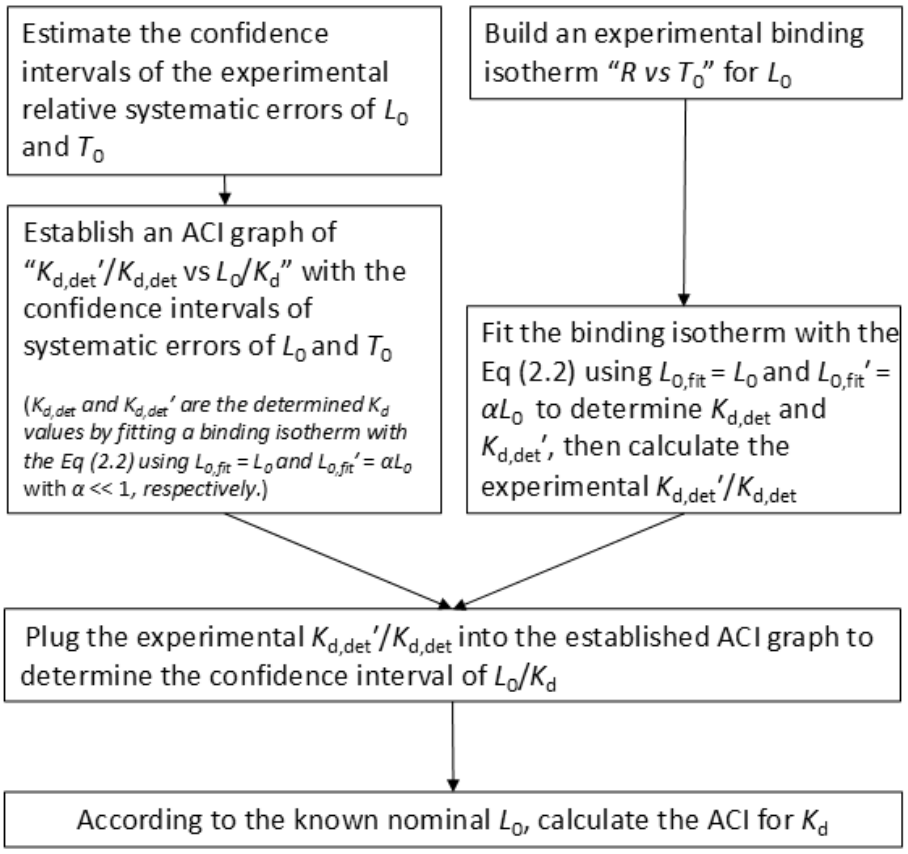


Figure 4.5. Step-by-step ACI-Kd workflow for finding ACI of K_d.

distribution of concentrations of stock solutions collapses to one single concentration, *i.e.*, the random error converts into the systematic error. With these considerations, the confidence intervals of $\Delta L_0/L_0$ and $\Delta T_0/T_0$ can be estimated from the relative random errors of stock L_0 and T_0 (*i.e.*, $\delta L_0/L_0$ and $\delta T_0/T_0$ for stock solutions). The values of $\delta L_0/L_0$ and $\delta T_0/T_0$ can be experimentally determined based on the spectroscopic signals (*e.g.*, light absorption or fluorescence intensity) from ligand and target in multiple preparations of stock solutions (examples are shown in the section describing applications of ACI-Kd to real binding experiments).

Based on the confidence intervals of $\Delta L_0/L_0$ and $\Delta T_0/T_0$, we can calculate the confidence level of ACI (**Figure 4.5**). In this work, the confidence levels of ACI and PCI will be harmonized, *i.e.*, presented with the same value. The confidence level of PCI can be varied by choosing different numbers of standard deviation (SD) as the uncertainty of the determined value $K_{d,det}$. Conventionally, PCI is presented as "mean value ± SD" with a confidence of 68.3%. To harmonize the confidence levels of PCI and ACI, our goal in this work was to use ACI-Kd to find ACI with 68.3% confidence. Since the probability distribution of ACI may not be normal, and therefore, cannot be presented as "mean value ± SD", we will present both PCI and ACI as ranges between the boundaries found with 68.3% confidence instead of the format of "mean value ± SD".[295]

To achieve a confidence level of 68.3% for ACI, by consulting the standard normal table (Z-score table) (**Figure C1.1**), we choose $\Delta L_0/L_0 \in [-1.36\delta L_0/L_0, 1.36\delta L_0/L_0]$ and $\Delta T_0/T_0 \in [-1.36\delta T_0/T_0, 1.36\delta T_0/T_0]$ as the confidence intervals of relative systematic errors of L_0 and T_0, respectively (*i.e.*, each concentration interval with a confidence level of 82.63%, and the combination of the concentration intervals with a confidence level of 82.63% × 82.63% ≈ 68.3%) (Note C1.4). Evidently, we can vary the confidence level of ACI by choosing different numbers of $\delta L_0/L_0$ and $\delta T_0/T_0$ in the confidence intervals of $\Delta L_0/L_0$ and $\Delta T_0/T_0$. Note that, to reduce the amount of experimental work required by the ACI-Kd workflow, we can estimate the confidence intervals of relative systematic errors of concentrations once and use these intervals for experiments conducted with similar samples prepared in the same way.

To establish the ACI graph of "$K_{d,det}'/K_{d,det}$ vs L_0/K_d", we need to determine the combinations of $\Delta L_0/L_0$ and $\Delta T_0/T_0$ that correspond to $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$, respectively. According to **Eq (4.10)** and the chosen confidence intervals of $\Delta L_0/L_0$ and $\Delta T_0/T_0$, we find that

$\mathbf{\Delta^*_{min}} = \{\Delta L_0/L_0 = 1.36\delta L_0/L_0,\ \Delta T_0/T_0 = -1.36\delta T_0/T_0\}$, and $\mathbf{\Delta^*_{max}} = \{\Delta L_0/L_0 = -1.36\delta L_0/L_0,\ \Delta T_0/T_0 = 1.36\delta T_0/T_0\}$ (Note C1.5), with which we can establish the two boundary curves in the ACI graph. To numerically build the two boundary curves, we first need to generate two simulated datasets of "R vs T_0" for a large range of L_0/K_d and for two systematic errors: $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$. Then, we need to fit the generated datasets with **Eq (2.2)** using $L_{0,fit} = L_0$ and $L_{0,fit}' = \alpha L_0$ ($\alpha << 1$, *e.g.*, $\alpha = 0.005$) to obtain $K_{d,det}$ and $K_{d,det}'$ at each L_0/K_d, respectively. As a result, the theoretical value of $K_{d,det}'/K_{d,det}$ at each value of L_0/K_d can be calculated, and the two boundary curves of "$K_{d,det}'/K_{d,det}$ vs L_0/K_d" in the ACI graph can be established (**Figure 4.3**). To be noted, we should use the nominal (instead of accurate) ligand concentration as L_0 for both fitting and calculating L_0/K_d in establishing the ACI graph, since the accurate L_0 is unknown in real experiments.

By now, we have both procedures required for the proposed ACI-Kd workflow (**Figure 4.5**) and the rest of the workflow is straightforward. When an experimental binding isotherm is obtained, it is fitted with **Eq (2.2)** twice, using $L_{0,fit} = L_0$ and $L_{0,fit}' = \alpha L_0$, to obtain experimental $K_{d,det}$ and $K_{d,det}'$, respectively. By plugging the experimental $K_{d,det}'/K_{d,det}$ into the established ACI graph, the confidence interval of L_0/K_d (with 68.3% confidence) is determined as illustrated in **Figure 4.3** (and **Eq (4.8)**). According to the computed confidence interval of the ratio of nominal L_0 to accurate K_d (L_0/K_d) and the known nominal L_0, ACI of K_d is calculated. Finally, an experimentalist can decide whether ACI is satisfactory or not based on the intended application of the ligand. If ACI is too large, the experiment must be performed with a lower L_0, which may require an instrument with a lower LOQ (an example is shown in the section describing applications of ACI-Kd to real binding experiments).

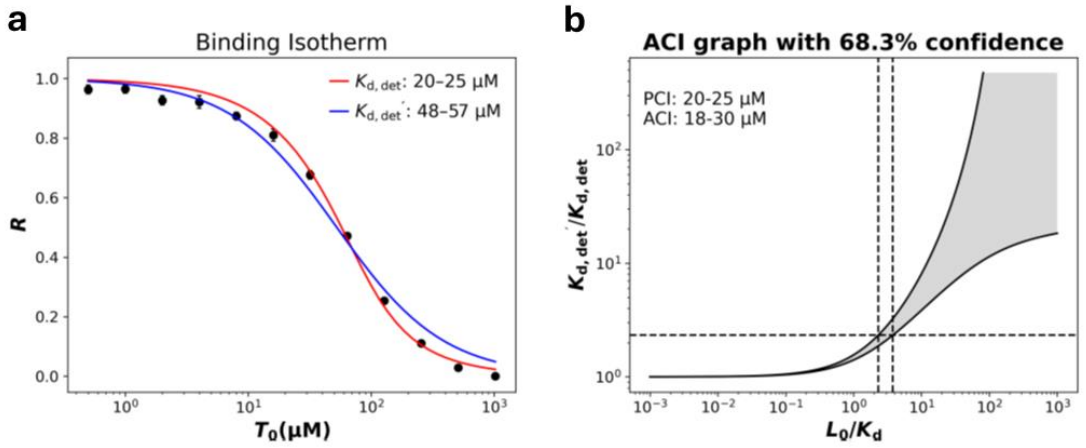


Figure 4.6. Example of binding isotherm **(a)** and ACI graph **(b)** produced by ACI-Kd webapp (http://aci.sci.yorku.ca). The binding isotherm was obtained from an ACTIS experiment using BAS–fluorescein as the binding pair, with a nominal ligand concentration (L_0) of 70 µM. Refer to the experimental demonstration section and Notes C1.8 and C1.9 for more experimental details.

Considering the computational complexity of the ACI-Kd workflow, we developed a dedicated program implemented as a user-friendly webapp (https://aci.sci.yorku.ca) to compute both ACI and PCI of K_d from a single binding isotherm. The detailed computational workflow (**Figure C1.2**) and algorithm (**Figure C1.3**) for the ACI-Kd webapp are shown in Note C1.6. The ACI-Kd webapp produces a report including the input and output data, a publication-quality binding isotherm, and the ACI graph (an example of the graphs is shown in **Figure 4.6**). This report can be downloaded in a PDF format ready for saving and distributing (*e.g.*, publishing) the results of the ACI analysis. For added flexibility in data interpretation, both ACI and PCI are computed for three confidence levels: 68.3, 95.5, and 99.7%.

Since the ACI-Kd workflow was developed to calculate the confidence interval of an accurate K_d by accounting for its fundamental sources of inaccuracy, specifically the systematic errors in

variables, this workflow can be applied to experimental data obtained using any K_d-determination method that employs nonlinear regression of binding isotherms. However, the current version of the ACI-Kd workflow is only applicable to one-to-one binding processes (**Eq (1.1)**) where K_d is the sole unknown parameter in the nonlinear regression (**Eq (2.2)**).

4.1.2.7. Verification of the ACI-Kd Workflow.

Verification of ACI-Kd requires showing that ACI-Kd is capable of producing theoretical results, which can be arranged without leaving the realm of solid theory. Since the accurate values (*e.g.*, K_d, L_0, T_0, and their systematic errors) in real experiments are fundamentally unknown and uncontrollable, the most reliable data source for verifying ACI-Kd is the simulated data with controllable variables and parameters. Thus, we simulated binding isotherms with **Eq (2.2)**; they are available as a supporting spreadsheet (Data S1 stored on Figshare website: https://doi.org/10.6084/m9.figshare.25777164.v3). In the simulations, the input values of K_d, L_0, and T_0 can be considered as true values without uncertainty. Simulating all the cases of accurate (true) K_d locating in a confidence interval with a known probability (*i.e.*, true ACI) was impractical because it required an infinite or a very large number of simulated binding isotherms. Thus, to simplify the verification process, we classified the infinite binding isotherms for true ACI into three scenarios: (*i*) true K_d was theoretically located inside the computed ACI, (*ii*) true K_d was theoretically close to one of the boundaries of computed ACI, and (*iii*) true K_d was theoretically outside the computed ACI.

First, we investigated if ACI-Kd can correctly produce the theoretically predicted results when the accurate K_d was theoretically located inside ACI. To conduct this investigation, with the spreadsheet (Data S1, https://doi.org/10.6084/m9.figshare.25777164.v3), we set $K_d = 1$ (unitless) and simulated the binding isotherms "R vs T_0" for a large range of L_0/K_d without

introducing any error. Then, we input the simulated binding isotherms, corresponding L_0 and reasonable relative random errors (*e.g.*, $\delta L_0/L_0 = \delta T_0/T_0 = 5\%$ in this study) of concentrations into ACI-Kd to compute ACI and PCI for each L_0/K_d. Since a zero error (used in the simulations) was at the centre of the assumed interval of systematic errors in concentrations (used in ACI-Kd), theoretically true $K_d = 1$ should always be located inside and close to the centre of the computed ACI. Because all the input concentrations in ACI-Kd were identical to the values used for simulating binding isotherms, the computed PCI should be very narrow and close to the true K_d value. Our results (**Figure 4.7**a) produced by ACI-Kd agreed with all the theoretical predictions.

Second, we studied whether ACI-Kd could correctly determine the ACI of K_d when the combinations of true systematic errors of L_0 and T_0 (*i.e.*, $\Delta L_0/L_0$ and $\Delta T_0/T_0$) were the boundary conditions, for which true K_d was theoretically located on the boundary of the computed ACI. To study the two boundary conditions, we simulated the binding isotherms using the spreadsheet (Data S1, https://doi.org/10.6084/m9.figshare.25777164.v3) by adding error combinations of "$\Delta L_0/L_0 = 1.36\delta L_0/L_0$ and $\Delta T_0/T_0 = -1.36\delta T_0/T_0$" ($\mathbf{\Delta^*_{min}}$) and "$\Delta L_0/L_0 = -1.36\delta L_0/L_0$ and $\Delta T_0/T_0 = 1.36\delta T_0/T_0$" ($\mathbf{\Delta^*_{max}}$) for a large range of L_0/K_d. Here, $\Delta L_0/L_0$ and $\Delta T_0/T_0$ were the true relative systematic errors added in the simulated binding isotherms, while $\delta L_0/L_0$ and $\delta T_0/T_0$ were the assumed relative random errors of L_0 and T_0 that would be input in ACI-Kd. By inputting the simulated binding isotherms, corresponding L_0, and the assumed relative random errors into ACI-Kd, we obtained that true $K_d = 1$ was close to the upper or lower limit of the computed ACI (depending on the introduced error combination in simulations) for the whole range of studied L_0/K_d (**Figure 4.7b1**, **b2**) as theoretically expected. According to **Figure 4.7b1**, **b2**, the computed PCI consistently shifted away from the accurate (true) K_d as L_0/K_d increased, as suggested by **Figure 4.4b** and **Eq (4.10)**. Additionally, the width of the PCI expanded with the

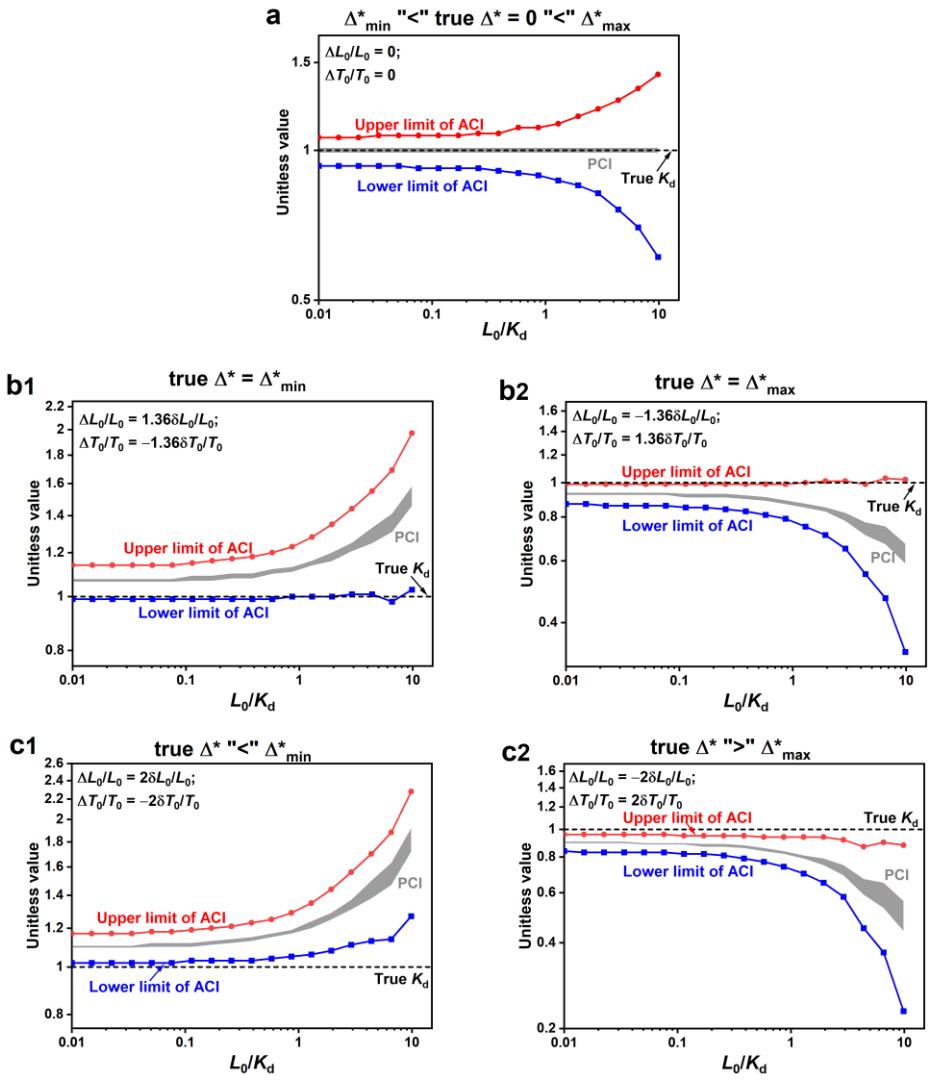


Figure 4.7. Verification of ACI-Kd with simulated binding isotherms — comparison of true K_d to computed ACI and PCI for: **(a)** true $\mathbf{\Delta^*} = \mathbf{0}$, which was between $\mathbf{\Delta^*_{min}}$ (*i.e.*, $\{\Delta L_0/L_0 = 1.36\delta L_0/L_0,\ \Delta T_0/T_0 = -1.36\delta T_0/T_0\}$ and $\mathbf{\Delta^*_{max}}$ (*i.e.*, $\{\Delta L_0/L_0 = -1.36\delta L_0/L_0,\ \Delta T_0/T_0 = 1.36\delta T_0/T_0\}$; **b1)** true $\mathbf{\Delta^*} = \mathbf{\Delta^*_{min}}$; **(b2)** true $\mathbf{\Delta^*} = \mathbf{\Delta^*_{max}}$; **(c1)** true $\mathbf{\Delta^*}$ "<" $\mathbf{\Delta^*_{min}}$ (*i.e.*, true $\mathbf{\Delta^*} = \{\Delta L_0/L_0 = 2\delta L_0/L_0,\ \Delta T_0/T_0 = -2\delta T_0/T_0\}$), and **(c2)** true $\mathbf{\Delta^*}$ ">" $\mathbf{\Delta^*_{max}}$ (*i.e.*, true $\mathbf{\Delta^*} = \{\Delta L_0/L_0 = -2\delta L_0/L_0,\ \Delta T_0/T_0 = 2\delta T_0/T_0\}$). For all verification steps, the input $\delta L_0/L_0 = \delta T_0/T_0 = 0.05$ in ACI-Kd, which was a reasonable value in experiments. To clearly show the detailed value change in large ranges of L_0/K_d, the results were presented in log-log scales.

increase in L_0/K_d (**Figure 4.7b1**, **b2**) due to the decreased precision of $K_{d,det}$, which resulted from increasing systematic errors in $K_{d,det}$ values calculated using different data points in a binding isotherm. Note that, in real experiments, there might be no specific trends for the dependence of the PCI width on L_0/K_d or the relationship between the widths of PCI and ACI. This is because either experimental random errors or experimental systematic errors can dominate the other type of errors for any given L_0/K_d.

Finally, we tested the cases of true K_d theoretically located outside the computed ACI. True K_d being outside the computed ACI meant that the combination of true systematic errors of L_0 and T_0 was beyond the interval of $\mathbf{\Delta^*}$ defined by $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$. To conduct this test, we simulated the binding isotherms across a large range of L_0/K_d, introducing two types of systematic errors. In the first scenario, we applied the error combination "$\Delta L_0/L_0 = 2\delta L_0/L_0 > 1.36\delta L_0/L_0$ and $\Delta T_0/T_0 = -2\delta T_0/T_0 < -1.36\delta T_0/T_0$" (*i.e.*, true $\mathbf{\Delta^*}$ "<" $\mathbf{\Delta^*_{min}}$) which is theoretically expected to cause the true K_d to fall below the lower limit of calculated ACI. In the second scenario, we used the error combination "$\Delta L_0/L_0 = -2\delta L_0/L_0 < -1.36\delta L_0/L_0$ and $\Delta T_0/T_0 = 2\delta T_0/T_0 > 1.36\delta T_0/T_0$" (*i.e.*, true $\mathbf{\Delta^*}$ ">" $\mathbf{\Delta^*_{max}}$) which theoretically results in the true K_d be greater than the upper limit of calculated ACI. Here, quotation marks are used for the ">" and "<" symbols to emphasize that vectors cannot be directly compared in magnitude. By inputting the corresponding data into ACI-Kd, we observed that true $K_d = 1$ was consistently outside the computed ACI across the entire range of L_0/K_d studied (**Figure 4.7c1**, **c2**), and the relationships between true K_d and the calculated ACI aligned with theoretical expectations. In **Figure 4.7c1**, **c2**, we also observed that with the increase of L_0/K_d, the computed PCI kept shifting away from the true K_d, and the PCI width increased with the increasing L_0/K_d (similar to **Figure 4.7b1**, **b2**).

To sum up the above, since the relationship between true K_d and computed ACI agreed with the theoretical predictions for all three scenarios, we confirmed that ACI-Kd can produce theoretically predicted results. The results in **Figure 4.7** also confirmed that the range of ACI increases with the increase of L_0/K_d, which was suggested by **Figure 4.3**. Additionally, **Figure 4.4b** and **Eq (4.10)** imply that, when there exists systematic error in L_0 and/or T_0, the difference between $K_{d,det}$ and true K_d (*i.e.*, $|K_{d,det} - K_d|$) increases with the increase of L_0/K_d. This dependence of $|K_{d,det} - K_d|$ on L_0/K_d was confirmed by the dependence of PCI (the center of PCI was $K_{d,det}$) on L_0/K_d in **Figure 4.7b1**, **b2**, **c1**, **c2**. Without adding any random errors to variables in the simulation, the PCI width increases with increasing L_0/K_d in **Figure 4.7b1**, **b2**, **c1**, **c2** because as L_0/K_d rises, $K_{d,det}$ values calculated with different data points in a binding isotherm are subject to larger systematic errors and thus distribute over a wider range, resulting in lower precision. However, this may not be the case in real experiments, which always include random errors in variables. Note that, since there was no systematic error in either L_0 or T_0 in the simulated binding isotherms for the study in **Figure 4.7a**, PCI in **Figure 4.7a** was always close to the true K_d with negligible width as expected. **Figure 4.7b1**, **b2**, **c1**, **c2** indicate that when there exist systematic errors in L_0 and T_0, which is typically the case, solely relying on PCI can potentially either highly underestimate or overestimate (depending on the systematic errors in independent variables) K_d values. Our results in **Figure 4.7** show that, compared to PCI, ACI computed with ACI-Kd is much more reliable to show the possible value range of accurate K_d when there are errors in concentrations (which is always the case in experiments). Now, let's apply ACI-Kd to real experiments and provide a comprehensive interpretation of the experimental results.

4.1.2.8. Examples of ACI-Kd Application to Results of Binding Experiments.

To demonstrate the application and verify the capability of ACI-Kd in K_d-determination experiments, we conducted both virtual and real experiments involving various binding pairs. The results of the virtual experiments (Note C1.7, including **Figure C1.4**) confirmed that ACI-Kd can accurately compute the confidence intervals of K_d when provided with the correct confidence intervals of relative systematic errors in concentrations. In the real experiments, we studied two binding pairs with two different K_d-determination methods: (*i*) BSA–fluorescein with the ACTIS method (Accurate Constant *via* Transient Incomplete Separation),[171, 172, 177, 178] and (*ii*) HSP90α-mNeonGreen–mScarlet-I-CDC37 with the FRET method (Förster resonance energy transfer).[320] Experimental details can be found in Notes C1.8–C1.12 (including **Figures C1.5**–**C1.7** and **Tables C1.2**–**C1.8**).

First, we demonstrate the application of ACI-Kd to the results of ACTIS experiments. We chose BSA and fluorescein as binding partners because this pair has been well studied and has a relatively large consensus K_d value in the micromolar range,[171, 321, 322, 323] which can give us enough room for varying L_0/K_d while still satisfying the requirement that L_0 should be greater than the LOQ of the ACTIS instrument for fluorescein (in the picomolar range). In four ACTIS experiments, the total fluorescein concentrations $[\text{fluorescein}]_0$ (nominal L_0) were prepared to be 0.10, 10, 100, and 300 μM. The purity of the product was 70% according to the product specification of the used fluorescein sodium salt. Thus, $[\text{fluorescein}]_0$ (nominal L_0) was corrected to be 0.070, 7.0, 70, and 210 μM, respectively. By measuring the signals of 5 independently prepared fluorescein stock solutions with a spectrophotometer, $\delta L_0/L_0$ was estimated to be 0.023 (**Table C1.2**). For all four experiments, total BSA concentration $[\text{BSA}]_0$ (nominal T_0) was varied from 0 (the lowest non-zero concentration was 0.1 μM or 0.5 μM) to 1,000 (or 1,024) μM, which was much greater than the expected K_d value to ensure the binding was saturated. The

high purity of the BSA protein used in this study was confirmed by SDS-PAGE (**Figure C1.5**). By measuring the UV-Vis absorbance of BSA stock solutions with a spectrophotometer in 5 independent sample preparations, $\delta T_0/T_0$ was estimated to be 0.082 (**Table C1.3**). Then, with ACTIS experiments, the experimental binding isotherms "R vs $[BSA]_0$" for different $[fluorescein]_0$ were produced (**Figure 4.8a**). By inputting $\delta L_0/L_0 = 0.023$, $\delta T_0/T_0 = 0.082$, and the

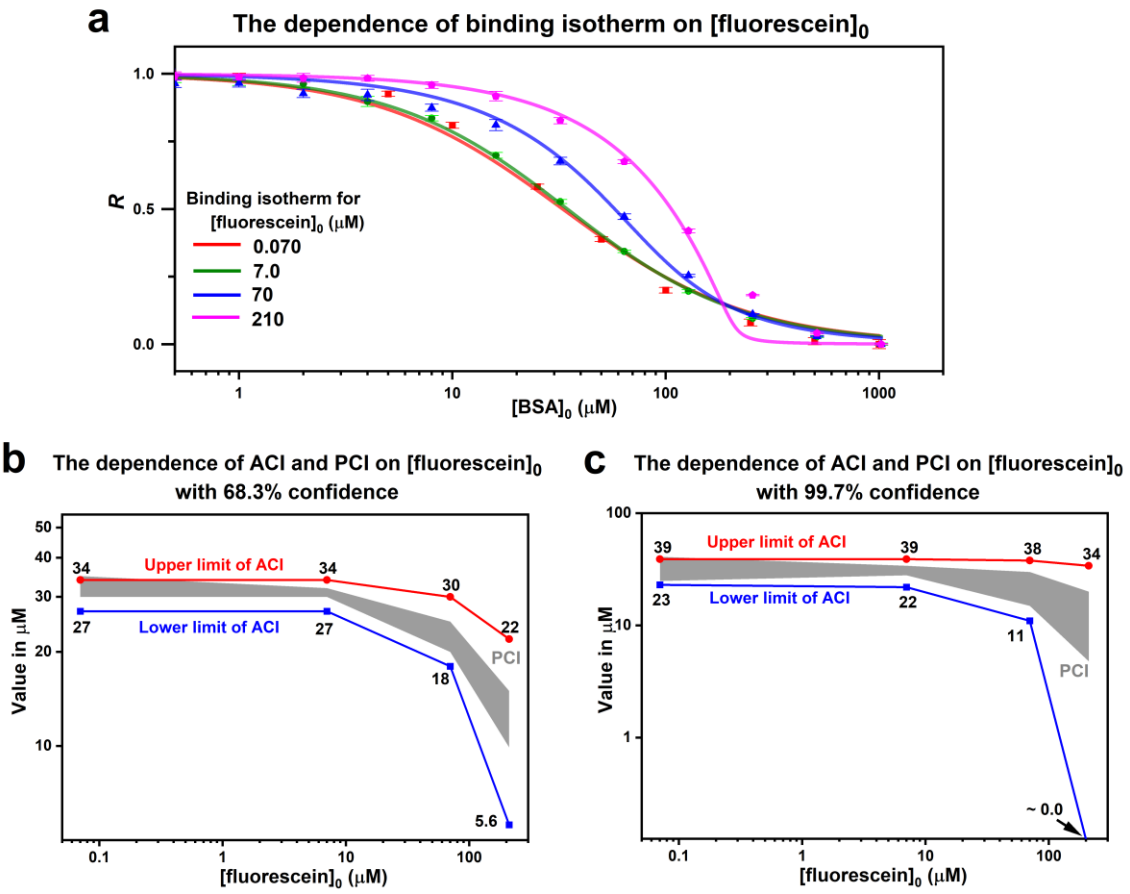


Figure 4.8. Application of ACI-Kd to results of ACTIS experiments with BSA–fluorescein as the binding pair: **(a)** the influence of change in $[fluorescein]_0$ (nominal L_0) on binding isotherms; **(b)** the dependence of computed ACI and PCI on $[fluorescein]_0$ with 68.3% confidence, and **(c)** the dependence of computed ACI and PCI on $[fluorescein]_0$ with 99.7% confidence. Note the log-log scale in panels **(b)** and **(c)**.

binding-isotherm data corresponding to [fluorescein]$_0$ = 0.070, 7.0, 70, and 210 μM into ACI-Kd, PCI and ACI (with 68.3% confidence) of K_d were found (**Figure 4.8b**). Note that, all the boundary values of ACI and PCI presented in **Figure 4.8** were with no more than two significant figures, and the rest of this paper will follow this rule because errors in K_d were unlikely smaller than 10%.

Figure 4.8b indicates that, with high ligand concentration (*e.g.*, [fluorescein]$_0$ = 210 μM), if we conclude the value of K_d only based on the underestimated narrow range of PCI (*e.g.*, 9.9–15 μM), we may overestimate the level of affinity between BSA and fluorescein. The large range of ACI (*e.g.*, 5.6–22 μM) with 68.3% confidence at [fluorescein]$_0$ = 210 μM suggests that the accurate K_d can be different from PCI with high probability, and researchers should pay attention to the large uncertainty of accurate K_d suggested by wide ACI and decide if additional experiments are required to obtain a narrower ACI. Therefore, our results clearly show the necessity of assessing ACI in K_d determinations since the conclusions based on PCI can be misleading. It was noticed (from **Figure 4.8b**) that, with 68.3% confidence, the upper limit of ACI (*i.e.*, 22 μM) determined with [fluorescein]$_0$ = 210 μM was smaller than the lower limit of ACI (*i.e.*, 27 μM) determined with [fluorescein]$_0$ = 0.070 or 7.0 μM. The possible reasons for this phenomenon include the underestimated width of confidence intervals for relative systematic errors in L_0 and T_0 and/or disregarding the systematic errors of R. This emphasizes that the ACI of K_d computed with ACI-Kd workflow is the narrower limit of the confidence interval of accurate K_d as we explained in the theory-development part. To obtain the range of accurate K_d with a higher confidence level, we increased the confidence level of ACI to 99.7%. With the increased confidence level of 99.7%, ACI of K_d for nominal [fluorescein]$_0$ = 0.070, 7.0, 70, and 210 μM were calculated to be 23–39 μM, 22–39 μM, 11–38 μM, and ~0.0–34 μM,

respectively (**Figure 4.8c**). As expected, with the higher confidence interval, the range of computed ACI became wider, and the ACI determined with different ligand concentrations were highly overlapped.

Then, to further confirm the capability of ACI-Kd in the research on molecular interactions, we applied ACI-Kd to the results of K_d-determination experiments conducted with the FRET method for a protein–protein binding pair of HSP90α-mNeonGreen–mScarlet-I-CDC37. Since HSP90α–CDC37 is a potential drug target in cancer, knowing the accurate K_d of this binding pair is crucial for corresponding drug (*e.g.*, inhibitor) development.[324, 325, 326] Therefore, our goal was to find a reasonably narrow ACI (68.3% confidence) of K_d with the ratio between the upper-limit value and lower-limit value being < 1.5 since two K_d values different by less than 1.5 folds are often considered indistinguishable. Note that, users of ACI-Kd can define their own targeted width of ACI based on the intended application of the evaluated ligand.

To find the ACI of K_d for HSP90α-mNeonGreen–mScarlet-I-CDC37, we first conducted the FRET K_d-determination experiment using a Tecan Infinite M200Pro plate reader (Männedorf, Switzerland). In this experiment, the total ligand concentration $[\text{HSP90α-mNeonGreen}]_0$ (*i.e.*, nominal L_0) was chosen to be 990 nM, while the total target concentration $[\text{mScarlet-I-CDC37}]_0$ (*i.e.*, nominal T_0) varied from 0 (the lowest non-zero concentration was 15 nM) to 4,500 nM. With SDS-PAGE, the purity of HSP90α-mNeonGreen protein was determined to be 74.5%, and the purity of mScarlet-I-CDC37 was determined to be 79.5% (**Figure C1.6**). Thus, the nominal L_0 was corrected to be 740 nM, and the nominal T_0 was corrected to be from 0 (the lowest non-zero concentration was 10 nM) to 3,500 nM. With a spectrophotometer, $\delta L_0/L_0$ and $\delta T_0/T_0$ for the protein stock solutions were estimated to be 0.024 and 0.054, respectively (**Tables C1.7** and **C1.8**). By inputting the obtained binding isotherm (**Figure 4.9a**) and relative errors of L_0 and T_0

into ACI-Kd, PCI of K_d was determined to be 180–250 nM and ACI of K_d was determined to be 150–220 nM. Both PCI and ACI demonstrated here are with 68.3% confidence. The ACI found here was slightly wider than our target range defined by the condition that the ratio between the

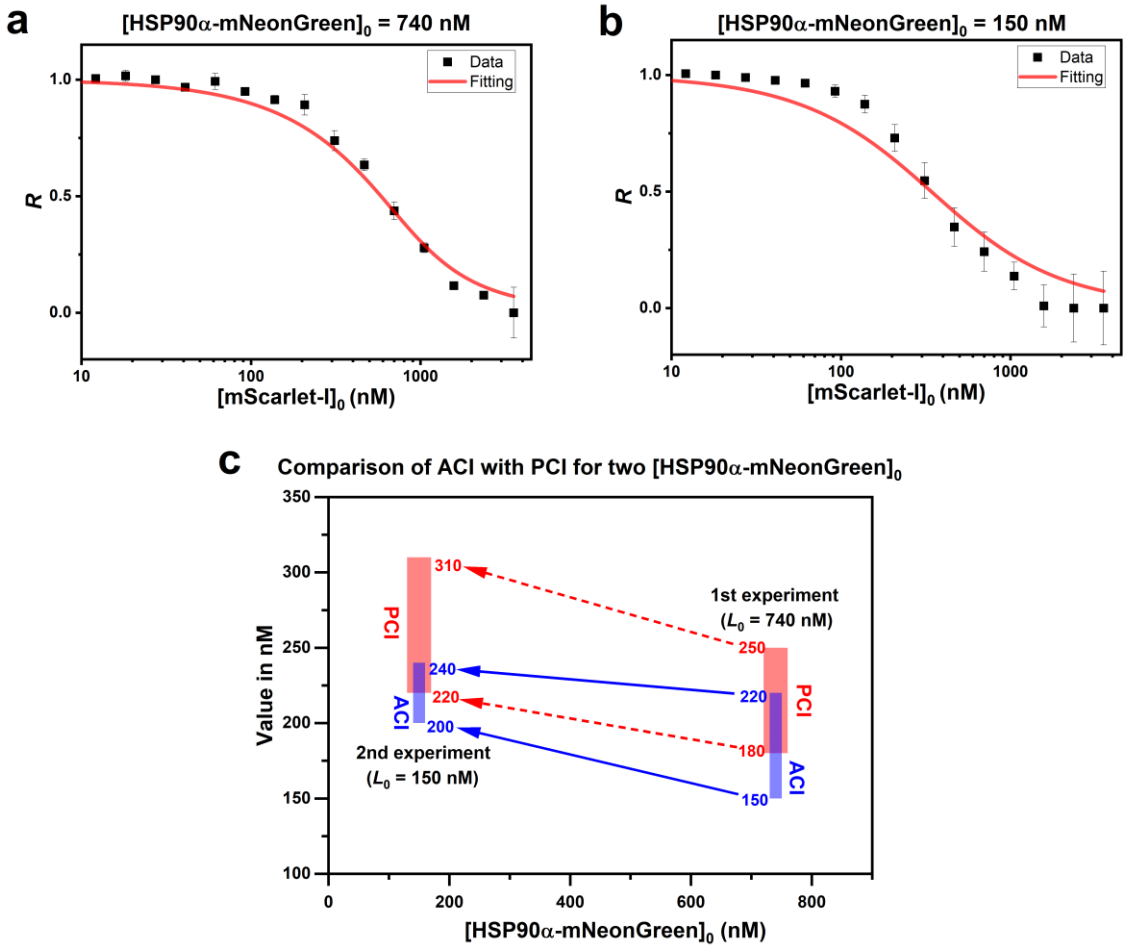


Figure 4.9. Application of ACI-Kd to results of FRET experiments with HSP90α-mNeonGreen–mScarlet-I-CDC37 as the binding pair: **(a)** the binding isotherm of the experiment with total ligand concentration L_0 = [HSP90α-mNeonGreen]$_0$ = 740 nM; **(b)** the binding isotherm of the experiment with total ligand concentration L_0 = [HSP90α-mNeonGreen]$_0$ = 150 nM, and **(c)** the comparison of ACI with PCI for the experiments with [HSP90α-mNeonGreen]$_0$ = 150 and 740 nM. All the ACI and PCI shown in panel **(c)** are with a confidence level of 68.3%.

upper-limit value and lower-limit value of K_d be < 1.5; in our case, this ratio is 220/150 ≈ 1.5. Since L_0 = 740 nM approaches the LOQ (≈ 630 nM) of Tecan Infinite M200Pro plate reader for HSP90α-mNeonGreen in the HSP90α-mNeonGreen–mScarlet-I-CDC37 binding experiment (**Figure C1.7a**), we decided to use a FRET instrument with a lower LOQ to re-conduct the K_d-determination experiment with lower L_0 (*i.e.*, $[\text{HSP90α-mNeonGreen}]_0$).

In the second experiment, we used a Tecan Infinite F200Pro plate reader (Männedorf, Switzerland) with LOQ ≈ 140 nM for HSP90α-mNeonGreen (**Figure C1.7b**) to conduct the study. Similar to the first experiment, the total target concentration $[\text{mScarlet-I-CDC37}]_0$ (*i.e.*, nominal T_0) was varied from 0 (the lowest non-zero concentration was 10 nM) to 3,500 nM in the second experiment, while the total ligand concentration $[\text{HSP90α-mNeonGreen}]_0$ (*i.e.*, nominal L_0) was lowered to 150 nM (after concentration correction based on the product purity) and kept constant. By inputting the binding isotherm obtained from the second experiment (**Figure 4.9b**), $\delta L_0/L_0 = 0.024$ and $\delta T_0/T_0 = 0.054$ into ACI-Kd, PCI and ACI (68.3% confidence) of K_d were determined to be 220–310 nM and 200–240 nM, respectively (demonstrated in **Figure 4.9c**). Since the ACI found in the second experiment achieved our targeted width (240/200 = 1.2 < 1.5), no more experiment was required, and we can finally conclude the K_d of HSP90α-mNeonGreen–mScarlet-I-CDC37 as 200–240 nM with 68.3% confidence. The comparison of ACI and PCI for the two FRET experiments (**Figure 4.9c**) reveals the following. First, the width of ACI decreased with decreasing the ligand concentration, which was expected. Second, ACI determined with the two different ligand concentrations greatly overlapped each other. Third, the width of PCI increased with decreasing ligand concentration due to the increased random errors for the data points with the lower ligand concentration. Notably, all the

observations of the experimental results for ACTIS and FRET experiments agree with our theory and the corresponding discussions.

Finally, we compared the confidence intervals produced by ACI-Kd with those obtained through bootstrapping, a statistical method based on Monte Carlo simulation. For this comparison, we developed a Python script utilizing publicly available packages, incorporating the fitting model for K_d determination (**Eq (2.2)**). The computational workflow for bootstrapping is illustrated in Note C1.13 (including **Figure C1.8**), which differs fundamentally from the computational algorithm for ACI-Kd shown in **Figure C1.2**. **Figure 4.10** presents the results of the comparison for two confidence levels: 68.3% and 99.7%. In **Figure 4.10**, the confidence intervals of K_d computed by both the ACI-Kd and bootstrapping methods are comparable and exhibit similar trends: the width of the confidence intervals increases, while the values of both

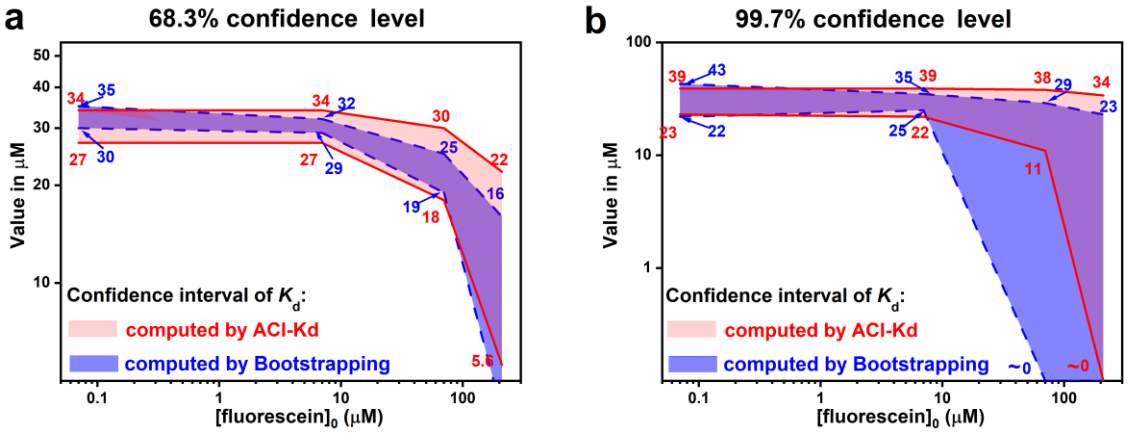


Figure 4.10. Comparison of the confidence intervals of K_d determined with ACI-Kd and bootstrapping approach for two confidence levels: 68.3% **(a)** and 99.7% **(b)**. The binding isotherms obtained from the ACTIS experiments for BSA–fluorescein were studied. A total of 5,000 iterations were used in the bootstrapping procedure, which was confirmed to be sufficient for achieving convergent results. Note the log-log scale in both panels.

interval bounds tend to decrease as the ligand concentration rises. We propose that the observed quantitative differences arise from the distinct strategies each method employs to define confidence levels. However, further investigation is necessary to fully understand the nature of these differences, which is beyond the scope of this study.

4.1.3. Concluding Remarks

Nonlinear regression is widely used for data analysis in physical sciences, engineering, and economics. This mathematical procedure can produce largely inaccurate values of determined parameters with apparently small uncertainties of regressions. The uncertainties of regression define PCI, which may exclude the accurate value of the sought parameter with a very high probability. To account for this inherent problem of nonlinear regression, we introduce ACI, a range that includes the accurate value of the determined parameter with a chosen confidence level (probability). We propose ACI-Concept, an approach for finding ACI, which uses language and concepts comprehensible to molecular scientists. ACI-Concept includes a strategy for developing practical workflows for computing ACI.

PCI and ACI are not correlated as they result from different types of errors: PCI results from random errors of the dependent variable, while ACI results from systematic errors of the variables. PCI has a normal (Gaussian) probability distribution (**Figure 4.1**) with the width depending on random errors of the dependent variable (assuming that the dependent variable does not have systematic errors and the regression model is correct). The random errors can be very small (*e.g.*, relative standard deviation < 10%) even for largely inaccurate values of the determined parameters. In contrast, ACI is expected to have a non-Gaussian probability distribution in most cases because the contributions of the normally distributed variables to the determined parameter are very likely to be unequal.[295] ACI depends only on how much the

systematic errors in all variables are magnified by a specific choice of the fixed independent variable. ACI may be wider than PCI for one range of fixed variables, and narrower for another range (*e.g.*, **Figure 4.9**).

At first glance, it may be attractive to integrate all errors in one parameter and produce a single confidence interval as a combination of ACI and PCI; however, this is a conceptual mistake as accuracy and precision are not "miscible". Therefore, it is recommended to use both ACI and PCI as independent characteristics (of accuracy and precision) of nonlinear regression. More studies are needed to better understand a multitude of potential combinations of ACI and PCI and ways to interpret them.

Our use of ACI-Kd revealed that PCI may overestimate or underestimate K_d, *i.e.*, relying only on PCI may lead to false-positive or false-negative results. Consequently, high-potency ligands can be dismissed, and low-potency ligands can instead be brought into further development of drugs and diagnostic probes if only PCI is used for decision-making in the hit-to-lead stage of drug discovery. Reporting both PCI and ACI of the determined K_d provides valuable insight into the precision and accuracy of the results. For a reliable K_d value, the PCI and ACI should be narrow and exhibit significant overlap.

The current ACI-Kd workflow computes the narrower limit of ACI of K_d, disregarding the systematic errors in R due to the lack of a strategy to estimate the confidence interval of ΔR. When such a strategy becomes available, ACI-Kd can be easily upgraded to compute ACI of K_d involving the systematic errors of all variables by slightly modifying the algorithm to establish the ACI graph. Without considering the systematic error in R, the lower limit of the relative systematic error in K_d values is a combination of relative systematic errors in L_0 and T_0, as suggested by error propagation rules.[179] As we have explained, systematic errors in L_0 and T_0

stem from their random errors when stock solutions of ligand and target are prepared once for the entire series of binding experiments, which is typically the case. The relative random errors in concentrations can hardly be below 1% even with the highest diligence in experiments. Our assessment shows that, even after careful method optimization and concentration correction based on product purity, the errors in protein concentrations can approach approximately 10%. When these errors are propagated, they will inevitably lead to errors in K_d, which are almost always greater than 10% (even without considering the potential contribution of ΔR to the errors of K_d). Therefore, it is impractical to present K_d values with more than two significant figures. As a rule, the boundaries of ACI and PCI must be presented with only one to two significant figures.

It is important to emphasize that systematic errors in concentrations can lead to an incorrect ranking of ligands in high-throughput screening despite the common belief that systematic errors shift all K_d values (or PCI of K_d using our terminology) in the same direction and, thus, do not affect the ranking order of ligands. It is obvious that ligands with low K_d can have high L_0/K_d values, and accordingly, judging only PCI can grossly overestimate K_d values, thus, mistakenly presenting such ligands as low-potency ones. Using both ACI and PCI for ligand ranking should greatly improve the quality of ranking and, thus, prevent waste of R&D efforts unavoidable when wrong leads are chosen for the development of drugs and diagnostic probes.

As finding ACI of K_d is now possible and simple, we recommend that reporting both ACI and PCI (ACI-Kd computes both) be adopted for publishing quantitative results of binding experiments. Taking on this recommendation will require reaching a consensus by the scientific community. We call for a constructive discussion on this subject within a public forum on Microsoft Teams (SC-ACI | General | Microsoft Teams) created for this purpose. Editors of

peer-reviewed journals who have expertise in this field are encouraged to discuss with their editorial boards adding the requirement of quantitative accuracy assessment for K_d to journal policies. The journals that frequently publish the results of binding experiments should advise their reviewers to pay attention to the quality of binding experiments and the need for accuracy assessments of K_d values. Developing standard requirements for binding experiments and the presentation of their results is an urgent need.[11] The joint effort of the authors, reviewers, and editors can help us to overcome the "pandemic" of inaccuracy in Kd of biomolecular complexes.

4.2. Extended Application of ACI: Introducing Quantitative Assessment of Michaelis Constant (K_m) Accuracy

The following material is reprinted from the article "**Wang, T. Y.**; Dhillon, P.; Schreiber, S.; Krylova, S. M.; Golemi-Kotra D.; Jose, J.; Kr ylov, S. N. Introducing Quantitative Assessment of Michaelis Constant (K_m) Accuracy" that has been submitted to *Analytical Chemistry* for peer review. My contributions to the article were: (*i*) conducting a portion of the mathematical analysis, (*ii*) contributing to program development, (*iii*) performing virtual experiments, (*iv*) preparing all figures, (*v*) interpreting the results, and (*vi*) writing the manuscript. Co-author Parmeetpal Dhillon discovered the mathematical similarity between K_d and K_m determinations. Co-author Sebastian Schreiber conducted the K_m-determination experiments for LDH–pyruvate pair.

4.2.1. Introduction: The Critical Need for Assessing K_m Accuracy

Enzymatic reactions are fundamental to most biochemical transformations requiring catalysis. These catalytic reactions proceed through the formation of an intermediate enzyme–substrate

complex. The Michaelis constant (K_m), a key quantitative parameter of enzymatic activity, provides insights into the affinity of the enzyme for the substrate and reflects the stability of the intermediate complex under pseudo-steady-state conditions.

Accurate K_m values are essential for decision-making across various applications, making the accuracy of K_m paramount. In industrial biotechnology, accurate K_m values ensure the selection of the right enzymes for optimum catalysis, leading to optimized production and cost efficiency.[327] In drug development, accurate K_m values assist in designing effective enzyme inhibitors and understanding drug metabolism, thereby influencing efficacy and dosage.[328, 329, 330, 331] Enzymes are also used as biopharmaceuticals, in which accurate K_m values also define the required dosage.[332] In metabolic engineering and synthetic biology, accurate K_m values facilitate the optimization of metabolic pathways, ensuring balanced fluxes and minimizing bottlenecks.[333] Environmental biotechnology relies on accurate K_m values to match enzyme activity with pollutant concentrations for effective bioremediation and waste treatment.[334] In the food and beverage industry, knowing the accurate K_m allows for precise control of enzymatic processes to achieve desired product qualities while maintaining consistency and quality. [335, 336] Overall, accurate K_m values are fundamental for enhancing efficiency, productivity, and specificity across various applications, underscoring the critical need for accurate K_m determination in decision-making processes.

The current standard for determining K_m is the nonlinear regression of the Michaelis-Menten (MM) plot, which graphically represents the relationship between the initial reaction velocity (v_0) and the initial substrate concentration (S_0).[337, 338, 339] Nonlinear regression fits the MM equation to experimental data, allowing for the estimation of both K_m and the maximum reaction velocity ($v_{0,max}$). The values of K_m and $v_{0,max}$ obtained through nonlinear regression are

accompanied by their respective standard errors, reflecting the precision of the estimated parameters. For example, experimentally determined K_m ($K_{m,det}$) is typically reported in the format "$K_{m,det} \pm \sigma$", where σ represents the standard error derived from the fit. The range defined by $K_{m,det} - \sigma$ to $K_{m,det} + \sigma$ is referred to as the Precision Confidence Interval (PCI), as it reflects the precision of the estimated parameter based on the standard error of the nonlinear regression.[10]

Several nonlinear regression models with slightly different performances have been suggested for fitting the experimental data to find PCI of K_m.[340, 341] It is common for $K_{m,det}$ values obtained for the same enzyme–substrate pair under similar conditions to differ significantly, meaning the PCIs do not overlap.[342, 343] Such discrepancies indicate that K_m determination is prone to systematic inaccuracies in addition to random errors. This limitation poses a problem, as it can lead to misinterpretation of enzyme affinity and incorrect decisions in applications where accurate K_m values are critical. The issue is further exacerbated by the absence of a practical, easy-to-use tool to quantitatively assess K_m accuracy without requiring additional experimentation. This work aims to address that problem.

Quantitative assessment of K_m accuracy involves determining the unknown systematic error in $K_{m,det}$, defined as the difference between the experimentally determined $K_{m,det}$ value and its true value (K_m). Since the true K_m value is fundamentally unknowable, assessing the accuracy of $K_{m,det}$ presents a significant challenge. This issue is not unique to K_m; similar challenges arise with most quantitative parameters determined via nonlinear regression.[3, 4, 5, 6, 7, 8] In particular, equilibrium dissociation constants (K_d) of affinity complexes can be grossly inaccurate when the constant (limiting) component is excessively large leading to the amplification of systematic errors in concentration that propagate to the systematic error in determined K_d during the fitting

process.[34, 179, 182]

It is important to note, however, that inaccuracies in K_m determination are generally not expected to reach the same magnitude as those in K_d determination, primarily due to the kinetic nature of K_m determination. Conditions that would result in significant inaccuracies in K_m often also disrupt the observation of the pseudo-steady state in enzymatic kinetics. This disruption typically prompts experimentalists to adjust conditions (*e.g.*, substrate or enzyme concentrations, reaction time), which not only helps restore pseudo-steady-state conditions but also improves the accuracy of K_m as a secondary benefit. Additionally, because K_m values for commonly studied enzyme–substrate pairs are generally much higher than K_d values for commonly studied ligand–target pairs, the experimental conditions required for accurate K_m determination are typically easier to achieve than those necessary for accurate K_d determination. Nevertheless, these inherent advantages of K_m over K_d do not eliminate the need for a rigorous quantitative assessment for the accuracy of $K_{m,det}$.

We recently proposed an approach to address this fundamental problem by introducing the concept of the Accuracy Confidence Interval (ACI) for any parameter determined using a correct regression model.[10] In this approach, the correct regression model serves as a reference standard in place of the true value of the parameter, which is inherently unknowable. The ACI represents a range, bounded by two values, within which the true value is expected to lie with a specified probability. These bounding values are computed by combining error-propagation and regression-stability analyses.[10]

The error-propagation analysis converts the experimentally determined random errors of variables (*e.g.*, concentrations) into estimates of the confidence intervals (CIs) for the systematic errors of those variables. These CIs are then used to define boundary conditions for determining

the systematic error of the sought parameter (*e.g.*, K_m). The validity of this conversion between random errors of variables and CIs of systematic errors depends on adherence to certain experimental protocols and assumptions.[10, 182]

The regression-stability analysis investigates, in silico, the stability of nonlinear regression for given values of variables and their associated errors. This analysis evaluates how variations in input variables influence the regression outcome. The resulting Accuracy Confidence Interval (ACI) defines the range within which the true value of the sought parameter is expected to lie, with the probability specified by the confidence levels used in the analysis.

In the mentioned recent proof-of-principle work on ACI, we used the ACI concept to develop a workflow for computing the ACI of K_d.[10] For this, we utilized results from an earlier study in which errors in variables (specifically, concentrations of reactants and the unbound fraction of the constant reactant) were propagated to calculate the errors in determined K_d for an affinity complex with 1:1 stoichiometry.[10, 179, 182]

The current work started with an attempt to perform similar error propagation, which is cumbersome but straightforward. While performing this analysis, we found that the K_m determination is mathematically identical to the K_d determination. To the best of our knowledge, this fundamental similarity has not been systematically studied, despite its significant methodological implications. First, this mathematical identity provides a universal theoretical basis for the determination of K_m and K_d. It also suggests that a classic regression model used for K_d determination through fitting a sigmoidal binding isotherm is applicable to K_m determination. Furthermore, this finding implies that the workflow developed by us for the computation of the ACI of K_d is immediately applicable to the determination of ACI of K_m upon renaming, for the convenience of users, the terms: ligand to enzyme, target to substrate, and K_d to K_m.

When applied to K_m determination, the binding-isotherm model uses a single unknown, K_m, assuming that $v_{0,max}$ can be determined experimentally. This assumption limits the applicability of our ACI approach to enzymatic reactions for which saturation can be reached in M-M plots. For such cases, the experimentally measured $v_{0,max}$ is approximately equal to the true $v_{0,max}$ and the accuracy of $K_{m,det}$ is improved by reducing the number of unknown parameters from two to one.[344, 345] This consideration suggests that the binding isotherm model should be used for MM plots with saturation, as it is not only more accurate than other models but also allows the assessment of K_m accuracy using the comprehensive ACI approach.

To facilitate quick and widespread adoption of our approach by the research community, we included the determination of ACI for K_m as an option in the ACI webapp (https://aci.sci.yorku.ca). The assessment of $K_{m,det}$ accuracy, along with the standard assessment of $K_{m,det}$ precision, will greatly improve the reliability of studies in enzyme kinetics and decrease the likelihood of misconceptions caused by inaccurate $K_{m,det}$ values.

4.2.2. Materials and Methods

4.2.2.1. Materials and Solutions

Both the NADH disodium salt and pyruvate acid sodium salt were purchased from Carl Roth GmbH + Co. KG (Karlsruhe, Germany). The lactate dehydrogenase (LDH) from rabbit muscle was purchased from Merck KGaA (Darmstadt, Germany). PBS buffer was obtained from Carl Roth GmbH + Co. KG (Karlsruhe, Germany). All solutions were prepared in the PBS buffer.

Kinetic Assays

The kinetic assays were performed in flat-bottom 96-well plates (Greiner Bio-One GmbH, Frickenhausen, Germany) using a microplate reader (Tecan Infinite M200Pro, Tecan, Männedorf, Switzerland). All measurements were performed in a final volume of 200 µL at

23 °C. The experiment was performed with an LDH concentration of 0.87 nM. The concentration of NADH was kept constant at 500 μM in all experiments. The pyruvate concentration was varied from the lowest nonzero concentration of 23 μM to 2000 μM in 1.5-fold increments, resulting in 13 data points, including one at 0 μM. After the addition of pyruvate, the 96-well plate was immediately placed in the microplate reader, and the formation of NAD^+ was measured at 340 nm every 10 s for a total of 490 s. The values of v_0 were determined from the change in NAD^+ concentration within the first 250 s and used to generate MM plots ("v_0 vs $[\text{pyruvate}]_0$"). All measurements were performed in triplicate.

4.2.3. Results and Discussion

4.2.3.1. Mathematics of K_m Determination

The mathematics of a classic way of K_d determination using a nonlinear regression of a binding isotherm has been thoroughly explained in ***Section 3.2.2.3*** of **Chapter 3**. In this section, we will explore the similarity between the mathematical frameworks used for K_m and K_d determination.

To formalize the mathematics of K_m determination, we consider a classic mechanism of enzymatic reaction:

$$E + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C \xrightarrow{k_{cat}} E + P \qquad (4.11)$$

where E is the enzyme, S is the substrate, C is the intermediate enzyme–substrate complex, P is the product, and k_{on}, k_{off}, and k_{cat} are complex formation rate constant, complex dissociation rate constant, and catalytic rate constant, respectively. This mechanism assumes that the rate-limiting step is that of product formation and the reaction rate is measured under the conditions where very little product is formed, so substrate depletion is negligible. The rate under these conditions

is referred to as the initial reaction velocity, $v_0 = k_{cat}C$.[346, 347] At the beginning of the pseudo-steady state, no product has been formed yet, enabling the determination of v_0.

In the pseudo-steady state, the rates of formation and consumption of C are assumed to be equal, resulting in:

$$\frac{dC}{dt} = k_{on}ES - \left(k_{cat} + k_{off}\right)C = 0 \tag{4.12}$$

From this equation, K_m is defined as:

$$K_m = \frac{k_{cat} + k_{off}}{k_{on}} = \frac{ES}{C} \tag{4.13}$$

which is mathematically analogous to **Eq (2.2)** for K_d.

Under the conditions where v_0 is measured (*i.e.*, when the amount of the product formed is negligible), the mass balance for the total concentrations of enzyme and substrate can be expressed as:

$$\begin{aligned} E + C &= E_0 \\ S + C &= S_0 \end{aligned} \tag{4.14}$$

where E_0 is the total enzyme concentration and S_0 is the total substrate concentration. This mass balance is equivalent to **Eq (3.10)** for mass balance in the binding process.

Let us introduce the fraction R of free enzyme:

$$R = \frac{E}{E_0} \tag{4.15}$$

Using the definitions of K_m and R (**Eqs (4.13)** and **(4.15)**) as well as mass balance equations (**Eq (4.14)**) we can link K_m to S_0, E_0, and R:

$$K_m = \frac{S_0 - E_0(1-R)}{(1/R-1)} \tag{4.16}$$

This expression is equivalent to **Eq (3.1)** for K_d. Similar to K_d, we need an experimental method to determine R. In enzymatic reactions, researchers typically measure v_0. We can express v_0 as the reaction rate (dP/dt) at the beginning of the pseudo-steady state and link R with the reaction velocity:

$$v_0 = \frac{dP}{dt} = k_{cat} C = k_{cat} E_0 (1 - R) = v_{0,max} (1 - R)$$
$$\Rightarrow R = \frac{v_{0,max} - v_0}{v_{0,max}} \tag{4.17}$$

where $v_{0,max} = k_{cat} E_0$. This expression for the determination of R in enzymatic reactions is analogous to **Eq (3.13)** for the determination of R in the binding process.

Eq (4.16) for K_m can be rewritten as a quadratic equation for R and solved to produce a theoretical dependence of R on K_m, S_0, and E_0:

$$R = -\frac{K_m + S_0 - E_0}{2E_0} + \sqrt{\left(\frac{K_m + S_0 - E_0}{2E_0}\right)^2 + \frac{K_m}{E_0}} \tag{4.18}$$

This completes the mathematical formulation for finding $K_{m,det}$ by nonlinear regression of a binding isotherm. The binding isotherm describes the dependence of R determined with **Eq (4.17)** on S_0 for a constant E_0. The determination of $K_{m,det}$ by nonlinear regression of a binding isotherm with **Eq (4.18)** is schematically depicted in **Figure 4.11a**.

4.2.3.2. Identity of Mathematics Describing K_m and K_d Determinations

It is instructive to emphasize that the equilibrium of binding (**Eq (1.1)**) and the pseudo-equilibrium of enzymatic catalysis (**Eq (4.11)**) are distinct processes. Despite this fundamental difference, the mathematics describing the K_m and K_d determinations exhibit strong analogies. A universal nonlinear regression model (**Eqs (2.2)** and **(4.18)**) can be used to fit the experimental binding isotherms; we term this model "the binding-isotherm model". Some previous studies

have recognized the use of the binding isotherm model (or its variants) for K_m determination, particularly when high enzyme concentrations (E_0) are used.[348, 349, 350] However, to our knowledge, no systematic study has comprehensively unified K_m and K_d determinations across theoretical foundations, experimental applications, and data analysis.

We would like to emphasize that the methodological significance of the identical mathematics underlying K_m and K_d determinations cannot be overstated. This mathematical similarity

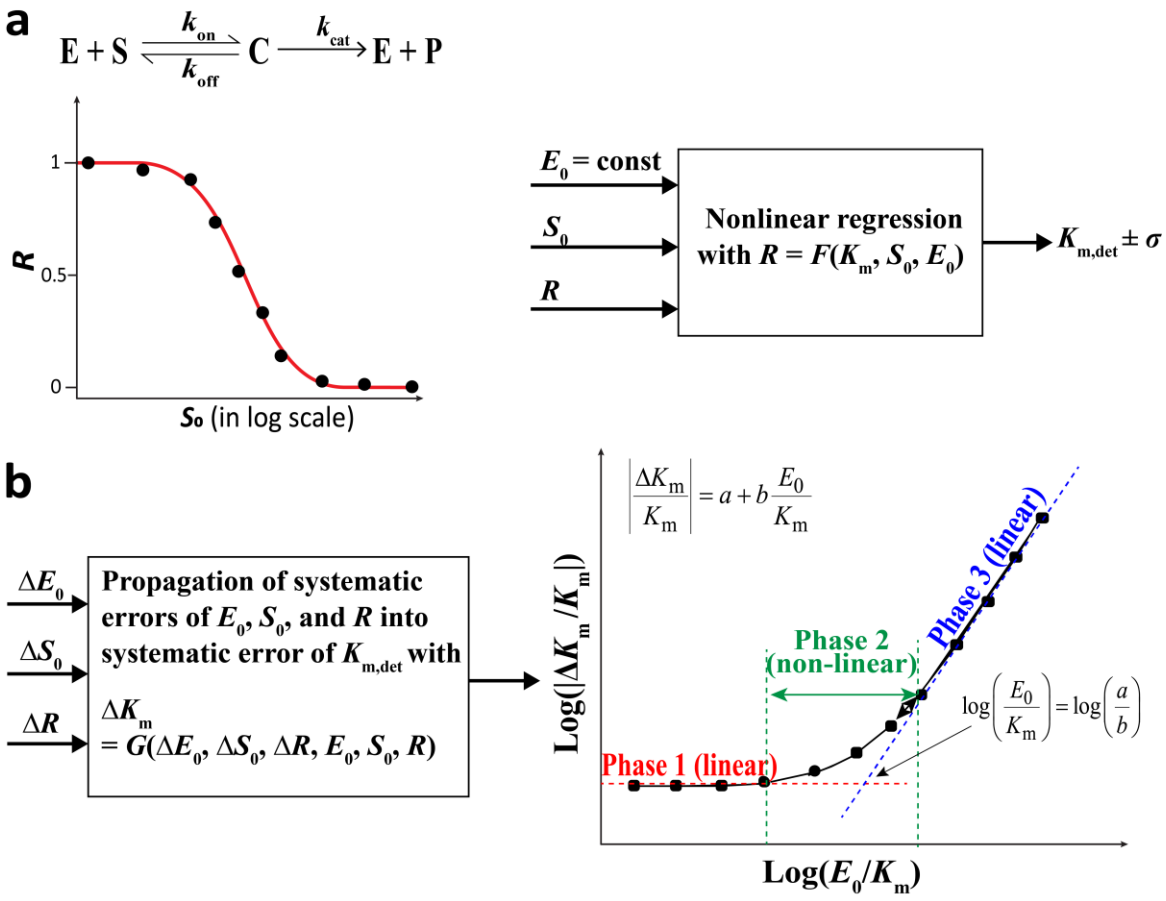


Figure 4.11. Schematic representations of **(a)** determination of $K_{m,det}$ and its standard deviation (σ) by nonlinear regression with three variables from the binding isotherm (E_0, S_0, and R) and **(b)** propagation of systematic errors in these variables leading to a triphasic dependence of the relative systematic error in $K_{m,det}$ on the ratio E_0/K_m in which K_m designates its true unknown value. (This figure is modified from **Figure 3.4**).

provides a unified framework for the quantitative assessment of binding and catalytic processes, streamlining the analytical approach for researchers. By demonstrating that the binding-isotherm model is applicable to both K_m and K_d determinations, we offer a versatile tool that enhances the reliability of nonlinear regression analyses across different biochemical contexts.

This consistency in mathematical treatment simplifies the validation and comparison of experimental results, as the same regression model can be employed without modification. Moreover, it allows for the direct transfer of computational tools and error-propagation techniques between studies focusing on enzymatic reactions and affinity binding, fostering greater efficiency and accuracy in biochemical research.

This unified approach also aids in the education and training of researchers by reducing the learning curve and minimizing the potential for methodological errors. Mastering a single, robust model enhances efficiency and ensures consistency across studies. Ultimately, identical mathematics enables advancements in one area of study to be readily adapted to others, accelerating innovation and improving the overall robustness of biochemical analyses. Here, we capitalize on the ability to directly transfer previously obtained error-propagation results and approaches for accuracy assessment from K_d to K_m.

4.2.3.3. Propagation of Systematic Errors in K_m Determination

The error propagation for K_d has been reported previously in **Section 3.1** of **Chapter 3**.[179] Transferring these findings to K_m involves a straightforward renaming of terms: replacing T_0 with S_0, L_0 with E_0, and K_d with K_m. Below, we detail the results of this transfer.

The accuracy of $K_{m,det}$ is defined by the accuracies of values of E_0, S_0, and R. If these values were exact (*i.e.*, free from systematic errors), K_m determination would be accurate and subject only to random errors. However, in practice, systematic errors in E_0, S_0, and R (denoted as ΔE_0,

ΔS_0, and ΔR) contribute to the systematic error in $K_{m,det}$ (denoted as ΔK_m), as explained below using an assumption that ΔE_0, ΔS_0, and ΔR are known.

The translation of ΔE_0, ΔS_0, and ΔR into ΔK_m follows error-propagation rules. These rules vary depending on whether ΔE_0, ΔS_0, and ΔR are strongly or weakly correlated. If they are strongly correlated (which occurs when similar procedures are used for enzyme and substrate solution preparation, and ΔR results solely from ΔE_0 and ΔS_0), the absolute value of ΔK_m can be approximated as:

$$\left|\Delta K_m\right| = \left|\left(\frac{\partial K_m}{\partial S_0}\right)\Delta S_0\right| + \left|\left(\frac{\partial K_m}{\partial E_0}\right)\Delta E_0\right| + \left|\left(\frac{\partial K_m}{\partial R}\right)\Delta R\right| \tag{4.19}$$

Applying this relationship to the dependence of K_m on E_0, S_0, and R (**Eq (4.16)**) gives:

$$\left|\frac{\Delta K_m}{K_m}\right| = a + b\frac{E_0}{K_m} \tag{4.20}$$

where a and b are error-dependent coefficients: a is determined by the relative systematic errors in substrate concentration and response signal, *i.e.*, $|\Delta S_0/S_0|$ and $|\Delta R/R|$, while b also incorporates the relative systematic error in enzyme concentration, $|\Delta E_0/E_0|$, making it a function of all three error terms.

If ΔE_0, ΔS_0, and ΔR are weakly correlated (*e.g.*, when enzyme and substrate solutions are prepared independently, and ΔR is independent of ΔE_0 and ΔS_0), the absolute value of ΔK_m can be approximated as:

$$\left|\Delta K_m\right| = \sqrt{\left(\frac{\partial K_m}{\partial S_0}\right)^2 {\Delta S_0}^2 + \left(\frac{\partial K_m}{\partial E_0}\right)^2 {\Delta E_0}^2 + \left(\frac{\partial K_m}{\partial R}\right)^2 \Delta R^2} \tag{4.21}$$

Applying this relationship to **Eq (4.16)** gives:

$$\left|\frac{\Delta K_m}{K_m}\right| = \sqrt{\alpha^2 + \lambda\frac{E_0}{K_m} + \beta^2\left(\frac{E_0}{K_m}\right)^2} \tag{4.22}$$

where α, λ, and β are error-dependent coefficients. Both α and λ depend solely on $|\Delta S_0/S_0|$ and $|\Delta R/R|$, while β depends on all three relative errors: $|\Delta S_0/S_0|$, $|\Delta E_0/E_0|$ and $|\Delta R/R|$. **Eqs (4.20)** and **(4.22)** were obtained with a single simplifying assumption: the value of R determined from measured reaction velocities is equal to 0.5, leading to the least erroneous $K_{m,det}$.[10, 101] Thus, **Eqs (4.20)** and **(4.22)** represent the lower limit for relative systematic errors in $K_{m,det}$. Exact dependencies of a and b as well as α, λ, and β on $|\Delta S_0/S_0|$, $|\Delta E_0/E_0|$, and $|\Delta R/R|$, which can be found elsewhere (and in Note C2.1),[179] are not required for our analysis here.

Eqs (4.20) and **(4.22)** share a triphasic shape for the dependence of $|\Delta K_m/K_m|$ on E_0/K_m, as illustrated in double-log coordinates (**Figure 4.11b**). In the first phase, low E_0/K_m values result in a linear relationship with a negligible dependence of $|\Delta K_m/K_m|$ on E_0/K_m as the first term in **Eq (4.20)** and the first two terms in **Eq (4.22)** dominate. In the third phase, high E_0/K_m values lead to another linear relationship: $|\Delta K_m/K_m| = b(E_0/K_m)$ and $|\Delta K_m/K_m| = \beta(E_0/K_m)$, respectively; showing high sensitivity to E_0/K_m. Between these linear phases lies a nonlinear transition range.

Both **Eqs (4.20)** and **(4.22)** indicate that minimizing $|\Delta K_m/K_m|$ involves reducing enzyme concentration E_0 (to lower the E_0/K_m ratio) and minimizing relative systematic errors in E_0, S_0, and R. However, decreasing E_0 is constrained by the instrument's limit of quantitation (LOQ) for the product of the enzymatic reaction.[351] Below this limit, excessively long times are required to accumulate sufficient product for reliable v_0 measurement, crucial for accurate calculation of R with **Eq (4.17)**. Similarly, systematic errors in concentrations cannot be reduced below the random error in stock-solution concentrations. Consequently, achieving highly accurate K_m determinations is inherently challenging, underscoring the need for quantitative methods to

assess the accuracy of $K_{m,det}$. The error-assessment approach developed for K_d can be directly applied to K_m, as utilized below.

4.2.3.4. Quantitative Assessment of K_m Accuracy by Computing its ACI

In general, accuracy is a quantitative term associated with a systematic error, which is defined as the deviation of the determined value from the true value. When the true value is fundamentally unknown, the systematic error cannot be calculated from this definition. As a result, accuracy assessment becomes a complex problem, even for parameters measurable against reference standards like length, mass, time, charge, and temperature.[26, 27, 28, 29] This challenge is even greater for parameters without reference standards, such as K_d and K_m.[3, 15, 292, 293, 294]

We recently suggested an approach to overcome this issue by introducing an Accuracy Confidence Interval (ACI) for parameters determined using correct regression models. In this concept, the correct regression model replaces the true value of the parameter as a reference standard.[10] The ACI defines a range within which the true value lies with a specified probability. For certain regression models, the bounds of this range can be computed by combining error-propagation and regression-stability analyses, provided that CIs for systematic errors in variables can be experimentally estimated.[10]

4.2.3.5. Error Propagation and $K_{m,det}$ Accuracy

The error-propagation analysis begins by transforming random errors in variables (*e.g.*, concentrations) into CIs for their systematic errors. Error-propagation rules are then applied to identify the systematic-error boundaries (of variables) defining the ACI of the sought parameter. The exact boundary conditions of the systematic errors may fall anywhere within the CIs of the

variables, depending on the error-propagation expression derived for specific cases (*e.g.*, $R = 0.5$).[10]

The conversion of random errors into systematic-error CIs requires valid assumptions based on experimental protocols.[10] Specifically, this conversion relies on the assumption that all "true" systematic errors in experimental variables — focusing here on concentrations — have been effectively eliminated through rigorous experimental procedures.[182] Given the similarity between K_d and K_m determinations, the approaches outlined in our recent K_d-focused study can be applied to minimize $|\Delta S_0/S_0|$ and $|\Delta E_0/E_0|$.[182] When these errors are minimized, we can reasonably assume that the CIs of $|\Delta S_0/S_0|$ and $|\Delta E_0/E_0|$ are determined by the relative random errors in stock-solution concentrations (*i.e.*, $\delta S_0/S_0$ and $\delta E_0/E_0$).[10, 182]

This assumption holds true when a single set of enzyme and substrate stock solutions is used to prepare all reaction mixtures in the entire K_m-determination experiment, which is a common experimental approach.[352, 353] The values of $\delta S_0/S_0$ and $\delta E_0/E_0$ can be experimentally determined based on the spectroscopic signals (*e.g.*, light absorption) from substrate and enzyme across multiple stock solution preparations. Alternatively, they can be estimated based on the accuracy uncertainties of the reagent purity and the equipment used for sample preparation.

Since no established method quantitatively assesses the CIs of $\Delta R/R$, we only consider the CIs of $\Delta S_0/S_0$ and $\Delta E_0/E_0$ to compute the narrower limit of the ACI of K_m. By applying error propagation rules (**Eqs (4.20)** and **(4.22)**) to a specific case of **Eq (4.16)** (*e.g.*, $R = 0.5$), we derive the values of $\Delta S_0/S_0$ and $\Delta E_0/E_0$ (in their corresponding CIs) that define the CIs of $\Delta K_m/K_m$ with specified probabilities (Note C2.2).

4.2.3.6. Regression-Stability Analysis

The regression-stability analysis quantitatively evaluates the stability of the $K_{m,det}$ value to changes in E_0 used in nonlinear regression. This process first computes the CI of E_0/K_m and subsequently determines the ACI of K_m.

Traditionally, researchers qualitatively assess $K_{m,det}$ accuracy experimentally by examining whether $K_{m,det}$ is insensitive to changes in E_0 in multiple MM plots.[354] However, this approach is

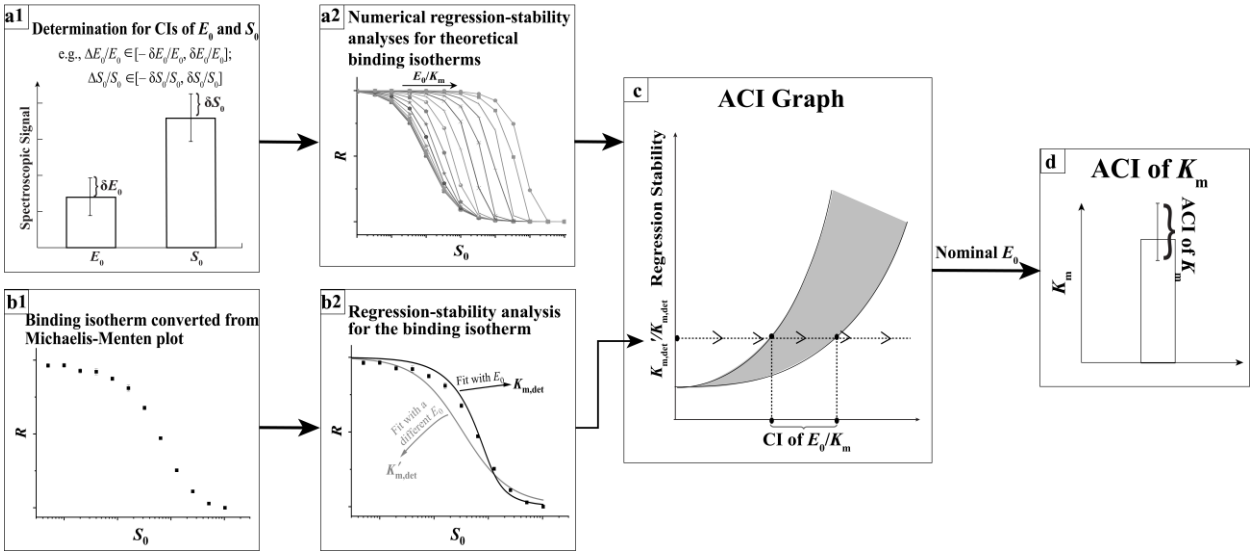


Figure 4.12. Conceptual workflow for computing the accuracy confidence interval (ACI) of K_m: **(a1)** Determine the confidence intervals (CIs) of relative systematic errors for enzyme and substrate concentrations by measuring the spectroscopic signals of enzyme and substrate stock solutions from multiple preparations. **(a2)** Perform a numerical analysis of the regression stability of theoretical binding isotherms, incorporating the determined CIs of concentrations, over a wide range of E_0/K_m ratios. **(b1)** Construct a binding isotherm ("R vs S_0") based on the experimental MM plot ("v_0 vs S_0"). **(b2)** Analyze the regression stability of the binding isotherm developed in **Step b1**. **(c)** Create an ACI graph of "regression stability vs E_0/K_m" using the results from **Step a2**, then apply the regression stability obtained in **Step b2** to the graph to determine the CI of the E_0/K_m ratio. **(d)** Use the CI of the E_0/K_m ratio and the nominal E_0 value to calculate the accuracy confidence interval (ACI) of K_m.

time-consuming, resource-intensive, and lacks quantitative rigor. Inspired by the mathematical similarity between K_d and K_m determinations, we developed a quantitative regression-stability analysis that reveals the ACI of K_m from a single binding isotherm ("R vs S_0"), which can be fitted by the binding isotherm model (**Eq (4.18)**).

We name this approach ACI-Km; the step-by-step conceptual workflow for ACI-Km is shown in **Figure 4.12**. ACI-Km requires a full MM plot ("v_0 vs S_0") obtained under pseudo-steady-state conditions. In this case, $v_{0,max}$ is experimentally achieved at the highest S_0. Using $v_{0,max}$ and **Eq (4.17)**, R values for each S_0 are calculated to construct the binding isotherm ("R vs S_0").

Since the effectiveness of ACI-Km depends on the regression-stability test, which requires sufficient sensitivity of $K_{m,det}$ to variations in E_0 during fitting, this approach is most reliable when the lower bound of the measured CI for E_0/K_m exceeds 0.001. When the true E_0/K_m falls below this threshold — a condition that is experimentally common in many K_m-determination experiments — the contribution of E_0 to K_m determination becomes negligible.[340] In such cases, the ACI of K_m is instead defined by the CI of S_0. This conclusion will be further validated in the program verification section.

4.2.3.7. ACI Graphs and Practical Implications

In ACI-Km, we numerically analyze the stability of $K_{m,det}$ against variations in E_0 used for nonlinear regression, based on theoretical binding isotherms generated by **Eq (4.18)** over a wide range of E_0/K_m. The resulting graph, which depicts the dependence of regression stability on E_0/K_m, is referred to as an ACI graph (**Figure 4.12c**). Importantly, when incorporating experimentally determined CIs for the relative systematic errors in concentrations, the relationship between regression stability and E_0/K_m is represented as an area rather than a single line, as shown by the grey region in the ACI graph (**Figure 4.12c**). For a binding isotherm

derived from an experimental MM plot ("v_0 vs S_0"), its regression stability can be calculated for the same E_0 variations used to construct the ACI graph. By plotting the experimental regression stability onto the ACI graph, we can determine the CI for E_0/K_m, which is then used to compute the ACI of K_m (**Figure 4.12c** and **4.12d**).

4.2.3.8. Advancement of Graphical Approach

Although several well-established Monte Carlo-based strategies exist for numerical error propagation in nonlinear regression,[140, 141, 142] we adopt the "graphical" approach here to intuitively demonstrate how systematic errors in variables can be amplified by a factor of E_0/K_m when propagated into the systematic error of $K_{m,det}$. This visualization provides researchers with a practical framework to estimate how much lower E_0 should be to achieve the desired accuracy in $K_{m,det}$. If the lowest feasible E_0 has already been utilized, further improvements in $K_{m,det}$ accuracy may necessitate the use of an instrument with a lower limit of quantification (LOQ) for the product. Notably, the similarity between ACIs obtained using this graphical approach and those from the Monte Carlo method has been confirmed in our previous work on ACI-Kd.[10]

4.2.3.9. ACI-Km Web Application

To ensure universal accessibility for calculating the ACI of K_m, we have integrated this functionality into the ACI web app (https://aci.sci.yorku.ca). It is important to emphasize that the primary requirement for using the web app to compute the ACI of K_m is that $v_{0,max}$ must be experimentally achieved with an excessively high S_0. By inputting the experimentally obtained MM plot "v_0 vs S_0" into the ACI-Km web app interface, the curve is automatically converted into a binding isotherm "R vs S_0" under the assumption that v_0 at the highest S_0 corresponds to $v_{0,max}$ (*e.g.*, conversion from **Figure C2.1C** to **Figure C2.1D**). The web app then generates a PDF report that provides the ACI and PCI of K_m for three confidence levels: 68.3%, 95.5%, and

99.7% (*e.g.*, **Figure C2.2**). In this study, we primarily focus on the ACI and PCI at 68.3% confidence for demonstration purposes.

4.2.3.10. Verification and Demonstration of ACI-Km

In this study, we verify the ACI-Km workflow by demonstrating its ability to produce results that align with theoretical expectations. Since the true values of key parameters, such as K_m, E_0, S_0, and their associated errors, are fundamentally unknown and uncontrollable in real experimental scenarios, the data of real experiments cannot be effectively used for the verification of ACI-Km. Therefore, we utilized computer-simulated data with well-defined, controllable variables and parameters. This approach enabled us to validate the workflow's accuracy and showcase the practical application of ACI-Km. The details of the simulations and data processing are provided in Note C2.3 (including **Figures C2.1** and **C2.2**).

Briefly, we simulated the kinetics of the enzymatic reaction using the following system of ordinary differential equations that describes the reaction mechanism outlined in **Eq (4.11)**:

$$\begin{aligned} \frac{dS}{dt} &= -k_{on}ES + k_{off}C \\ \frac{dC}{dt} &= k_{on}ES - (k_{off} + k_{cat})C \\ \frac{dP}{dt} &= k_{cat}C \end{aligned} \tag{4.23}$$

In the simulations, we defined a set of rate constants k_{on}, k_{off}, and k_{cat}, resulting in $K_m = 1$ (unitless). Reaction progress curves ("P vs time") were generated for a wide range of nominal E_0 values (0.005, 0.01, 0.05, 0.1, and 0.5 (unitless)) and a given range of nominal S_0 values (0.01, 0.033, 0.1, 0.33, 1.0, 3.3, 10, 33, 100, 333, and 1000 (unitless)). Given that the validity of the steady-state assumption is widely accepted when $E_0 << S_0 + K_m$,[355] we did not further increase the nominal enzyme concentration in this study. Systematic errors were

introduced with two combinations: (*i*) $\Delta E_0/E_0 = 0.2$ and $\Delta S_0/S_0 = -0.2$ and (*ii*) $\Delta E_0/E_0 = -0.2$ and $\Delta S_0/S_0 = 0.2$. A normally distributed relative random error of 1% was applied to the product concentrations in the synthetic progress curves. Representative synthetic reaction progress curves are shown in **Figure C2.1a**. The pseudo-steady-state phases of the simulated progress curves were analyzed to determine the v_0 for different nominal S_0 at each nominal E_0 (*e.g.*, **Figure C2.1B**). From these data, the MM plots "v_0 vs S_0" were constructed for each E_0 and a combination of systematic errors in E_0 and S_0 (*e.g.*, **Figure C2.1C**).

The ACI-Km web app was used to compute the PCI and ACI of K_m for each condition. As explained in our previous work,[10] to ensure that the ACI (with 68.3% confidence level) includes the true K_m, the minimum relative random errors for E_0 and S_0 should be $\delta E_0/E_0 = |\Delta E_0/E_0|/1.36$ and $\delta S_0/S_0 = |\Delta S_0/S_0|/1.36$, which leads the input $\delta E_0/E_0 = \delta S_0/S_0 = 0.2/1.36 = 0.15$. To validate the ACI-Km workflow, we entered the nominal E_0

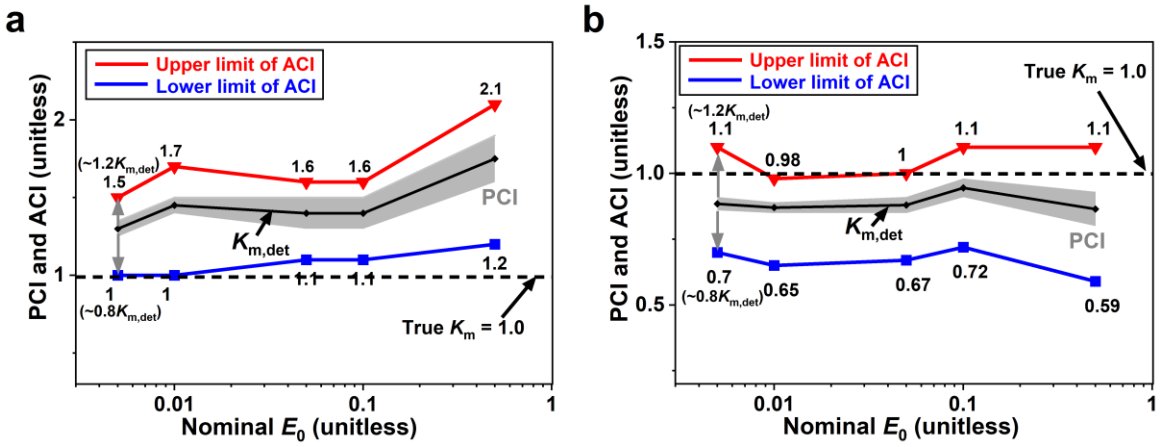


Figure 4.13. PCI and ACI determined using ACI-Km for two sets of MM plots derived from the simulated kinetic data under different conditions: **(a)** $\Delta E_0/E_0 = 0.2$ and $\Delta S_0/S_0 = -0.2$, with nominal E_0 values of 0.005, 0.01, 0.05, 0.1, and 0.5 (unitless) and **(b)** $\Delta E_0/E_0 = -0.2$ and $\Delta S_0/S_0 = 0.2$, with nominal E_0 values of 0.01, 0.05, 0.1, and 0.5 (unitless). For both cases, $\delta E_0/E_0 = \delta S_0/S_0 = 0.15$ was input into the ACI web app (https://aci.sci.yorku.ca) to calculate both the ACI and PCI of K_m.

values, the corresponding relative random errors, and the MM plots data into the web app (https://aci.sci.yorku.ca). If both the web app and ACI-Km method perform as expected and there are no systematic errors in R values, the true K_m should fall at or near one of the boundaries of the computed ACI for K_m (with 68.3% confidence), depending on the specific combination of errors introduced in the MM plots obtained from the simulated data.

Figure 4.13 summarizes the comparison between the true/input K_m and the ACI and PCI of K_m, calculated from simulated data using ACI-Km. The results presented in **Figure 4.13** align well with our four theoretical predictions: (*i*) when $E_0 << S_0 + K_m$ (*e.g.*, $E_0/(S_0 + K_m) < 0.1$) is satisfied for all S_0 used in the simulation (*i.e.*, $E_0 \leq 0.1$), the true K_m lies near or at the boundary of the computed ACI for K_m; (*ii*) when $E_0 > 0.1$, the true K_m shifts away from the boundary of ACI, due to the difficulty in accurately measuring the initial reaction velocity without satisfying the steady-state assumption, introducing unquantifiable systematic errors in the calculated R values; (*iii*) as E_0 increases, the ACI tends to widen due to the magnification of systematic errors, as predicted by **Eqs (4.20)** and **(4.22)**; and (*iv*) ACI of K_m are often much wider than PCI of K_m.

The results in **Figure 4.13** suggest that when the condition $E_0 << S_0 + K_m$ is satisfied for all experimental S_0, and CIs for systematic errors in concentrations are accurately estimated, the ACI-Km method can compute the ACI that includes the true value of K_m. Even if this condition is not strictly met for some data points, calculating the ACI remains valuable. A wide ACI at the experimental E_0 can serve as an indicator for experimentalists to refine their experimental conditions, such as reducing the enzyme concentration to improve accuracy. Notably, the PCI is often significantly narrower than the ACI, meaning that relying solely on the PCI can result in a substantial underestimation of the uncertainty in K_m, potentially causing inaccurate under- or

overestimation of K_m. This underscores the importance of incorporating the ACI in K_m characterization to provide a more accurate representation of the true uncertainty.

Further analysis reveals that when the ratio of nominal E_0 to the true K_m was ≤ 0.01, the ACI of K_m closely matched the range from $0.8K_{m,det}$ to $1.2K_{m,det}$ (**Figure 4.13**), which could be determined solely by setting $\delta S_0/S_0 = 0.15$ while neglecting errors in E_0 (*i.e.*, setting $\delta E_0/E_0 = 0$) (Note C2.4). This finding suggests that for $E_0/K_m \leq 0.01$, the ACI of K_m is primarily dictated by the relative systematic error in the substrate concentration ($|\Delta S_0/S_0|$). For instance, at a 68.3% confidence level, the ACI is given by $[K_{m,det}(1–1.36\delta S_0/S_0), K_{m,det}(1+1.36\delta S_0/S_0)]$ (Note C2.4 and **Figure C2.3**). This effect arises because, at very low E_0/K_m ratios, the contribution of the relative systematic error in E_0 ($|\Delta E_0/E_0|$) to the overall systematic error in $K_{m,det}$ becomes negligible (refer to **Eq (4.20)** or **(4.22)**). Furthermore, this observation suggests that while the direct calculation of ACI for K_m from the ACI graph is not feasible when the lower bound of the CI for E_0/K_m falls below 0.001 due to the limited sensitivity of $K_{m,det}$ to the E_0 values used in fitting, the ACI of K_m can still be estimated. Specifically, $K_{m,det}$ can be determined by fitting the binding isotherm "R vs S_0" (derived from the experimental MM plot) using **Eq (4.18)**, after which the ACI of K_m can be computed as $[K_{m,det}(1-N\delta S_0/S_0), K_{m,det}(1+N\delta S_0/S_0)]$, where N takes values of 1.36, 2.27, and 3.18 for confidence levels of 68.3%, 95.5%, and 99.7%, respectively (confirmed in **Figure C2.3**).[10]

Based on these findings, we have improved the ACI-Km web app with the following updates. When the lower bound of the CI for $E_0/K_m \geq 0.01$, the ACI of K_m is calculated using the ACI-Km workflow as shown in **Figure 4.12.** When the lower bound of the CI for $E_0/K_m < 0.01$, the ACI-Km web app first determines $K_{m,det}$ by fitting the converted binding isotherm using **Eq (4.18)**. The ACI of K_m is then computed using the equation above. Regardless of E_0/K_m, the PCI can always be calculated based on the fitted $K_{m,det}$ and its standard deviation (σ).

4.2.3.11. Application of ACI-Km to Experimental Data

After verifying and updating the ACI-Km web app, we utilized it to evaluate the ACI and PCI of experimentally determined K_m values. To demonstrate its applicability in K_m determination, we first applied ACI-Km to a published dataset for the enzyme–substrate pair AnaAT–acetyl-coenzyme A. Additionally, we performed an experiment using lactate dehydrogenase (LDH) with pyruvate as the substrate and NADH as a co-substrate. The ACI-Km program was then employed to analyze our experimental data, further illustrating its effectiveness in estimating reliable $K_{m,det}$.

To assess the effectiveness of ACI-Km in K_m determination, we first applied it to a published dataset for the enzyme–substrate pair AnaAT–acetyl-coenzyme A from Yin *et al*. published in 2008.[356] This dataset represents a case where conventional Michaelis-Menten and Lineweaver-Burk models were unsuitable, whereas the ACI-Km program based on the binding isotherm model (**Eq (4.18)**) can provide a more accurate framework for analysis.

In the published experiment, the enzyme concentration $[AnaAT]_0$ (E_0) was 0.01 μg/μl. The molecular weight of AnaAT was determined to be 43 kDa using size-exclusion filtration, corresponding to a nominal enzyme concentration of approximately 0.23 μM. Substrate concentrations $[acetyl\text{-}coenzyme\ A]_0$ were varied at 0, 5, 10, 20, 50, 200, 500, and 1000 μM. By analyzing the product formation rate under each condition, experimental progress curves "P vs time" were generated, from which the MM plot "v_0 vs S_0" was established (**Figure 4.14a**). The MM plot was then transformed into a Lineweaver-Burk plot "$1/v_0$ vs $1/S_0$" (**Figure 4.14b**) to find $K_{m,det}$, yielding a value of 96 μM without uncertainty being reported.

The Lineweaver-Burk plot is a linearized transformation of the MM plot, which assumes $E_0 << K_m$.[354] If the E_0/K_m ratio is sufficiently small, the contribution of E_0 to K_m determination is

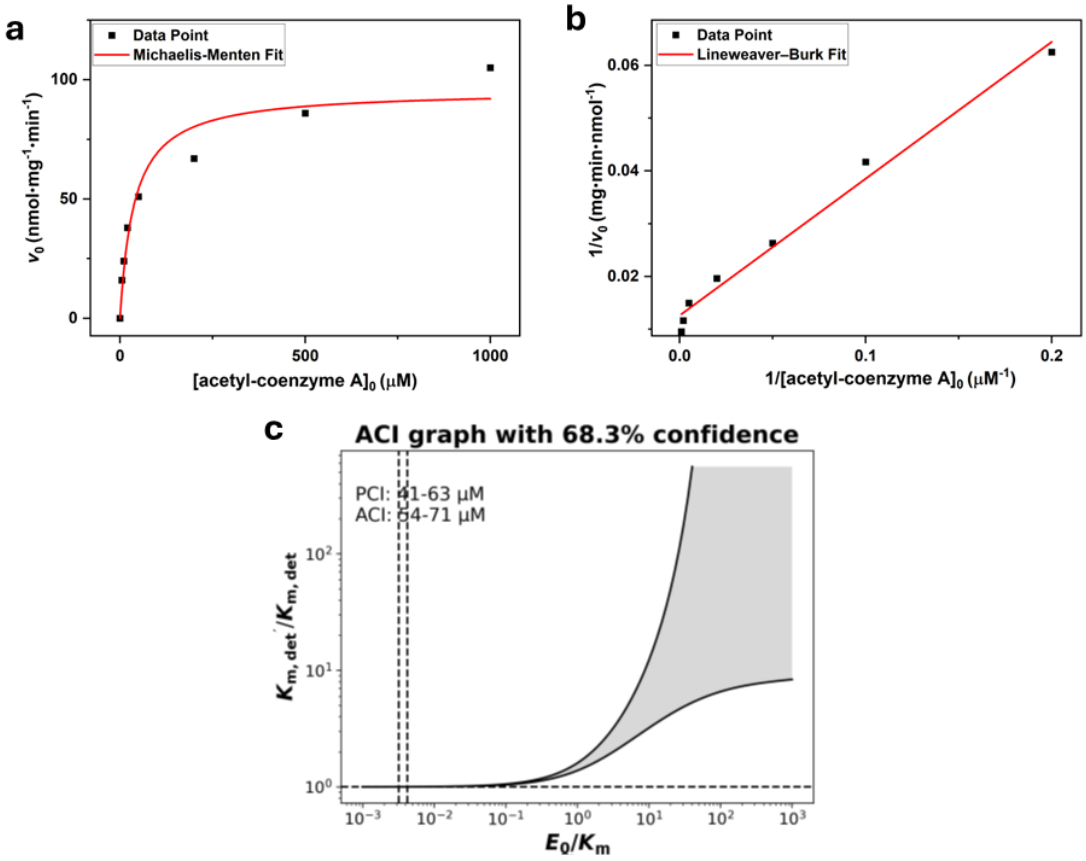


Figure 4.14. Application of ACI-Km to a retrospective K_m-determination dataset for the AnaAT–acetyl-coenzyme A pair. **(a)** Experimental MM plot "v_0 vs [pyruvate]$_0$" fitted using the MM model. **(b)** Lineweaver–Burk transformation of the pseudo-steady-state curve "1/v_0 vs 1/[pyruvate]$_0$" fitted with the linear Lineweaver–Burk model. **(c)** ACI graph (68.3% confidence) generated by ACI-Km, from which the PCI of K_m was determined as 41–63 μM, and the ACI of K_m as 54–71 μM. (Panels **a** and **b** are adapted from **Figure 6D** in **ref 356**)

considered negligible, resulting in an approximately linear Lineweaver-Burk plot. However, as shown in **Figure 4.14b**, the plot deviated significantly from linearity, suggesting that this assumption was not valid. This observation highlights the risks of simply assuming E_0/K_m is small enough to disregard E_0 in K_m determination without verification.

To address this limitation, we applied ACI-Km to calculate the ACI and PCI of K_m for AnaAT–acetyl-coenzyme A using the same published dataset. Unlike the conventional MM or Lineweaver-Burk approach, ACI-Km employs the binding isotherm model (**Eq (4.18)**), which is derived from first principles without excluding E_0 from the fitting equation. As shown in **Figure 4.14a**, the experimental MM plot approaches the plateau region at the highest substrate concentration ([acetyl-coenzyme A]$_0$=1000 μM), indicating that the measured v_0 at this condition was nearing $v_{0,max}$. This saturation suggests that the necessary conditions for applying the binding isotherm model were satisfied, supporting the validity of using ACI-Km for K_m determination.

To utilize the ACI-Km program, we needed to account for the uncertainty of inaccuracy in reagent concentrations. Since the original study did not report the systematic uncertainties associated with enzyme and substrate concentrations, we estimated the lower confidence interval (CI) limits for their relative systematic errors to perform the ACI calculations. Systematic errors in enzyme concentration can arise from factors such as pipetting inaccuracies, dilution errors, sample impurities, and adsorption to surfaces, which are particularly significant at low enzyme concentrations.[357] Based on these considerations, the standard deviation ($\delta E_0/E_0$) for the CI of $|\Delta E_0/E_0|$ is unlikely to be below 10%.[358]

Given that the substrate concentration range spans from micromolar to millimolar levels, the relative systematic error in substrate concentration is expected to be lower than that of the enzyme. However, considering the inherent uncertainties in mass and volume measurements during sample preparation, reducing the standard deviation ($\delta S_0/S_0$) for the CI of $|\Delta S_0/S_0|$ below 5% remains challenging.[359] Therefore, in this study, we assigned relative standard deviations of

10% and 5% to the CIs of $|\Delta E_0/E_0|$ and $|\Delta S_0/S_0|$, respectively. These values were also applied in the analysis of our experimental data.

By inputting the published dataset, the nominal enzyme concentration of 0.23 μM, and the proposed relative standard deviations for $|\Delta E_0/E_0|$ and $|\Delta S_0/S_0|$ (10% and 5%, respectively) into the ACI-Km web app, the ACI and PCI for K_m were computed based on the numerically established ACI graph (*e.g.*, **Figure 4.14c**). At a 68.3% confidence level, the ACI and PCI for K_m were determined to be 54–71 μM and 41–63 μM, respectively. Both intervals have similar widths and show substantial overlap, yet the ACI provides a more statistical significance. The ACI estimate indicates that the accurate K_m value for AnaAT–acetyl-coenzyme A falls within the range of 54–71 μM with 68.3% confidence, which is substantially lower than the originally reported 96 μM, which was determined using the linearized MM model, specifically the Lineweaver-Burk plot.

In our previous study, we numerically demonstrated that employing a simplified fitting model, such as the K_d fitting model (mathematically equivalent to the MM equation) incorrectly assuming the ligand concentration to be negligible in K_d determination, can result in an overestimation of the K_d value.[182] The present experimental results for K_m determination further support this concept, confirming that a similar bias arises when using a simplified model, specifically the MM equation, for K_m determination. This finding highlights the importance of employing the unified binding isotherm model for both K_d and K_m determinations, facilitating cross-applicability between these areas of study.

To further explore the application of the ACI concept in K_m determination, we conducted an experiment using lactate dehydrogenase (LDH) and pyruvate as the enzyme–substrate pair, with NADH as a co-substrate. In this experiment, the conversion of NADH to NAD^+ was monitored

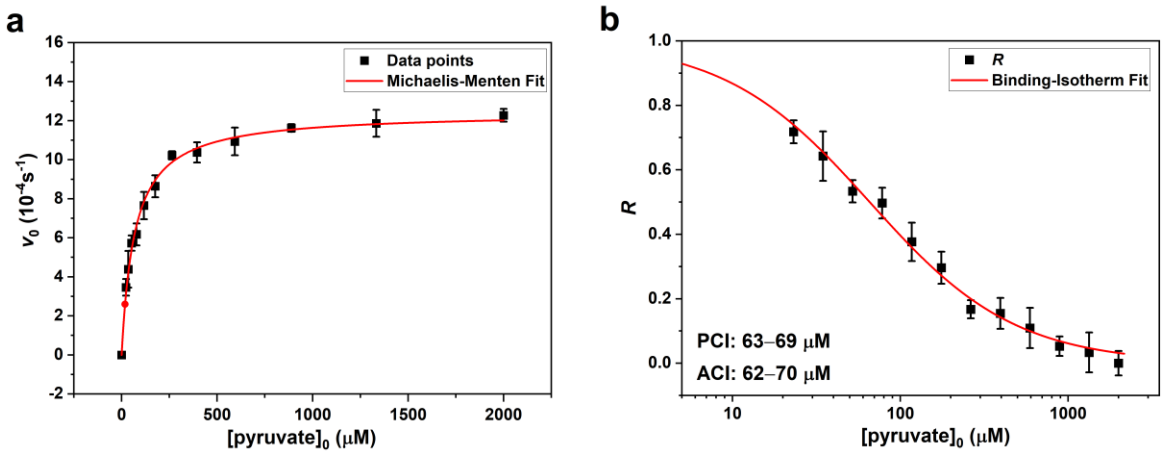


Figure 4.15. Application of the binding-isotherm model and ACI-Km to a K_m-determination experiment using LDH–pyruvate as the enzyme–substrate pair. **(a)** Experimental MM plot "v_0 vs [pyruvate]$_0$" with an MM fit. **(b)** Transformation of the MM plot into the binding-isotherm plots, fitted using the binding-isotherm model (**Eq (4.18)**) to compute the PCI and ACI of K_m under conditions where the E_0/K_m ratio is lower than 0.001.

by measuring the decrease in absorbance at 340 nm, reflecting the rate of product formation. A nominal enzyme concentration of $[LDH]_0 = 0.87$ nM was used. The progress curves "P vs time" were recorded at initial pyruvate concentrations ($[pyruvate]_0$), from 23 to 2000 μM in 1.5-fold increments. From the recorded progress curves, we determined v_0 for product formation at each pyruvate concentration, which resulted in the MM plot "v_0 vs $[pyruvate]_0$" (**Figure 4.15a**). Additional experimental details can be found in the "Materials and Methods" section (**Section 4.2.2**) and Note C2.5, which includes **Figure C2.4**. For comparison purposes, the fitting line to the MM plot using the MM equation is shown in **Figure 4.15a**, yielding $K_{m,det} = 68 \pm 4$ μM (PCI: 64–72 μM) and $v_{0,max} = (12.4 \pm 0.2) \times 10^{-4}$ s^{-1}. The agreement between the predicted $v_{0,max}$ and the measured v_0 ($(12.3 \pm 0.3) \times 10^{-4}$ s^{-1}) at the highest initial pyruvate concentration ($[pyruvate]_0 = 2000$ μM) confirms the applicability of ACI–K_m to this experimental MM plot.

The MM plot was analyzed using the ACI-Km program to compute the ACI and PCI of K_m for LDH–pyruvate. To determine the lower limit of the ACI for K_m, we assigned standard deviations of 10% and 5% to the CIs of $|\Delta E_0/E_0|$ and $|\Delta S_0/S_0|$, respectively. Although the lower bound of the CI for E_0/K_m fell below 0.001, the updated ACI-Km web app successfully estimated the ACI of K_m as 62–70 μM, which was slightly wider than the PCI of 63–69 μM, both at 68.3% confidence (**Figure 4.15b**). Notably, at such an extremely low E_0/K_m ratio, when $v_{0,max}$ is experimentally measured, the PCI ranges determined by both the MM model and the binding-isotherm model closely agree, as theoretically expected. It is important to emphasize that this ACI represents its lower bound, assuming no significant systematic errors in substrate concentration due to improper experimental practices. In this context, any residual systematic errors are presumed to stem from random errors in sample preparation.

4.2.4. Concluding Remarks

In enzymology, nonlinear regression is commonly applied to Michaelis–Menten plots to estimate $K_{m,det}$. Although the $K_{m,det}$ values determined for the same enzyme–substrate pair using the same method are often precise (having narrow PCI), they can still vary significantly, suggesting inaccuracies and potential misconceptions.[342, 343]

Traditionally, enzymologists strive to use E_0 values much smaller than S_0 to satisfy the pseudo-steady-state condition $E_0 << S_0 + K_m$.[342] However, in some cases, due to the properties of the enzyme and the limitations of detection techniques, using relatively high enzyme concentrations becomes unavoidable. [348] In such scenarios, the pseudo-steady-state assumption can be considered valid and the determined $K_{m,det}$ is deemed accurate only if it remains consistent across different E_0 values. [354] However, this traditional "regression-stability analysis", which involves experimentally varying E_0, is time-consuming, resource-intensive, and non-

quantitative.

In this study, we found the mathematical similarities between K_m and K_d determinations and concluded that all methods developed to maximize and quantitatively assess the accuracy of K_d can be directly applied to K_m-determination experiments and data analysis. Recently, we proposed an approach to compute the ACI of K_d determined by nonlinear regression with a correct regression model.[10] Here, we reiterated the conceptual workflow of ACI (**Figure 4.12**) and verified its capability of computing the ACI of K_m (**Figure 4.13**).

Unlike K_d determination, where the experimentally used ligand concentration is sometimes significantly higher than the true K_d due to the instrument's high LOQ values,[10, 179] achieving $E_0/K_m < 1$ is generally more feasible, as commonly studied K_m values (within the micromolar range) are usually much larger than typical K_d values (in the picomolar to nanomolar range). Nevertheless, calculating the ACI of K_m remains crucial, as our previous study indicated that obtaining an accurate K_m depends on satisfying the condition $E_0/K_m < a/b$ (**Figure 4.11b**) rather than $E_0/K_m < 1$. Since the coefficients a and b are defined by the relative systematic errors of the variables, the ratio a/b can be significantly less than unity if the variables are subject to large systematic errors, which is possible in K_m determination.[179, 360, 361] Consequently, a wide ACI may still be obtained even when $E_0/K_m < 1$. In such cases, experimentalists may need to further reduce E_0 to achieve a narrower ACI for K_m that aligns with the accuracy required for their specific enzymatic application.

To promote the rapid and seamless adoption of our ACI calculation approach within the enzymology research community, we have incorporated the option to determine the ACI for K_m in the ACI web app (https://aci.sci.yorku.ca) and expanded its applicability across the entire range of E_0/K_m. In this study, we validated the ACI-Km web app using synthetic data and

demonstrated its practical utility through real K_m-determination experiments.

Our retrospective analysis of published experimental data revealed that even when $E_0 << S_0 + K_m$ (and $E_0/K_m < 1$), conventional data analysis using the MM or Lineweaver–Burk models can result in an overestimation of K_m. By incorporating the ACI and PCI of K_m, as determined using ACI-Km with the binding isotherm model (**Eq (4.18)**), the accuracy and reliability of K_m determination can be assessed and subsequently improved.

Incorporating ACI calculations for K_m alongside PCI assessment can significantly enhance the reliability of enzyme kinetics studies. This combined approach mitigates the risk of misinterpretations arising from inaccuracies in K_m determination and provides a more comprehensive understanding of enzyme–substrate intermediate stability under pseudo-steady-state conditions.

Chapter 5. Application of ACI to Isothermal Titration Calorimetry (ITC)

5.1 Deterministic Error Propagation in ITC: Revealing Multi-Fold Errors in K_d Values Under Standard Conditions

The following material is reprinted from the article "**Wang, T. Y.**; Krylov, S. N. Deterministic Error Propagation in ITC: Revealing Multi-Fold Errors in K_d Values Under Standard Conditions. *Biophysical Chemistry* **2025**, *323*, 107455." Copyright 2025 Elsevier. My contributions to the article were: (*i*) conducting all mathematical analysis and equation derivations, (*ii*) preparing all figures, (*iii*) interpreting the results, and (*iv*) writing the manuscript.

5.1.1. Introduction: Potential Inaccuracies in ITC Measurements

As we discussed in **Chapter 3**, to improve and evaluate the accuracy of $K_{d,det}$, it is essential to first maximize the accuracy of the input variables involved in K_d determination, such as analyte concentrations and experimental signals.[182] This requires minimizing systematic errors and estimating the confidence intervals of any remaining systematic errors in these variables.[10, 182] These intervals should then be propagated into the final $K_{d,det}$ values to account for their cumulative effect.[10] Moreover, as we suggested in **Chapter 4**, given that the probability distribution of $K_{d,det}$ is often non-Gaussian,[10, 295] it is scientifically more rigorous to report $K_{d,det}$ as a confidence range accompanied by a specified confidence level (*e.g.*, 95%).

The most common methods for K_d determination, for which specialized commercial equipment is available, include isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), bio-layer interferometry (BLI), fluorescence anisotropy (FA), and microscale thermophoresis (MST).[12, 182] Each method offers unique advantages, *e.g.*, SPR and BLI can also

provide kinetic data, while MST requires smaller sample volumes.[66, 185, 362] However, many of these methods, such as SPR, BLI, FA, and MST often require labeling or immobilization, which can potentially alter the natural interaction between binding partners.[66, 185, 363, 364, 365] ITC, on the other hand, has gained widespread use in the past few decades due to its ability to directly measure the heat changes associated with binding without the need for labels or immobilization.[53, 184, 185] Additionally, ITC provides a comprehensive thermodynamic profile, including enthalpy ($\Delta H°$), Gibbs free energy ($\Delta G°$), and entropy ($\Delta S°$), making it a preferred method for thoroughly studying molecular interactions in their native, unmodified states.[55, 56] Major limitations of ITC include low throughput and the need for high amounts of materials.[53, 185] This balance of advantages and limitations makes ITC well-suited for confirmatory studies aiming at obtaining accurate and comprehensive binding data for binders identified in high-throughput screening by other methods.[52, 53]

In a classic (non-titration) equilibrium approach to determining K_d, *e.g.*, FA and MST, multiple equilibrium mixtures are prepared, where the concentration of one component (the ligand) is kept constant, while the concentration of the other component (the target) is varied over a wide range.[66, 366] For each mixture, the fraction of the free (unbound to the target) ligand is measured. Finally, the resulting binding isotherm, depicting the dependence of this fraction on the target concentration, is analyzed using nonlinear regression to obtain $K_{d,det}$.[34, 66, 101, 181, 366]

We have discussed the error-propagation analysis for this classic approach in **Section 3.1**.[179] This analysis has shown that the relative error in $K_{d,det}$ becomes increasingly sensitive to systematic errors in the variable — both concentrations and the measured unbound fraction — as the ratio between the true (and, thus, unknown) values of ligand concentration and K_d grows.[179] This ratio is called a *c*-value in ITC; we will use this notation for the non-titration approach

discussed here as well. Even small inaccuracies in the variables can be greatly amplified when propagated into error in K_d.[179] This amplification leads to significant systematic errors in $K_{d,det}$ when the *c*-value is high.

It is generally recommended that the *c*-value be kept below unity for the non-titration approach.[34, 181, 225] However, this recommendation is an oversimplification and may not always be applicable. The error-propagation analysis (**Section 3.1**) revealed that the *c*-value should instead be compared to a threshold determined by the magnitude of errors in the variables — concentrations and the unbound fraction.[179] This threshold can be significantly smaller than unity if these errors are substantial.[179] In other words, when the errors in the variables are high, the relative systematic errors in $K_{d,det}$ can be significant, even when $c < 1$. Experimental observations have shown that differences in $K_{d,det}$ for varying *c*-values can span several orders of magnitude despite the relative standard deviation remaining below 50%.[34, 179]

Unlike the classic approach discussed above, ITC utilizes a titration method enabling dynamic variation in target concentration over successive injections. The concentrated solution of the target (titrant) is titrated into the ligand (titrand) solution so that the concentration of the ligand gradually and slightly decreases while the concentration of the target increases over a wide range. The successive injections of the target generate a binding isotherm from the heat release or consumption measured after each injection.

For ITC, it is recommended that the *c*-values for a 1:1 binding process fall within the range of 10 to 100.[186] This range ensures that the ITC binding curve contains sufficient information to reliably determine thermodynamic parameters, such as the enthalpy change ($\Delta H°$), while avoiding overly steep curves that could compromise the accurate determination of K_d at higher *c*-values.[55, 186] Interestingly, this recommendation to maintain $c > 10$ stands in stark contrast to the

consensus recommendation of satisfying $c < 1$ in non-titration methods, suggesting that standard conditions for ITC may potentially compromise the accuracy of $K_{d,det}$.

The nonlinear regression model in ITC differs from the one (**Eq (2.2)**) used in the non-titration approach. Instead of the target concentration, it utilizes the molar (target-to-ligand) ratio and usually includes three unknowns: K_d, $\Delta H°$, and the binding stoichiometry (n).[55, 186] Therefore, the analytical solution obtained for error propagation in the non-titration K_d-determination approach (connecting the relative systematic error in $K_{d,det}$ to systematic errors in concentration and fraction unbound) is not directly applicable to ITC. This represents a significant methodological gap.

To address this limitation, we developed an analytical solution for error propagation in K_d determined with ITC, as presented here. This solution specifically focuses on how systematic errors in key variables, such as the concentrations of the ligand and target, and the measured heat changes, propagate through the nonlinear regression model to affect $K_{d,det}$. Our approach involves deriving equations that explicitly link these systematic errors to the resulting relative systematic error in $K_{d,det}$, taking into account the unique aspects of the ITC model, including its dependence on the molar ratio and additional unknown parameters, such as $\Delta H°$.

We incorporated experimentally realistic systematic errors in reactant concentrations and ITC-measured heat into the derived analytical solutions to evaluate the magnitude of relative errors in $K_{d,det}$ across a wide range of c-values. For the recommended range of c-values (10 to 100), we found that the accuracy of $K_{d,det}$ can be poor, with deviations from the true K_d values reaching multi-fold levels, consistent with certain experimental findings reported in the ITC field.[56, 313] This finding underscores that ITC can be prone to gross inaccuracies in $K_{d,det}$ within its recommended operating conditions. It also highlights the necessity for a quantitative

evaluation of $K_{d,det}$ accuracy in ITC experiments. While this evaluation has been addressed in a number of publications,[176, 367, 368, 369] it remains underutilized by practitioners, largely due to a lack of awareness of the potential magnitude of errors in $K_{d,det}$.[11]

5.1.2. Results and Discussion

5.1.2.1. Mathematics of K_d Determination with ITC

The mathematical foundation of a classic method for determining K_d through nonlinear regression of a binding isotherm has been thoroughly explained in ***Section 3.2.2.3*** of **Chapter 3**. This completes our discussion of the core mathematical principles underlying K_d determination by nonlinear regression in the non-titration approach. In this section, we will examine the similarities between the mathematical frameworks used in ITC-based K_d determination and those employed in the non-titration method.

In ITC, the binding process is monitored by measuring the heat change during each injection of the target into the ligand.[53, 55, 184] For clarity, in this work, we refer to the injected reactant (titrant) — whose concentration varies throughout the experiment — as the target, while the reactant (titrand) in the ITC sample cell, which remains at a near-constant concentration, is referred to as the ligand, aligning with the terminology used in our previous studies.[9, 11, 32]

In ITC experiments, the total heat absorbed/released after i-th injection (Q_i) is directly proportional to the total amount of complex formed:[370, 371]

$$Q_i = nV_0\Delta H°C \tag{5.1}$$

Here, n represents the number of binding sites per ligand molecule (analyte in the ITC sample cell), V_0 is the sample cell volume, $\Delta H°$ denotes the enthalpy change of complex formation (heat absorbed/released per mole of the complex formed), and C is the equilibrium concentration of the complex in the sample cell after the i-th injection. In this study, we focus on the one-site

binding model (**Eq (1.1)**) with n theoretically equal to unity, as this is the most common interaction type in both biological and technological contexts.[180, 241]

By combining **Eqs (3.10)** and **(2.1)**, we obtain:

$$C = (1 - R)L_0 \tag{5.2}$$

Then, for n = 1, **Eq (5.1)** can be rewritten as:

$$Q_i = V_0 \Delta H^\circ L_0 (1 - R) \tag{5.3}$$

By substituting the dependence of R on K_d, L_0, and T_0 (**Eq (4.18)**) into **Eq (2.2)**, the fundamental equation used to derive the practical nonlinear regression models for ITC data is:

$$Q_i = V_0 \Delta H^\circ \left[\left(K_d + L_0 \Gamma + L_0 \right) / 2 - \sqrt{\left(K_d + L_0 \Gamma - L_0 \right)^2 / 4 + K_d L_0} \right] \tag{5.4}$$

with K_d and ΔH° as sought parameters and the molar ratio $\Gamma = T_0/L_0$ used instead of T_0 as the varied independent variable.[371] Note that, according to **Eq (5.4)**, if the ITC measurements reach saturation (no unbound ligand remains) at the highest target concentration used in the experiment ($T_0 >> L_0$, K_d), the R values can be back-calculated from Q_i, and the total heat changes measured at the initial (zero) and saturation target concentrations (Note D1.1).

In most ITC data-fitting processes, the first derivative of **Eq (5.4)** with respect to T_0 serves as the fitting model.[55, 371] Furthermore, due to slight dilution of reactants during the titration process and the overflow design of modern ITC instruments, small corrections to L_0 and T_0 after each injection are introduced when applying nonlinear regression to process the experimental data.[55, 371] Nonetheless, these adjustments do not alter the fact that the practical fitting models used for ITC data analysis fundamentally originates from **Eq (5.4)** and aligns with the principles of the classic non-titration equilibrium K_d-determination approach explained in ***Section 3.2.2.3*** of **Chapter 3**. This equivalence underscores that the insights derived from the classic method are

applicable to ITC. To streamline this study, here we conduct a semi-quantitative analysis leveraging the close relationship between **Eq (5.4)** and **Eq (2.2)**, which is central to interpreting experimental data from the classic K_d-determination approach.

5.1.2.2. Propagation of Systematic Errors for K_d Determination in ITC

Error propagation in K_d determination using the non-titration equilibrium approach has been explored in **Chapters 3** and **4**.[10, 179, 182] In these studies, error propagation rules for strongly and weakly correlated error sources were applied to the dependence of K_d on L_0, T_0, and R (**Eq (3.1)**) to establish the relationships between the accuracy of $K_{d,det}$ ($|\Delta K_d/K_d|$, where $\Delta K_d = K_{d,det} - K_d$) and the c-value, defined as the ratio between the true values of L_0 and K_d (L_0/K_d).[179] As a result, we found that regardless of the relationships between error sources, the dependence of $|\Delta K_d/K_d|$ on the c-value (L_0/K_d) exhibits a triphasic shape, best seen in double-log coordinates (**Figure 4.4b**). In the first phase, at low c-values, $|\Delta K_d/K_d|$ demonstrates a linear relationship with a negligible dependence of $|\Delta K_d/K_d|$ on c. In the third phase, at high c-values, another linear relationship emerges, showing a high sensitivity of $|\Delta K_d/K_d|$ to the c-value. Between these two linear phases lies a nonlinear transition phase. These dependencies were derived for a typical case where the experimentally determined R is 0.5 assuming the relative systematic errors in L_0, T_0, and R (denoted as $\Delta L_0/L_0$, $\Delta T_0/T_0$, and $\Delta R/R$) remain constant across the entire range of studied c-values.[179]

This triphasic dependence (**Figure 4.4b**) suggests that minimizing $|\Delta K_d/K_d|$ requires lowering the c-value through reducing L_0 and/or minimizing the relative systematic errors in L_0, T_0, and R.[179] Since decreasing L_0 is limited by the instrument's limit of quantitation (LOQ) for detecting ligand signals, in the work shown in **Section 3.2**, we have systematically explored strategies to

reduce the relative systematic errors in L_0, T_0, and R, and to determine their confidence intervals.[182]

As demonstrated earlier (***Section 3.2.2.3*** of **Chapter 3** and **Eqs (5.1)–(5.4)**), the core principles of data analysis remain consistent between ITC and the classic non-titration approach. Therefore, the results of error propagation for ITC data can be derived based on the findings (results in **Section 3.1** of **Chapter 3**) from error-propagation studies of non-titration equilibrium methods.

In this study, which focuses exclusively on analyzing the accuracy of K_d, we assume that the parameter $\Delta H°$ is directly measurable from the data points of the initial injections and set $n = 1$ without systematic error. Here, n is treated as the number of binding sites on the ligand, rather than a compensatory parameter used to address inaccuracies in ligand concentration, as is commonly practiced by many ITC practitioners.[368, 372, 373] Pre-setting n to its theoretical value of 1 (for the studied 1:1 binding process) does not imply that the ligand concentration is accurate. Instead, systematic errors in ligand concentration are explicitly included in the error propagation analysis. Under these assumptions, K_d becomes the sole parameter being determined. The systematic error in the $K_{d,det}$ (ΔK_d) arises exclusively from the systematic errors in the measured heat (Q_i), measured $\Delta H°$, nominal total ligand concentration (L_0), and the calculated molar ratio (Γ), denoted as ΔQ_i, $\Delta(\Delta H°)$, ΔL_0, and $\Delta\Gamma$, respectively.

Through comprehensive mathematical analysis (Notes D1.2 and D1.3), we found that the relative systematic errors in T_0, L_0, and R (which are the error sources for non-titration equilibrium methods) are interconvertible with the relative systematic errors in Q_i, $\Delta H°$, L_0, and Γ. As a result, if ΔQ_i, $\Delta(\Delta H°)$, ΔL_0, and $\Delta\Gamma$ are strongly correlated (which occurs when similar procedures are used for ligand and target solution preparation, and ΔQ_i and $\Delta(\Delta H°)$ are mainly

the consequence of ΔL_0 and $\Delta\Gamma$), the relationship between the accuracy of the determined K_d ($|\Delta K_d/K_d|$) and the c-value is derived as a linear equation (Notes D1.2):

$$\left|\Delta K_d / K_d\right| = A + Bc \tag{5.5}$$

where both constants A and B depend on the relative systematic errors in all four variables: $|\Delta Q_i/Q_i|$, $|\Delta(\Delta H°)/\Delta H°|$, $|\Delta L_0/L_0|$, and $|\Delta\Gamma/\Gamma|$. Note that c is the true c-value, defined as the ratio between the true values of L_0 and K_d (L_0/K_d), which are unknown due to the unavoidable systematic errors in them.

In contrast, if ΔQ_i, $\Delta(\Delta H°)$, ΔL_0, and $\Delta\Gamma$ are weakly correlated (occurring when different procedures are used for preparing ligand and target solutions, and ΔQ_i and $\Delta(\Delta H°)$ are independent of ΔL_0 and $\Delta\Gamma$), the dependence of the accuracy of $K_{d,det}$ on the true c-value (for weakly correlated error sources) is derived to be a square root of a quadratic function for the c-

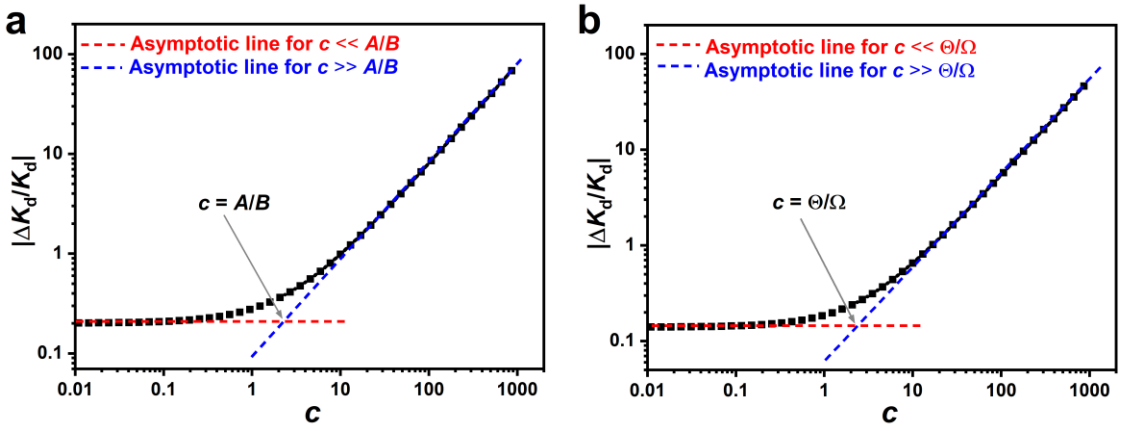


Figure 5.1. General trends in the dependence of $|\Delta K_d/K_d|$ on c-value for constant relative systematic errors in variables. Reasonable values were used for A, B, Θ, Λ, and Ω, calculated from typical values of $|\Delta Q_i/Q_i|$, $|\Delta(\Delta H°)/\Delta H°|$, $|\Delta L_0/L_0|$, and $|\Delta\Gamma/\Gamma|$ as explained in ***Section 5.1.2.5***. These values are: $|\Delta Q_i/Q_i| = |\Delta(\Delta H°)/\Delta H°| = 0.01$, $|\Delta L_0/L_0| = 0.05$, and $|\Delta\Gamma/\Gamma| = 0.1$. The dependence of $|\Delta K_d/K_d|$ on c is shown for panel **(a)** using **Eq (5.5)** and for panel **(b)** using **Eq (5.6)**. Data are presented on a log-log scale to enhance the clarity of details for low c-values.

value (Notes D1.3):

$$\left|\Delta K_{\mathrm{d}} / K_{\mathrm{d}}\right| = \sqrt{\Omega^2 c^2 + \Lambda c + \Theta^2} \tag{5.6}$$

with constants Ω, Λ, and Θ being defined by different combinations of $|\Delta Q_i/Q_i|$, $|\Delta(\Delta H°)/\Delta H°|$, $|\Delta L_0/L_0|$, and $|\Delta\Gamma/\Gamma|$, respectively.

To illustrate these relationships, we plotted the dependencies described by **Eqs (5.5)** and **(5.6)** for constant A and B (strong correlation) and for constant Ω, Λ, and Θ (weak correlation) in **Figure 5.1** (and **Figure D1.1**). As expected, the resulting curves resemble the triphasic behavior shown in **Figure 4.4b**, characterized by two linear phases at low and high c-values, separated by a nonlinear transition phase. The threshold, defined as the intercept of the two linear regions, depends on the relative errors in the variables and is given by A/B for **Eq (5.5)** (**Figure 5.1a**) and Θ/Ω for **Eq (5.6)** (**Figure 5.1b**).

5.1.2.3. Changing Relationship between Error Sources with Changing c-Value

In the earlier analysis, we assumed that error sources were either fully correlated or completely uncorrelated across the entire range of c-values, leading to the derivation of **Eq (5.5)** and **Eq (5.6)**, respectively. However, as we explain below, in practical ITC experiments, the relationship between error sources evolves with changing c-values. Specifically, **Eq (5.5)** applies to high c-values, where the systematic errors in all variables are predominantly influenced by inaccuracies in analyte concentrations. In contrast, **Eq (5.6)** is appropriate for low c-values, where weaker heat signals and incomplete binding isotherms become the dominant sources of errors in the measured heat, independent of the error sources for concentrations.

While modern ITC instruments are capable of accurately measuring the experimental heat change for low c-values (*e.g.*, $0.1 < c < 1$),[374] it is widely accepted that the lowest c-value suitable for producing a relatively complete binding curve to measure $\Delta H°$ reliably is around

10.[374, 375, 376, 377] The accuracy and precision of measured $\Delta H°$ decrease rapidly when $c < 10$.[374] Therefore, for $c \geq 10$, it is reasonable to assume that the systematic errors in measured heats (ΔQ_i and $\Delta(\Delta H°)$) are primarily driven by ΔL_0 and $\Delta \Gamma$, and error propagation can be derived using **Eq (5.5)**. However, when $c < 10$, the primary source of systematic errors in the measured heat shifts toward weaker heat signals and incomplete binding isotherms. This results in a gradual increase in the magnitude of ΔQ_i and a rapid increase in the magnitude of $\Delta(\Delta H°)$, both of which become independent of ΔL_0 and $\Delta \Gamma$. In this case, **Eq (5.6)** should be used to produce error propagation.

Notably, for a molecular binding pair, since the true K_d value is a constant, a decrease in the true c-value (L_0/K_d) corresponds to a decrease in ligand concentration L_0. Consequently, to maintain the same molar ratio range, the target concentration in the syringe must also decrease if the injection volume and number of injections remain unchanged. This reduction in concentration poses greater challenges in accurately preparing low-concentration samples and amplifies the impact of unavoidable reagent loss during sample preparation. As a result, lower c-values are often associated with increased systematic errors in concentrations and, consequently, in molar ratios.[378] Nevertheless, for $0.1 < c < 10$, compared to the rapid increase of $\Delta(\Delta H°)$, the increase of systematic errors in concentrations and Q_i is much less significant.[368] For $c \geq 10$, relative systematic errors in concentrations and heat measurements are primarily constrained by the maximum accuracy limits of the experimental equipment, which remain approximately constant.

In this study, we focus on the broader c-value range of $0.1 < c < 1000$, which extends significantly beyond the commonly recommended range of $10 \leq c \leq 100$. Within this range, we assume that $|\Delta L_0/L_0|$, $|\Delta \Gamma/\Gamma|$, and $|\Delta Q_i/Q_i|$ remain constant. However, $|\Delta(\Delta H°)/\Delta H°|$ exhibits a

rapid increase as the c-value decreases within $0.1 < c < 10$. For $10 \leq c < 1000$, we approximate $|\Delta(\Delta H°)/\Delta H°| \approx |\Delta Q_i/Q_i|$.

5.1.2.4. Realistic Relative Systematic Errors in Variables for ITC Experiments

To analyze how the relative systematic error in $K_{d,det}$ ($|\Delta K_d/K_d|$) determined by ITC depends on the true c-value (L_0/K_d), we first need to examine the experimentally realistic relative systematic errors in the variables: $|\Delta Q_i/Q_i|$, $|\Delta(\Delta H°)/\Delta H°|$, $|\Delta L_0/L_0|$, and $|\Delta \Gamma/\Gamma|$.

Modern ITC instruments offer high precision in injection volumes, leading to highly precise calculated values of T_0 and L_0, as well as their molar ratios, with relative random errors typically within 1–2%.[106] However, as described in ***Section 3.2.3.3*** of **Chapter 3**, after eliminating common sources of systematic errors in analyte concentrations (e.g., reagent impurity and adsorption), the accuracies of T_0 and L_0 are constrained by sample preparation procedures.[182] Given the accuracy limitations of preparation instruments (*e.g.*, analytical balances and pipettes) and the unavoidable impurities in reagents, the relative systematic error in these concentrations ($|\Delta L_0/L_0|$ and $|\Delta T_0/T_0|$) is unlikely to fall below 5%.[106, 359] When considering error propagation in calculating molar ratios from the initial ligand concentration (in the sample cell before titration starts) and the target concentration (in the syringe), along with injection and sample cell volumes, achieving a relative systematic error below 10% in the computed molar ratio ($|\Delta \Gamma/\Gamma|$) becomes a considerable challenge, given the additive nature of error propagation ($|\Delta \Gamma/\Gamma| > |\Delta L_0/L_0| + |\Delta T_0/T_0|$). However, with a well-calibrated ITC instrument and careful baseline adjustment, the relative systematic error in measured heat change can be effectively controlled within 1–2%.[379]

Regarding the relative systematic error of $\Delta H°$ ($|\Delta(\Delta H°)/\Delta H°|$), we assume the following: (*i*) when $0.1 < c < 10$, the relative systematic error for measured $\Delta H°$ ($|\Delta(\Delta H°)/\Delta H°|$) increases

exponentially as the c-value decreases, such as $|\Delta(\Delta H°)/\Delta H°| = \exp(-4.6c)$ and (*ii*) when $c \geq 10$, $|\Delta(\Delta H°)/\Delta H°|$ remains constant and equal to a typical relative systematic error of Q_i, *i.e.*, 1%. The constant 4.6 in assumption (*i*) defines the rate at which $|\Delta(\Delta H°)/\Delta H°|$ increases as the c-value decreases (for $0.1 < c < 10$) and, importantly, ensures that $|\Delta(\Delta H°)/\Delta H°|$ approaches 0.01 (1%) when $c = 10$ for a smooth transition between the two parts of the dependence.

5.1.2.5. Dependence of Relative Systematic Errors in K_d on c-Value in a Wide Range of c-Values.

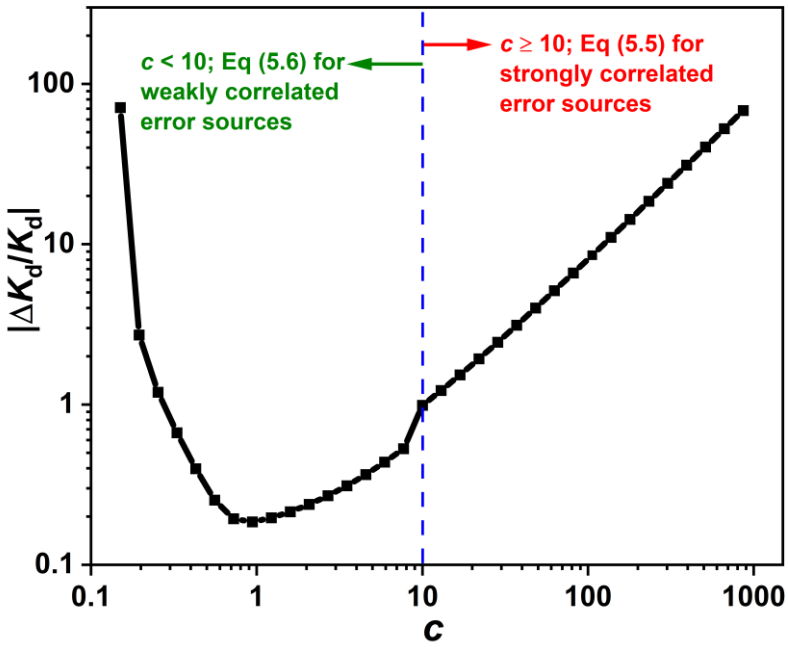


Figure 5.2. General trends in the dependence of $|\Delta K_d/K_d|$ on c-value under realistic experimental conditions. The parameters A, B, Θ, Λ, and Ω were calculated using typical values for $|\Delta Q_i/Q_i|$, $|\Delta(\Delta H°)/\Delta H°|$, $|\Delta L_0/L_0|$, and $|\Delta\Gamma/\Gamma|$, consistent with those in **Figure 5.1**, and held constant across the entire c-value range. For $|\Delta(\Delta H°)/\Delta H°|$, we assume $|\Delta(\Delta H°)/\Delta H°| = \exp(-4.6c)$ for $c < 10$ and $|\Delta(\Delta H°)/\Delta H°| = 0.01$ for $c \geq 10$. The dependence of $|\Delta K_d/K_d|$ on c was calculated using **Eq (5.6)** for $0.1 < c < 10$ and **Eq (5.5)** for $10 \leq c < 1000$. Data are presented on a log-log scale to enhance the clarity of details for low c-values.

Based on the discussions in the previous section, we assign reasonable relative systematic errors to ligand concentration, molar ratio, and heat change as $|\Delta L_0/L_0| = 0.05$, $|\Delta \Gamma/\Gamma| = 0.1$, and $|\Delta Q_i/Q_i| = 0.01$, respectively, and establish the dependence of $|\Delta K_d/K_d|$ determined by ITC on the *c*-value, as illustrated in **Figure 5.2** (and **Figure D1.2**). In this analysis, we assume that all common and controllable sources of systematic error, such as those arising from reagent impurities, dilution effects, and instrument mis-calibration, have been eliminated. The remaining systematic errors are assumed to originate from unavoidable limitations in sample preparation instruments and baseline selection. To streamline the semi-quantitative analysis, we introduce realistic systematic errors only into L_0, Γ, Q_i, and $\Delta H°$, while omitting potential errors from injection and cell volumes. This simplified approach allows us to illustrate the general dependence of $|\Delta K_d/K_d|$ on the true *c*-value and to estimate the potential magnitude that $|\Delta K_d/K_d|$ can reach.

The results in **Figure 5.2** demonstrate that the dependence of $|\Delta K_d/K_d|$ on the true *c*-value in ITC follows a U-shaped trend. At very low *c*-values (*e.g.*, $0.1 < c < 1$), $|\Delta K_d/K_d|$ increases as *c* decreases, primarily because the exponentially increasing relative systematic error of $\Delta H°$ becomes the dominant factor in error propagation. As *c* increases to a certain range (*e.g.*, around $c \approx 1$), $|\Delta K_d/K_d|$ reaches its minimum, suggesting that this range offers an optimal balance for accurate K_d determination, providing sufficient accuracy in the measured $\Delta H°$ while the amplification of systematic errors in variables (by *c*-value) remains insignificant. However, as *c*-values exceed this optimal point, $|\Delta K_d/K_d|$ starts to increase once again. This is due to the significant magnification of systematic errors at high *c*-values, as suggested by **Eq (5.5)**.

In this demonstration shown in **Figure 5.2**, we did not examine the dependence of $|\Delta K_d/K_d|$ on *c* for $c < 0.1$, as this range is well below the typical lower limit ($c = 10$) for standard ITC

operating conditions. At such extremely low c-values ($c < 0.1$), $|\Delta(\Delta H°)/\Delta H°|$ increases rapidly, and the relative systematic errors in Q_i and concentrations also rise significantly, albeit at unknown rates.

The general trend in the dependence of $|\Delta K_d/K_d|$ on the c-value, as illustrated in **Figure 5.2**, qualitatively agrees with the results of previous numerical studies.[56, 313, 380, 381] To our best knowledge, however, this is the first time this trend has been revealed through analytical error propagation. It is important to emphasize that the results presented in **Figure 5.2** may differ quantitatively from actual experimental outcomes of K_d accuracy assessments, which are typically presented as the dependence of $K_{d,det}$ on the determined c-value. Several factors can contribute to such quantitative differences, and we list a few of them below.

First, the relative systematic errors assigned to the variables used to generate **Figure 5.2** may differ from the actual errors in experiments. Second, the error propagation in this study is applied only to the specific case where the fraction of unbound ligand R is close to 0.5, rather than across the full range of R in nonlinear regression. Third, we assume that $\Delta H°$ is directly measurable, whereas, in real ITC experiments, $\Delta H°$ is a parameter to be determined. Finally, ITC data analysis often includes stoichiometry n as an additional parameter to estimate, alongside K_d and $\Delta H°$, even though n is theoretically equal to unity in a 1:1 binding model.[368, 372, 373]

These four factors prevent a direct comparison of the theoretical results in **Figure 5.2** with the experimental dependencies of ITC-determined K_d values on experimentally determined c-values. However, none of these factors significantly alters the overall trend depicted in **Figure 5.2**, as they cannot eliminate the systematic errors in the experimental variables embedded within the constants of **Eq (5.5)** and **Eq (5.6)**. For instance, while the mutual interactions among the

uncertainties of the three common ITC-derived parameters — K_d, $\Delta H°$, and n — are difficult to capture through analytical error propagation, there is robust evidence that these interactions do not negate the trends shown in **Figure 5.2**.[368, 382] Specifically, in most cases, allowing n to float as a fitting parameter tends to further increase the uncertainty in the determined K_d.[368] Additionally, the uncertainty in K_d has been shown to correlate positively with the uncertainty in $\Delta H°$,[382] indicating that the uncertainties of these two parameters propagate to one another rather than compensating each other.

Furthermore, in ITC data fitting, some researchers incorporate additional parameters, such as a constant or polynomial representing the heat of dilution or baseline offset, into the regression model.[378] While increasing the number of fitting parameters may improve the apparent fitting quality, there is no mathematical evidence that this approach enhances the accuracy of critical fitted parameters, such as K_d and $\Delta H°$. In fact, unknown interactions between the systematic errors of various parameters may degrade the reliability of all fitted values.[368, 378] To ensure robust and meaningful results, the number of parameters included in the ITC data fitting model should be minimized.[383] Parameters that can be determined experimentally or independently should be measured separately rather than estimated during the fitting process. For example, the heat change resulting from dilution is best addressed by conducting a control titration to directly measure the dilution heat and then subtracting it from the ITC data.[384, 385] Similarly, baseline determination is an essential step in ITC data analysis, but should be performed independently of the binding isotherm fitting process. This can be achieved through manual baseline selection or rigorous baseline-determination procedures,[168] rather than incorporating baseline as an additional parameter in the regression model. Thus, separating baseline determination and other extraneous

factors from the fitting process helps avoid overfitting and ensures that the determined binding parameters, such as K_d and $\Delta H°$, are more accurate and reliable.

Therefore, the purpose of **Figure 5.2** here is not to provide quantitative guidance on an optimal *c*-value for practitioners because the actual experimental systematic errors for each variable and the true *c*-values (without systematic errors) are unknown in real experiments. Instead, **Figure 5.2** demonstrates that within the traditionally recommended optimal *c*-value range of 10 to 100, the K_d values determined by ITC may still deviate significantly from the true K_d.

5.1.2.6. Improving the Reliability of Experimentally Determined K_d Values in ITC.

The findings in **Figure 5.2** underscore that even for the widely regarded K_d-determination method, ITC, significant systematic errors can persist in the determined K_d values, even when experiments are conducted under recommended conditions. This highlights the critical need to rigorously assess the accuracy of ITC-derived parameters and also to develop other reliable techniques for K_d-determination.

While several Monte Carlo-based tools, such as SEDPHAT and *pytc*, have been developed for assessing the accuracy or precision of ITC-derived parameters,[176, 367, 368, 369] they are underutilized by ITC practitioners. This lack of adoption can be attributed to two primary factors: limited awareness of the potential magnitude of inaccuracies in $K_{d,det}$ and the relatively steep learning curve required to effectively use these programs. Moreover, many existing statistical methods fail to clearly define the necessary input information for accurately assessing ITC-derived parameters. Consequently, there is a strong demand for a comprehensive study on the necessary inputs for accuracy assessment and the development of a more user-friendly, accessible statistical tool tailored to the needs of the field.

It is important to emphasize that, although ITC is widely regarded as the primary confirmatory method, other techniques exist that offer comparable or even superior accuracy. A detailed review of these alternative methods is provided in **Section 1.4.1** of **Chapter 1**. Each of these methods has distinct advantages and limitations. Achieving the highest accuracy in K_d determinations often requires employing multiple complementary methods with meticulous attention to experimental design, implementation, and data analysis.[12]

5.1.3. Concluding Remarks

Building on the work in **Section 3.1** of **Chapter 3**,[179] we investigated the theoretical relationship between the accuracy of ITC-derived K_d ($|\Delta K_d/K_d|$) and the true *c*-value using analytical error propagation. Our analysis reveals a U-shaped dependence of $|\Delta K_d/K_d|$ on the *c*-value (**Figure 5.2**). At very low *c*-values, $|\Delta K_d/K_d|$ increases significantly as *c* decreases. This trend is primarily due to the dominance of the rapidly increasing relative systematic error in $\Delta H°$, which drives the overall propagation of errors. As *c* increases, $|\Delta K_d/K_d|$ reaches a minimum in a certain range, indicating an optimal balance where the amplification of systematic errors in variables by the *c*-value (according to **Eq (5.5)**) is insignificant, and $\Delta H°$ measurements remain sufficiently accurate. However, as *c*-values exceed this optimal range, $|\Delta K_d/K_d|$ increases once more due to the amplification of systematic errors at high *c*-values. These findings confirm conclusions from the previous numerical studies but, to our knowledge, represent the first demonstration of the described trends using analytical error propagation.

While our theoretical results may differ quantitatively from experimental observations due to variations in experimental and data processing methods, the overall trends shown in **Figure 3** remain robust. These trends underscore that even within the traditionally recommended *c*-value range of 10 to 100, ITC-derived K_d values can deviate significantly — by several folds — from

the true K_d values. These findings highlight the critical need for assessing the accuracy of K_d values derived from ITC experiments.

Although Monte Carlo-based tools for accuracy assessment are available, they are largely underutilized by ITC practitioners. This is partly due to a tendency to overestimate the accuracy of ITC measurements, and the challenging learning curve associated with mastering these software programs. Additionally, many statistical methods lack clear input requirements for accuracy evaluation. While a recently introduced tool for computing the accuracy confidence interval (ACI) for K_d using the classic non-titration approach represents a significant advancement,[11] it is not directly applicable to ITC. This highlights the urgent need for a user-friendly, accessible tool and a clearer understanding of necessary inputs for assessing the accuracy of ITC-derived parameters.

5.2 A Browser-Based Tool for Assessing Accuracy of ITC-Derived Parameters: K_d, $\Delta H°$, and n

The following material is reprinted from the article "**Wang, T. Y.**; Bijlani, A.; Chao, H. P.; Johnson, P. E.; Krylov, S. N. A Browser-Based Tool for Assessing Accuracy of ITC-Derived Parameters: K_d, $\Delta H°$, and n. *ChemBioChem*, 2500194, published online." My contributions to the article were: (*i*) conducting all the mathematical analysis, (*ii*) contributing to program development, (*iii*) performing virtual experiments, (*iv*) preparing all figures, (*v*) interpreting the results, and (*vi*) writing the manuscript. Co-author Amit Bijlani converted the Python code to a web application. Co-author Emily Hoi Pui Chao conducted the ITC experiments.

5.2.1. Introduction: The Need for an Accessible Tool to Assess ITC Accuracy

Isothermal titration calorimetry (ITC) is broadly regarded as the gold standard for validating K_d values determined through high-throughput screening.[52, 53, 54] Unlike many other methods, ITC quantifies binding thermodynamics directly in solution without requiring any reactant modifications such as labeling or immobilization.[53, 184, 185] By measuring heat changes during binding events, ITC provides a comprehensive thermodynamic profile, including K_d, enthalpy change ($\Delta H°$), and binding stoichiometry (n). Additionally, ITC instrumentation ensures highly reproducible titration and measurements, providing exceptional precision in the determined parameters.[53, 386] However, ITC has limitations, such as requiring relatively large quantities of reactants and having low throughput.[53] Despite these drawbacks, ITC remains a critical tool for confirming K_d values derived from high-throughput methods.

While being highly precise, ITC is prone to inaccuracies in the determined K_d values due to systematic errors in experimental variables, such as analyte concentrations and measured heat.[56, 106] These inaccuracies can significantly affect the reliability of the resulting K_d values. Under typical operating conditions, ITC is recommended to be used at c-values (the ratio between the true titrand concentration and the true K_d value) significantly exceeding unity to ensure sufficient curvature in the binding isotherm for reliable parameter fitting. However, multiple studies (including our study shown in **Section 5.1**) suggest that such high c-values can amplify systematic errors in input variables, leading to significant deviations of the determined K_d values from their true values.[373, 378] These findings indicate that even under standard experimental conditions, ITC-derived K_d values may be more prone to inaccuracies than broadly appreciated. Without their quantitative assessment, such inaccuracies can result in misleading conclusions about binding affinities and compromise downstream decision-making. By quantifying these

inaccuracies, researchers can better leverage ITC's strengths while mitigating potential errors in critical decision-making steps.

To address the challenges associated with assessing the accuracy of K_d values in the classic non-titration approach (commonly used in techniques such as fluorescence anisotropy, microscale thermophoresis, BLI, and SPR), we introduced the ACI-Kd (and ACI-Km) tool (**Chapter 4**).[10] The term Accuracy Confidence Interval (ACI) refers to a confidence interval (CI) that specifically accounts for systematic errors, which is why the word "accuracy" is included in its name. ACI-Kd effectively assesses systematic errors in non-titration methods using analytical (non-numerical) error propagation to assess K_d accuracy avoiding the complexities of statistical numerical methods and offering a practical, accessible approach.[10] Moreover, ACI-Kd is implemented as a user-friendly web application (https://aci.sci.yorku.ca), requiring only a web browser and making it broadly accessible.[10] However, ACI-Kd is not designed for ITC, which requires accounting for additional unknowns, such as $\Delta H°$ and n.

Monte Carlo simulations have long been proposed as a robust statistical method for assessing the accuracy of ITC-derived K_d values.[141, 176, 367, 373] These simulations estimate confidence intervals (CIs) and error distributions by conducting numerous virtual experiments that incorporate random variations in experimental parameters, such as analyte concentrations and measured heat.[141, 176, 367, 373] By simulating a wide range of conditions, Monte Carlo methods can identify how systematic biases impact derived parameters.[141, 176, 367, 373]

While Monte Carlo simulations are relatively straightforward for propagating random errors (indicators of precision), they may be excessive for precision evaluation, as standard nonlinear regression already provides precision estimates in the form of standard deviations.[141, 176, 367, 373] In contrast, applying Monte Carlo simulations to assess K_d accuracy is fully justified but

significantly more challenging, as it involves propagating systematic errors — less obvious and more complex parameters. Estimating these systematic errors requires a deep understanding of their nature, origins, and dependence on experimental design. This complexity, combined with the often-overlooked potential for large inaccuracies in ITC-derived K_d values and the technical expertise required to apply Monte Carlo techniques, likely explains their limited use. Simplified tools are essential to bridge these gaps and make accuracy assessments accessible to a broader research community.

Here, we address the challenges of quantitatively assessing the accuracy of ITC-derived K_d values. We present ACI-ITC, a Monte Carlo-based tool that calculates the ACI for K_d, $\Delta H°$, and n. Unlike traditional Monte Carlo approaches, ACI-ITC is optimized for usability and requires no programming or advanced statistical knowledge, making it accessible to a broad range of molecular scientists. To further enhance usability, ACI-ITC is integrated into the ACI web application (https://aci.sci.yorku.ca), which requires no software installation and provides comprehensive reports that integrate accuracy with precision for K_d, $\Delta H°$, and n, along with publication-quality graphs. Additionally, we offer comprehensive instructions for estimating input parameters, such as systematic errors in concentrations and heat, to facilitate practical implementation.

By simplifying the complexities of propagating systematic errors, ACI-ITC enables researchers to generate reliable data for K_d, $\Delta H°$, and n from ITC experiments. This advancement enhances the reliability of ITC-derived K_d values and their applications in drug discovery, allowing researchers to draw more accurate conclusions about molecular interactions.

5.2.2. Materials and Methods

5.2.2.1. Materials

All DNA aptamer samples were obtained from Integrated DNA Technologies (Coralville, IA, USA) as a lyophilized powder and used without further purification. Salts and buffer reagents were obtained from BioShop (Burlington, ON, Canada). All small molecule ligands were obtained from Sigma Aldrich (Oakville, ON, Canada). DNA samples were dissolved in distilled deionized H_2O (ddH_2O), exchanged three times using a 3 kDa molecular weight cut-off concentrator with 1.0 M NaCl, and washed three times with ddH_2O. The concentrations of the aptamers were determined by measuring the OD_{260} in a UV-Vis spectrometer using the extinction coefficients provided by IDT.

5.2.2.2. ITC Experimental Procedure

ITC was performed using a MicroCal-ITC 200 (Malvern, Worcestershire, UK) instrument. Samples were degassed and centrifuged to remove any bubbles and aggregates before setting up the experiment. Titrations were performed with the aptamer samples (ligand) loaded in the cell and the target as the titrant was in the syringe. Binding experiments were performed at 20 °C with both the DA-3BP aptamer (5'-ACG TCA GTT TGA AGG TTC GTT CGC AGG TGT GGA GTG ACG T-3'),[387] and the Theo2201 aptamer (5'-GAC GAC GAT TGT GGT CTA TTC ATA GGC GTC CGC TGA GTC GTC-3').[388] ITC was performed with the DA-3BP concentration set at 450 μM and the dopamine (target) concentration at 7,020 μM (7.02 mM). The Theo2201 aptamer concentration was set at 60 μM and the theophylline (target) concentration at 936 μM.

A typical binding experiment consisted of 19 successive injections spaced every 180 s apart where the injection volume was 2 μL with 10 μcal/s reference power and a 750 rpm stir speed. The data was preliminarily analyzed using Microcal PEAQ software to produce the binding isotherm "Δ*H* vs Molar ratio". The software automatically sets the baseline and integrates every

peak from the baseline, but manual integration is often needed as software integration is not always performing optimally.

5.2.3. Results and Discussion

5.2.3.1. Browser-Based ACI Calculator

To evaluate the accuracy of ITC-derived binding parameters considering systematic errors in experimental variables, such as concentrations and measured heat changes, we first developed a Python program named ACI-ITC. In ITC measurements, the heat change per injection is typically normalized to the heat change per mole of the injected target, denoted as the enthalpy change ΔH, with units of kcal/mol. To clarify, ΔH will be used throughout this text to refer to

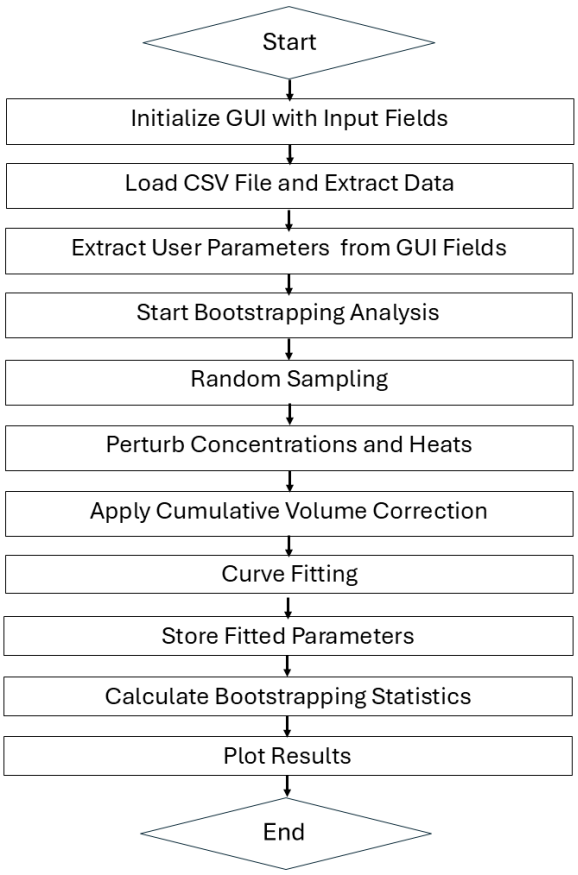


Figure 5.3. Streamlined computational algorithm of the ACI-ITC program.

the “measured heat changes”, while $\Delta H°$ will specifically denote the enthalpy change per mole of the formed complex, which is one of the parameters determined in ITC experiments and reported.

The ACI-ITC program employs Monte-Carlo-based bootstrapping to compute the ACI for three binding parameters — K_d, $\Delta H°$, and n — with 95% confidence. The statistical systematcrationale and methodological foundations of this approach, along with a comparison to conventional nonlinear regression, are detailed in Note D2.1. A streamlined computational algorithm for ACI-ITC is depicted in **Figure 5.3**, while a more detailed version of the algorithm is presented in **Figure D2.1**. To ensure broad accessibility and ease of use for individuals with diverse technical backgrounds, we integrated the ACI-ITC program into the browser-based ACI web application (https://aci.sci.yorku.ca), providing a user-friendly platform for the field and integrating other tools for assessing the accuracy of physicochemical parameters. This free web app is accessible from any browser on any device and requires no programming skills. It accepts ITC binding isotherms, represented as “ΔH vs Molar Ratio,” directly from commercial software integrated with the ITC instrument for calculations. The web app produces and downloads a report including the input and output data, a high-quality binding curve “Cumulative ΔH vs Molar ratio” with bootstrapping fitting lines, and the frequency-distribution histograms of the determined values of K_d, $\Delta H°$, and n. An example of the graphs produced by ACI-ITC is shown in **Figure 5.4**. Notably, “Cumulative ΔH” refers to the total heat change measured up to the i-th injection. It is the summation of the measured ΔH values for the first i injections at the molar ratio corresponding to the i-th injection. From this point forward, we will denote the cumulative ΔH after the i-th injection as $\Delta H_{cum,i}$.

5.2.3.2. Necessary Inputs for Computing ACI

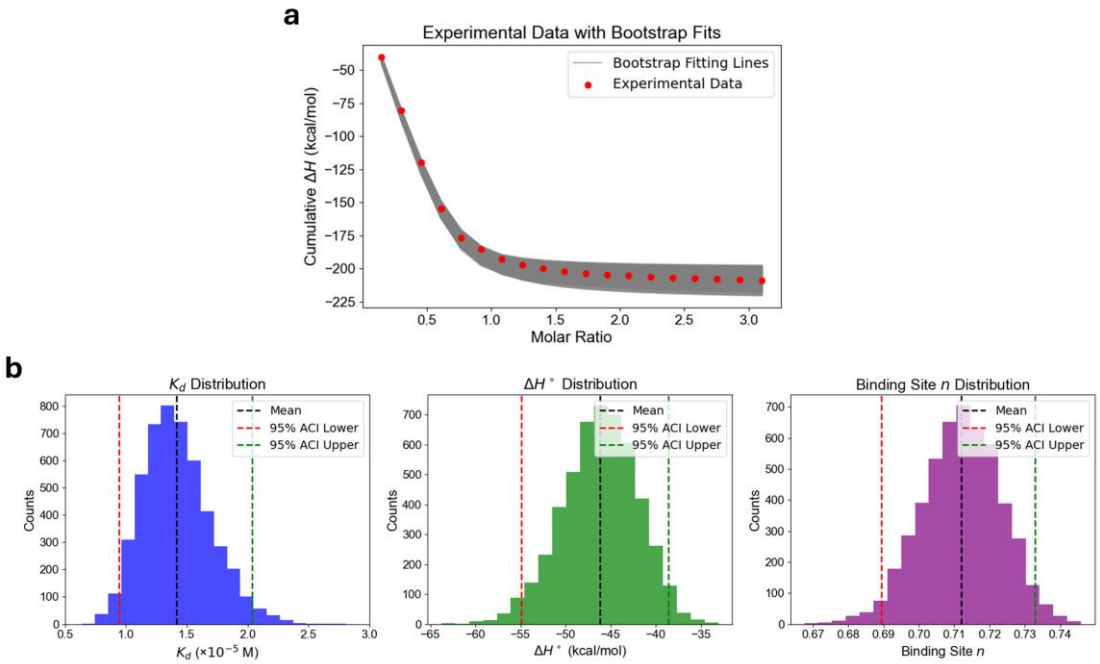


Figure 5.4. Example of graphs generated by the ACI-ITC web application (http://aci.sci.yorku.ca): **(a)** Bootstrap fitting results for an ITC binding curve "Cumulative ΔH vs Molar ratio**"**, where the red dots represent experimental data, and the gray-shaded region contains bootstrap-fitted curves. The ITC binding curve was derived from an ITC experiment using the DA-3BP aptamer–dopamine binding pair. See the experimental demonstration section for more details. **(b)** Frequency distributions of the determined K_d, $\Delta H°$, and n, derived from 5000 bootstrap iterations. The y-axis in each histogram represents the number of occurrences for each parameter value.

To use the ACI-ITC program, users must input the following information: the nominal concentration of the target in the syringe (T_0), the nominal initial concentration of the ligand in the sample cell (L_0), sample/reference cell volume (V_0), injection volume per step (V_{inj}), total number of injections (N), relative random errors in the concentrations of analyte stock solutions ($\delta T_0/T_0$ and $\delta L_0/L_0$), random error for the measured heat changes ($\delta(\Delta H)$), and the preferred

number of bootstrapping iterations. The reason for $\delta T_0/T_0$, $\delta L_0/L_0$, and $\delta(\Delta H)$ to be used rather than systematic errors in T_0, L_0, and ΔH is explained in the next section.

The default minimum number of bootstrapping iterations is set at 2,000 to ensure data convergence. However, users can input a higher value to obtain more detailed frequency distributions for the determined parameters. After users enter or load the experimental binding isotherm "ΔH vs Molar ratio", the web app completes the ACI calculations within approximately 10 s for 2,000 bootstrapping iterations; the time increases with the increasing number of iterations.

The binding isotherm "ΔH vs Molar ratio" is a typical output from modern ITC instruments.[53] In the ACI-ITC program, this data is internally converted to the binding curve "$\Delta H_{\mathrm{cum},i}$ vs Molar ratio", which is then fitted using **Eq (5.7)**:

$$\Delta H_{\mathrm{cum},i} = \frac{\Delta H_{\mathrm{cum,max}}}{2nL_0}\left[\left(K_\mathrm{d} + L_0\Gamma + nL_0\right) - \sqrt{\left(K_\mathrm{d} + L_0\Gamma + nL_0\right)^2 - 4n\Gamma L_0{}^2}\right] \qquad (5.7)$$

Here, Γ represents the molar ratio of target to ligand in the sample cell; $\Delta H_{\mathrm{cum,max}}$ denotes the maximum cumulative ΔH value, which theoretically corresponds to $\Gamma = \infty$, at which all ligands in the sample cell bind to targets, forming complexes. K_d, $\Delta H_{\mathrm{cum,max}}$, and n are the three parameters to be determined. Notably, rather than representing binding stoichiometry, the parameter n is often interpreted as a correction factor for the ligand concentration in the 1:1 binding model.[313, 372] It is worth emphasizing that using the first derivative of **Eq (5.7)** (with respect to the amount of injected target) has been shown to offer advantages over directly fitting **Eq (5.7)**.[370] However, in this work, we use **Eq (5.7)** as the fitting equation because it is directly derived from the mass balance of analytes and the fundamental definitions of K_d and the measured bound/unbound fraction of the ligand.[25] This formulation provides a clear and intuitive framework, making it the most easily interpretable fitting model for biochemists.

Based on the determined $\Delta H_{cum,max}$, the enthalpy change per mole of the formed complex ($\Delta H°$) is calculated with **Eq (5.8)**:

$$\Delta H° = \frac{\Delta H_{cum,max} V_{inj} T_0}{n V_0 L_0} \tag{5.8}$$

where all parameters retain their definitions as described above.

Therefore, except for the random errors $\delta T_0/T_0$, $\delta L_0/L_0$, and $\delta(\Delta H)$, all other inputs are either known based on the experimental design or obtained from standard ITC measurements. Next, we will explore the relationship between the random errors of T_0, L_0, and ΔH ($\delta T_0/T_0$, $\delta L_0/L_0$, and $\delta(\Delta H)$) and their systematic errors. We will also understand how the values of $\delta T_0/T_0$, $\delta L_0/L_0$, and $\delta(\Delta H)$ are determined.

5.2.3.3. Assessing CIs of Systematic Errors in Variables

The systematic errors in ITC-derived parameters K_d, $\Delta H°$, and n arise from inaccuracies in experimental variables, including analyte concentrations, detected heat changes, injection volumes, and sample cell volume. However, because the systematic errors associated with injection and sample cell volumes are tied to inherent instrumental accuracies that cannot be mitigated and measured by experimentalists, this work focuses solely on assessing the systematic error ranges for analyte concentrations and detected heat changes, and on propagating them into the ACI of the determined parameters. Disregarding systematic errors in the injection and sample cell volumes leads to the resulting ACI being narrower than the real ACI.

The actual systematic errors in the experimental variables are inherently unknown. If these errors were known, assessing the accuracy of the parameters would be unnecessary, as the input values could be mathematically corrected to yield accurate parameters. Therefore, only the CIs of the systematic errors in the experimental variables can be determined.

As has been explained in detail elsewhere,[10, 182] the CIs of relative systematic errors in analyte concentrations can be estimated by determining the relative random errors in the concentrations of stock solutions. The conversion between relative random errors and confidence intervals for systematic errors of the variables used in K_d determination requires a valid assumption under specific experimental conditions.[10, 182] First, all known means should be applied to minimize systematic errors in the variables, such as analyte concentrations and measured signal, *e.g.*, heat change in ITC experiment. Given the experimental similarities between K_d determination using ITC and other equilibrium methods, the error-minimization approaches detailed in our previous work,[182] are applicable for reducing systematic errors of variables in ITC experiments.

After minimizing systematic errors in analyte concentrations and measured signals, it is reasonable to assume that the CIs for the remaining relative systematic errors in analyte concentrations (*i.e.*, $\Delta T_0/T_0$ and $\Delta L_0/L_0$) are primarily governed by the relative random errors observed across independently prepared sample solutions (*i.e.*, $\delta T_0/T_0$ and $\delta L_0/L_0$). In non-titration equilibrium approaches, these sample solutions refer to independently prepared stock solutions of the target and ligand used to prepare each equilibrium mixture.[182] In contrast, for ITC experiments, the relevant solutions are those loaded into the syringe and sample cell, namely, the concentrated target solution and the ligand solution, used consistently throughout the entire titration. As a result, the random error associated with preparing the target and ligand solutions, typically involving multiple steps, effectively manifests as the systematic error in analyte concentrations within a single ITC experiment. The CI for the systematic error in the measured heat change (ΔH) is primarily determined by the random errors in baseline adjustments, as a single baseline adjustment is applied to the entire dataset from an ITC titration experiment.[10, 53, 182] Consequently, the random error associated with repeated stock solution

preparations effectively becomes the systematic error in analyte concentrations within a single ITC experiment. The CI for the systematic error in the measured heat change (ΔH) is primarily determined by the random errors in baseline adjustments, as a single baseline adjustment is applied to the entire dataset from an ITC titration experiment.[168, 389]

Experimentally, $\delta T_0/T_0$ and $\delta L_0/L_0$ can be determined by measuring spectroscopic signals (*e.g.*, absorbance) of the target and ligand across multiple independent sample.[10, 182] In contrast, assessing confidence intervals for heat changes remains considerably more challenging. Ideally, the buffer compositions of the target and ligand should be carefully matched to minimize heat changes arising from buffer mixing.[390] Moreover, control titration experiments should be meticulously designed and executed to account for heat changes unrelated to binding events, such as those caused by dilution or stirring.[54] In addition to these experimental optimizations and data corrections, the selection of the baseline for ITC thermograms during data analysis plays a critical role in the accuracy of determined heat change for each injection.[168, 389]

Multiple approaches have been developed to improve the accuracy of baseline determination;[168, 389 391] however, the true baseline is fundamentally unknown, and its selection is often subject to methodological or subjective biases. Unlike some other experimental parameters, the baseline cannot be determined independently through control experiments,[168] and no robust method currently exists to rigorously assess the confidence interval of baseline offsets. Nevertheless, reasonable estimates can be made based on the maximum heat pulse detected in the titration experiment,[106] which will be discussed in the next section.

Because the baseline is selected only once for the entire titration, any random error in its selection propagates as an absolute systematic error in the measured heat changes. This systematic error impacts all heat changes used for downstream data analysis, introducing bias

and compromising the overall accuracy of the analysis. The lack of a reliable method to assess the CI of baseline offsets remains a critical challenge in the ITC field, directly impacting the accuracy of binding thermodynamic measurements. Researchers are encouraged to develop innovative approaches to address this gap, advancing the accuracy and reliability of ITC data analysis.

It is important to emphasize that in conventional ITC data analysis using least squares regression (*e.g.*, Levenberg-Marquardt algorithm), all of the systematic errors discussed above are typically ignored (assumed to be zero).[53] Furthermore, the frequency (probability) distributions of the determined parameters are conventionally assumed to follow a Gaussian shape.[392] However, when accounting for the CIs associated with systematic errors in the variables, these assumptions no longer hold true. As a result, the frequency distributions of the determined parameters can change significantly.

5.2.3.4. Minimum Values of Systematic Errors in Input Variables

Modern ITC instruments deliver highly precise injection volumes, resulting in precise values of T_0 and L_0, as well as their molar ratios, with relative random errors typically within 1–2%.[393] However, the accuracies of T_0 and L_0 are fundamentally constrained by the limitations of sample-preparation methods. These limitations stem from the inherent precision of preparation instruments (*e.g.*, analytical balances and pipettes) and unavoidable reagent impurities. Consequently, the relative systematic errors in these concentrations ($|\Delta L_0/L_0|$ and $|\Delta T_0/T_0|$) are unlikely to fall below 5%.[359]

With a well-calibrated ITC instrument and careful baseline adjustments, systematic errors in measured heat changes ($|\Delta(\Delta H)|$) can typically be maintained within 2% of the maximum absolute value of observed ΔH.[106] This maximum value is often observed in the first reliable

measurement (commonly corresponding to the second injection) and is denoted here as $|\Delta H_1|$. Based on this, it is reasonable to assume that the probability distribution of $\Delta(\Delta H)$ is normal and has a standard deviation (random error) of $\delta(\Delta H) = 0.02|\Delta H_1|$. Users can therefore input $\delta(\Delta H) = 0.02|\Delta H_1|$ as the minimum standard deviation for measured heat changes until a robust method for assessing CIs of baseline offsets becomes available.

For inputting information into the ACI-ITC program, users should provide $\delta T_0/T_0$ and $\delta L_0/L_0$ values based on the relative random errors determined from multiple independent target and ligand solution preparations, with minimum values of 0.05 (5%). In cases of retrospective data where CIs for systematic errors in concentrations are unavailable, users may reasonably assume $\delta T_0/T_0 = \delta L_0/L_0 = 0.05$ as the minimum value. At present, due to the lack of reliable methods for determining CIs for detected heat changes, users are advised to input $\delta(\Delta H) = 0.02|\Delta H_1|$ as a practical estimate to compute the narrower limits for the ACI of K_d, $\Delta H°$, and n. The ACI-ITC utilizes the mentioned values as preset defaults that can be changed by users.

5.2.3.5. Verification of ACI-ITC

We rigorously verified the ACI-ITC program and its web application using synthetic ITC data generated with a Python-based simulator. These simulations utilized predefined (true) values for K_d, $\Delta H°$, and n ($n = 1$), initial concentrations L_0 and T_0, and their associated systematic errors, as well as the systematic errors in the heat change per injection. The complete verification process and results are detailed in Note D2.2. The outcomes from the ACI-ITC web app for the synthetic data were in full agreement with theoretical predictions, confirming the reliability and accuracy of both the ACI-ITC program and its web app (http://aci.sci.yorku.ca).

5.2.3.6. Application of ACI-ITC to Experimental Data

Following the verification of the ACI-ITC web app, we applied it to evaluate the ACI for the parameters determined in real ITC experiments. Our goal was to compare the ACI of K_d, ΔH°, and n computed by ACI-ITC and the CIs of these parameters computed in conventional data analysis utilizing the Levenberg-Marquardt algorithm. In this study, we conducted ITC experiments using a MicroCal Auto-ITC200 instrument to investigate the binding interactions of two molecular pairs: Theo2201 aptamer–theophylline and DA-3BP aptamer–dopamine. An additional application of the ACI-ITC tool to retrospective ITC data is provided in the Supporting Information (Note D2.3)

For the Theo2201 aptamer–theophylline system, the nominal initial aptamer concentration in the sample cell ($[\text{Theo2201}]_0$, nominal L_0) was 60 μM, while the nominal initial theophylline concentration in the syringe ($[\text{theophylline}]_0$, nominal T_0) was 936 μM. The high purity of both the aptamer and theophylline, with no detectable impurities, was confirmed by liquid chromatography–mass spectrometry (LC–MS) analysis (Notes D2.4.1 and D2.4.2). Using a spectrophotometer (NanoDrop 1000, Thermo Scientific), the relative random error associated with independently prepared theophylline solutions at 936 μM was determined to be 6% (Note D2.5.1). Since experimentally determining the batch-to-batch variability of the aptamer concentration was impractical, we assigned a minimum practical value of 5%, as discussed in a previous section, as the standard deviation (SD) of the CI for the relative systematic error in aptamer concentration when calculating the ACI of the fitted parameters, despite the relative random error from preparations made from the same stock being only 4% (Note D2.5.2). The maximum observed absolute enthalpy change ($|\Delta H_1|$) during the titration was 30.3 kcal/mol. To set a lower bound for the uncertainty in ΔH, we used $\delta(\Delta H) = 0.02 \times 30.3 \approx 0.61$ kcal/mol.

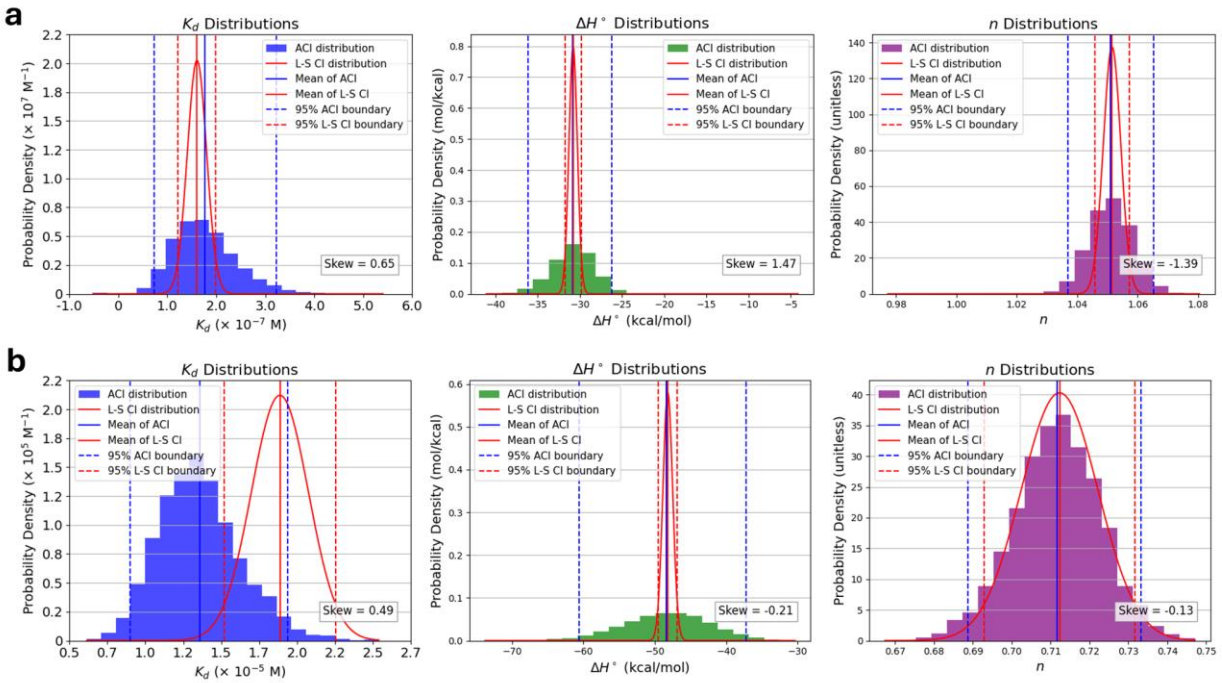


Figure 5.5. Comparison of the probability density distributions for the ACI calculated using ACI-ITC versus the CIs obtained from conventional least-squares (L-S) regression for **a)** the Theo2201 aptamer–theophylline binding pair and **b)** the DA-3BP aptamer–dopamine binding pair. ACI histograms were generated from 5,000 bootstrap iterations. Gaussian probability distributions were constructed using the mean and standard deviation values obtained from L-S nonlinear regression of the ITC binding isotherms using the Levenberg–Marquardt algorithm. The mean values and the boundaries of the 95% confidence intervals determined by both data analysis approaches are indicated on the graphs. Skewness values are reported in each panel to indicate distribution asymmetry.

The ITC experiment involved 19 injections, each of 2 µL, into a 200 µL sample cell. Using the experimental binding isotherm ("ΔH vs Molar ratio") and inputting all the values listed above into the ACI-ITC web application, we calculated the ACI for K_d, $\Delta H°$, and n, along with their frequency distributions (**Figure 5.5a**). In all ACI-ITC applications presented in this study, results were generated using 5,000 bootstrap iterations. For comparison, the 95% CIs of these

parameters were also determined using conventional least-squares (L-S) nonlinear regression with the Levenberg-Marquardt algorithm, assuming Gaussian probability distributions.

The results, summarized in **Figure 5.5a**, demonstrate that ignoring the systematic error contributions to variable uncertainties in conventional L-S analysis can significantly underestimate the uncertainties (*i.e.*, the CI widths) of ITC-derived parameters. For instance, the 95% ACI for K_d calculated by ACI-ITC was 73–320 nM, with a mean of 180 nM and a non-Gaussian distribution with a skewness of 0.65. In contrast, the CI derived from L-S analysis was much narrower, at 120–200 nM with a mean of 160 nM, incorrectly suggesting a Gaussian distribution. Understanding the correct probability distribution of the equilibrium dissociation constant K_d is essential for accurately ranking drug candidates, as binding affinity comparisons rely on this distribution to conclude which ligand binds the target tighter with a higher probability.[394] This result emphasizes the importance of incorporating systematic error considerations to ensure reliable parameter estimation in ITC analysis. Similar implications apply to the determination of the other two parameters: $\Delta H°$ and n.

To further investigate the differences between ACI and conventionally determined CIs, we applied ACI-ITC to an ITC experiment examining the DA-3BP aptamer–dopamine binding pair.[395] The nominal initial concentrations were L_0 = [DA-3BP]$_0$ = 450 μM and T_0 = [dopamine]$_0$ = 7,020 μM. The purity of the DA-3BP aptamer was determined to be 96% by LC–MS (D2.4.3), resulting in a corrected DA-3BP concentration in the ITC cell of 432 μM. Immediately after preparation, the high purity of the dopamine solution was confirmed by LC–MS (Note D2.4.4). Given the susceptibility of dopamine to oxidation, its stability over the duration of the ITC experiment was assessed by monitoring the time-dependent decrease in absorbance at 280 nm (OD_{280}) over a 90-minute period (Note D2.4.5). The results showed no significant change in

OD_{280}, indicating that dopamine oxidation was negligible under the experimental conditions and unlikely to impact the effective dopamine concentration during the measurement. This evaluation strategy can be generalized to other unstable reagents (*e.g.*, glutathione). In such cases, freshly prepared solutions should be used immediately before measurements, and degradation kinetics should be experimentally characterized, such as by monitoring absorbance at characteristic wavelengths or assessing purity over time using analytical methods like in situ LC–MS, especially when anti-degradation measures (*e.g.*, pH adjustment or stabilizing agents) are impractical. For retrospective analyses lacking experimental stability data, concentration corrections may be estimated based on published degradation kinetics.

The relative random error in dopamine concentration, based on independently prepared solutions measured by UV spectrophotometry, was determined to be 11% (Note D2.5.3). Since experimental determination of batch-to-batch variability in aptamer concentration was impractical, a conservative standard deviation of 5% was assigned for the relative systematic error in aptamer concentration when calculating the ACI, as discussed earlier. This estimate exceeds the observed 3% random variability in concentration measurements from the same stock (Note D2.5.4). Based on the maximum observed absolute heat change (ΔH_1) in the experiment, the uncertainty in heat measurement, $\delta(\Delta H)$, was set to 0.80 kcal/mol. The corrected binding isotherm was then analyzed using both ACI-ITC and conventional least-squares (L-S) regression methods, with comparative results presented in **Figure 5.5b**.

The analysis of data in **Figure 5.5b** reveals that while the ACI distribution can, in some cases, appear nearly symmetrical ($|\text{skewness}| < 0.5$), significant differences often persist between the ACI and the conventionally determined CI. For instance, in analyzing the ITC binding isotherm for the DA-3BP aptamer–dopamine pair, the ACI distribution exhibited a relatively small

positive skewness of 0.49, indicating near symmetry. The 95% confidence ACI for K_d spanned a range of 9.0–20 μM, with a width of 11 μM, 38% wider than the 95% CI determined by L-S regression, which spanned 15–23 μM with a width of 8 μM. More importantly, the mean of the ACI (14 μM) was substantially lower than that of the conventional CI (19 μM), representing a 26% difference. These results underscore a considerable divergence between the two confidence intervals, emphasizing that accounting for systematic errors in experimental variables during ITC data analysis can significantly alter the probability distribution of K_d.

Notably, while the mean values of ACI and L-S determined CI for $\Delta H°$ and n were similar for both binding pairs, the ACI were often considerably wider. In this study, we place greater emphasis on K_d, as it is a critical parameter in early-stage drug discovery, particularly for ranking drug candidates. The potential impact of inaccuracies in $\Delta H°$ and n is not extensively discussed here.

In **Figure 5.5**, it is observed that, for both examples, the uncertainties in concentrations and measured heats have a much greater impact on the determined mean K_d values than on the mean values of $\Delta H°$ and n. This phenomenon can be qualitatively explained by examining **Eqs (5.7)** and **(5.8)**. The parameter n acts as a correction factor associated with the ligand concentration (L_0), while $\Delta H°$ is derived from $\Delta H_{cum,max}$, which is constrained by the extrapolation of the binding curve. Therefore, whether using conventional L-S regression or ACI-ITC, the mean values of $\Delta H°$ and n (*i.e.*, the values close to the ones with the highest probabilities) are primarily governed by the nominal concentrations and measured heats, which have limited freedom to vary. In contrast, K_d is a parameter that is not directly tied to the concentrations or measured heats and thus retains greater flexibility to adjust during fitting. As a result, the

distribution of K_d is much more sensitive to uncertainties in these variables, and its mean value can shift significantly when uncertainties in concentrations and heats are taken into account.

By addressing both random errors (reflected in data point fluctuations) and systematic errors (such as inaccuracies in concentrations and heat measurements), ACI-ITC provides a more reliable determination of ITC-derived parameters. **Figure 5.5** highlights the significant consequences of disregarding systematic errors in analyte concentrations and heat changes, which can lead to misjudged confidence intervals for K_d. Such oversights may result in either overestimation or underestimation of K_d, or misinterpretation of its frequency (probability) distribution. In the context of drug-lead selection, these inaccuracies can disrupt the affinity ranking of drug candidates, potentially leading to the preference of weaker binders while overlooking stronger binders — decisions that could result in costly mistakes.

To compare two widened K_d probability distributions determined by ACI-ITC, it is useful to first transform the K_d values to a logarithmic scale. This transformation makes the distributions more symmetrical and facilitates direct comparison.[295] After transformation, a probabilistic comparison can be performed by quantifying the likelihood that one binder exhibits a lower/higher K_d than the other, based on their respective distributions. This approach allows for a more meaningful assessment than relying solely on overlapping confidence intervals, particularly when the distributions are broad and partially overlapping. A detailed development of statistical comparison tools for such non-Gaussian distributions is beyond the scope of the current study and will be addressed in future work.

Overall, the findings in this work highlight the importance of assessing ACI for ITC-derived parameters. Relying solely on conventional least-squares nonlinear regression, which neglects

systematic errors, may lead to misleading conclusions. To ensure accurate and reliable interpretations, ITC-derived parameters should be evaluated within the framework of ACI.

5.2.4. Concluding Remarks

To enhance the reliability of ITC-derived parameters, we developed a Monte Carlo-based program, ACI-ITC, which calculates ACI for key parameters commonly determined by ITC: K_d, $\Delta H°$, and n. To ensure accessibility and ease of use for all ITC practitioners, we integrated ACI-ITC into the ACI-Kd web app (http://aci.sci.yorku.ca), creating a browser-based tool.

This study provides practical guidance on determining the critical inputs required for calculating the ACI of ITC-derived parameters. These inputs include the CIs of systematic errors in analyte concentrations and measured heat changes. The CIs for analyte concentrations can be estimated by analyzing spectroscopic signals of the target and ligand across multiple independently prepared samples, following minimization of systematic errors, *e.g.*, correcting stock concentrations based on measured purities. This approach is appropriate because ITC experiments typically employ a single concentrated target solution and a single ligand solution throughout the titration. In cases where independent sample preparation is not feasible, for example, when only a single stock solution or batch (*e.g.*, a DNA aptamer batch from a manufacturer) is available, the CI for the reagent concentration can be estimated by propagating the known systematic uncertainties of the instruments used in sample preparation, combined with a reasonable assumption for batch-to-batch variability.

The CI for systematic error in measured heat changes is primarily influenced by the accuracy of baseline adjustment, as a single baseline adjustment is applied to the entire dataset in an ITC titration. Although a rigorously established method for assessing the CI of baseline offsets is currently lacking, reasonable estimates can be derived based on the maximum heat pulse

observed during the titration experiment. These estimates serve as a practical approach to account for baseline-related systematic errors in ACI calculations.

By evaluating the sources of systematic errors in analyte concentrations and detected heat changes, we estimated that the minimum SD input for the CIs of relative systematic errors in analyte concentrations should be set at 5%, while the minimum SD for the CIs of relative systematic errors in heat changes should be set at 2% of the maximum absolute value of the measured heat pulse. These values provide essential inputs for assessing narrower limits of the ACI of ITC-derived parameters, particularly when analyzing retrospective data. They are set as default values in the ACI-ITC web tool.

The ACI-ITC web tool was rigorously validated using synthetic ITC data and subsequently applied to real experimental datasets to demonstrate its practical utility. The experimental demonstration also illustrates the procedures for minimizing systematic errors and for estimating the CIs of the remaining systematic errors in reagent concentrations. The results reveal that the conventional least-squares fitting of ITC binding isotherms, which neglects systematic errors in variables, can underestimate the uncertainties of ITC-derived parameters and misrepresent their probability distributions. While the current work suggests initial strategies for comparing the widened ACI, further studies are needed to develop comprehensive ranking methods.

Overall, the findings in this work highlight the crucial importance of assessing ACI to ensure accurate and reliable interpretation of ITC-derived parameters. By enabling the incorporation of systematic error considerations, the browser-based tool of ACI-ITC substantially improves the reliability of ITC data analysis and strengthens the conclusions drawn from these analyses. Its application is expected to greatly enhance the accuracy and confidence in ITC-derived parameter estimations across a wide range of experimental conditions.

Chapter 6. Conclusions and Future Directions

6.1. Conclusions

Affinity interactions play a foundational role in essential processes in biology, medicine, and technology, with the equilibrium dissociation constant (K_d) serving as the key quantitative measure of the stability of the resulting complexes. Despite its critical importance, accurately determining K_d remains challenging due to methodological limitations, the absence of systematic studies on the fundamental sources of inaccuracy, the difficulty of defining systematic errors in experimental variables (*e.g.*, analyte concentrations and experimentally detected signals), and the lack of practical tools for evaluating the reliability of results. My PhD research represents a comprehensive effort to address these challenges by advancing both experimental methodologies and theoretical frameworks to enhance the accuracy of experimentally determined K_d values.

My research began by confirming the robustness and ruggedness of the instrumentation used in an inherently accurate K_d-determination method, Accurate Constant *via* Transient Incomplete Separation (ACTIS) (**Section 2.1**). We then successfully expanded the method's applicability beyond its initial limitation to systems with large differences in diffusion coefficients between the unbound ligand and the target–ligand complex (**Section 2.2**). Notably, we demonstrated that ACTIS can accurately determine K_d even for complexes with minimal diffusivity contrast, such as protein–DNA and protein–protein interactions, thereby significantly broadening its scope and establishing it as a promising reference method for diverse molecular systems.

Then, in **Section 3.1**, we conducted the first systematic theoretical investigation into the fundamental determinants of K_d accuracy, independent of experimental technique. Using analytical error-propagation analysis, we developed a universal framework that links K_d accuracy to two key factors: the ratio of the true ligand concentration to the true K_d (L_0/K_d), and

systematic errors in reagent concentrations and in the experimentally determined fraction of unbound ligand (R value). This analysis revealed a triphasic relationship between the relative systematic error in K_d ($|\Delta K_d/K_d|$) and L_0/K_d. At low L_0/K_d values, $|\Delta K_d/K_d|$ remains minimal and largely insensitive to changes in ligand concentration. As L_0/K_d increases, $|\Delta K_d/K_d|$ becomes highly sensitive and rises rapidly with increasing L_0. These two regimes are separated by a nonlinear transition zone. These insights underscore that improving K_d accuracy requires not only minimizing systematic errors in experimental variables but also employing the lowest feasible ligand concentration, constrained by the instrument's limit of quantitation (LOQ). This theoretical framework offers practical guidance for optimizing the design of non-titration equilibrium K_d-determination experiments. To further support experimental accuracy, in **Section 3.2**, we systematically investigated common sources of systematic error in K_d-determination variables, such as analyte concentrations and measured signals. We numerically evaluated and tabulated their impacts on K_d accuracy under various conditions. More importantly, we comprehensively summarized strategies for mitigating these errors. Together, these findings serve as a practical guide for improving K_d accuracy in all non-titration equilibrium binding assays.

To enable rigorous and accessible assessment of parameter accuracy, in **Section 4.1**, we introduced the Accuracy Confidence Interval (ACI) as a new statistical descriptor that complements but is distinct from the traditional Precision Confidence Interval (PCI). We developed the ACI-Kd algorithm and a corresponding web application, which allows researchers to estimate the ACI of K_d from a single binding isotherm. This tool represents a methodological advancement in the evaluation and reporting of accuracy in molecular interaction studies.

Leveraging mathematical parallels between equilibrium binding and enzyme kinetics, we extended the ACI framework to the Michaelis constant (K_m) (**Section 4.2**), providing the same level of statistical rigor to enzymology through the ACI-Km tool. In **Chapter 5**, we further adapted the ACI framework to isothermal titration calorimetry (ITC), where conventional parameter estimates frequently obscure substantial inaccuracies due to high ligand concentrations (explained in **Section 5.1**). Implemented through Monte Carlo simulation (**Section 5.2**), the ACI-ITC tool provides rigorous, error-sensitive estimates of the three key ITC-derived parameters: the equilibrium dissociation constant (K_d), the enthalpy of complex formation ($\Delta H°$), and the binding stoichiometry (n).

All three tools (ACI-Kd, ACI-Km, and ACI-ITC) are freely available through a unified, browser-based platform at https://aci.sci.yorku.ca, offering the scientific community a practical resource for enhancing the reliability and interpretability of affinity-related thermodynamic parameters.

Researchers in pharmaceutical development often assume that any systematic bias in K_d measurements will shift all candidate compounds equally, preserving their rank order of potency. However, our findings indicate that this assumption does not hold: the magnitude of systematic error in a determined K_d can vary significantly between different true K_d values. In practice, stronger-binding interactions (with lower true K_d) tend to suffer disproportionately large errors under typical assay conditions, whereas weaker binders (higher true K_d) experience smaller relative errors. This means a truly high-affinity hit may appear artificially weak (an overestimated K_d), while a moderate-affinity hit may appear comparatively stronger than it really is. Consequently, inaccurately determined K_d values can lead to wrong ranking orders of drug leads and, as a result, to disregarding high-potency hits while advancing low-potency ones. Such

rank reversals undermine the efficiency of hit triage and lead optimization in drug discovery, potentially causing the pharmaceutical industry to miss its best candidates.

It is sometimes argued in the industry that absolute accuracy in binding constants is not critical, on the reasoning that *in vivo* complexities (protein binding, metabolism, cell permeability, *etc*.) can overshadow *in vitro* affinity differences, and that ultimately a drug's value is proven in clinical trials. While it is true that a compound's *in vivo* efficacy depends on many factors beyond K_d, dismissing K_d accuracy is short-sighted. Early screening data inform which leads advance, and overestimated K_d values (*i.e*., underestimating affinity) can produce false-negative outcomes — resulting in the overlooking of promising leads in pharmaceutical development. In other words, a potent drug candidate could be erroneously dropped because an inaccurate experiment made it seem to bind weakly. Conversely, underestimating a K_d (overstating affinity) can promote a less-effective compound (a false positive), wasting resources until later tests correct the mistake. Moreover, selecting a lead with an erroneously high K_d (weaker binder) in place of a truly stronger binder means a higher dose of the drug will be required to achieve therapeutic effect. Higher required doses narrow the therapeutic window and increase the risk of off-target interactions and side effects — a direct consequence of advancing a suboptimal lead. Therefore, improving the accuracy of K_d measurements is not just an academic exercise; it has tangible benefits for drug discovery by ensuring that the best candidates are recognized and pursued, ultimately de-risking the development pipeline.

It is also important to consider how our focus on K_d accuracy relates to the potency metrics commonly used in pharmaceutical research, notably the half-maximal inhibitory concentration IC_{50}. In drug discovery assays, IC_{50} is the concentration of a compound required to inhibit a biological activity by 50%, and it is a key readout of compound potency. Generally, there is a

positive correlation between binding affinity and potency: a drug with a lower K_d (tighter binding) often yields a lower IC_{50}, meaning it achieves the same effect at a smaller dose. However, IC_{50} is an operational parameter dependent on the conditions of the assay. Factors such as ligand concentration, cell context, or pathway feedback can cause two compounds with the same true K_d to exhibit very different IC_{50} values. Indeed, it is a common mistake to assume that a lower IC_{50} always signifies stronger binding, in reality, IC_{50} and K_d are not interchangeable. One compound might appear to have a lower IC_{50} than another due to assay-specific effects rather than intrinsic affinity. This is why accurate K_d determination remains crucial: K_d is the fundamental thermodynamic descriptor of binding affinity, free from the confounding influences that affect functional assays. By obtaining reliable K_d values, researchers gain a stable point of reference for a drug's binding strength, which in turn aids in interpreting IC_{50} data correctly. Ultimately, the advances from my work, from minimizing systematic errors in binding measurements to providing tools like ACI-Kd for assessing accuracy, equip the pharmaceutical community to trust the rank ordering of drug hits based on affinity. This ensures that lead selection is grounded in true molecular performance rather than artifacts, thereby improving the chances that the most potent, selective compounds are advanced for further development. Although drug candidates must ultimately demonstrate efficacy in complex biological systems, beginning with the most accurate affinity measurements strengthens the entire drug discovery process. It reduces the risk of overlooking high-quality leads, enhances the reliability of structure–activity relationship analyses, and helps prevent costly late-stage failures by prioritizing compounds with a true mechanistic advantage in binding performance.

In summary, this work addresses a long-standing gap in the accurate quantification of molecular affinities by offering both conceptual and practical solutions that are already

influencing research practices. By integrating theoretical insights, experimental innovations, and accessible digital tools, we establish a robust foundation for maximizing and assessing the accuracy of commonly determined physicochemical parameters in molecular binding studies. The tools and frameworks introduced here are poised to elevate the rigor and credibility of binding analyses, ultimately enabling more reliable discoveries in biophysics, biomedicine, and fundamental biological research.

6.2. Future Directions

While this body of work lays a strong foundation, several key directions remain for future research:

1. **ACTIS-MS:** Although ACTIS with fluorescence detection has been demonstrated to be both rugged and robust,[177] the required fluorescent labelling of the ligand may still perturb the native binding affinity between the ligand and its target. To advance ACTIS as a truly label- and immobilization-free technique, preliminary studies have explored its integration with mass spectrometry (ACTIS-MS).[171] However, this approach remains in its early stages. Future research should include systematic investigations to further assess the robustness and reliability of ACTIS-MS. Additionally, its applicability across diverse classes of binding pairs warrants thorough evaluation.
2. **Automation of ACTIS:** At present, the ACTIS setup consists of multiple discrete components dispersed across the laboratory bench, occupying considerable space and requiring frequent manual intervention and close monitoring by the experimentalist.[177] Additionally, several separate software platforms must be coordinated to operate the system and acquire data, introducing inefficiencies and increasing the risk of human and instrumental error. To enhance usability, reliability, and adoption, future development must

focus on integrating all hardware components into a compact, unified instrument, controlled by a single, automated software interface. Such advancements will significantly improve the operational efficiency of ACTIS, reduce user burden, and establish a strong foundation for its broader adoption and eventual commercialization.

3. **Expansion of the Study Toward More Complex Binding Models:** Future research will extend the scope of this work to more complex binding models beyond the 1:1 interaction studied here. Many biologically relevant systems involve 1:*n* stoichiometries, cooperative binding, or sequential multi-site interactions, where the binding of one ligand influences the affinity of subsequent binding events. These models present additional challenges in both experimental design and data analysis, as they require more sophisticated mathematical treatment, increased parameterization, and often higher-quality or multi-dimensional datasets to ensure reliable interpretation. Moving forward, we aim to develop and validate analytical strategies capable of accurately handling such complexity. This will involve refining model selection criteria, integrating global fitting routines that accommodate multiple equilibria, and employing orthogonal techniques, such as CE, ITC, and native MS, to cross-validate stoichiometry and cooperativity. By advancing the methodological framework established in this dissertation, future work will enable more accurate and comprehensive characterization of complex binding systems, which are critical for understanding biological regulation and for the rational design of multi-valent therapeutics.
4. **Development of an Approach for Assessing the Confidence Interval of the *R* Value:** In this study, we introduced practical methods for estimating the confidence intervals (CIs) of the systematic errors in analyte concentrations.[182] However, a reliable approach for assessing the CI of systematic errors in *R* values remains underdeveloped. Future research is

needed to establish a comprehensive and universal framework for evaluating the CI of systematic errors in R, which can then be modified and adapted to method-specific contexts. Once such an approach is available, it will be incorporated into all ACI-calculation tools, enabling the systematic uncertainty in R to be fully considered in binding parameter determination.

5. **Advanced Model Fitting and Statistical Methodology for ACI:** In this study, the current ACI-calculation tools utilize the standard binding isotherm equation as the fitting model.[10] However, this model may not be optimal for certain techniques, such as isothermal titration calorimetry (ITC), where more suitable and thermodynamically consistent equations are available and should be adopted. Additionally, our current statistical approach relies on basic Monte Carlo methods, specifically bootstrapping, to estimate the ACI of ITC-derived parameters. While effective, more advanced statistical frameworks, such as Bayesian inference, offer greater flexibility and deeper insight into parameter uncertainty. Future research should focus on implementing more appropriate ITC-specific models and integrating advanced statistical tools into ACI computation. A systematic evaluation of different modeling and statistical strategies is essential to identify the most robust and reliable approaches for ACI calculation across diverse experimental platforms.
6. **Expansion of ACI Concept to Kinetic Methods and other Related Quantitative Parameters:** The current ACI tools have been successfully applied to equilibrium-based methods for estimating equilibrium or pseudo-equilibrium parameters such as K_d and K_m. However, the ACI framework is not intrinsically limited to these applications. Its core principles can be extended to a broader range of techniques, including kinetic methods such as surface plasmon resonance (SPR) and bio-layer interferometry (BLI). Likewise, the scope

of analysable parameters can be expanded to include key quantitative metrics in molecular and biochemical studies, such as kinetic rate constants (*e.g.*, k_{on} and k_{off}) and Hill coefficients (cooperativity). Applying ACI to these techniques and parameters would establish a unified and systematic framework for assessing the accuracy and reliability of experimentally derived values, particularly under conditions involving systematic uncertainties. This expansion will enhance the interpretability of kinetic and pharmacological data while promoting more rigorous, uncertainty-aware analysis across diverse disciplines, including enzymology, drug discovery, and biophysical sensing technologies.

List of Publications

12. **Wang, T.Y.**; Bijlani, A.; Chao, E. H. P.; Johnson, P.E.; Krylov, S.N. A Browser-Based Tool for Assessing Accuracy of ITC-Derived Parameters: K_d, $\Delta H°$, and *n.* **2025**, *ChemBioChem, 2500194,* published online.

11. **Wang, T. Y.**; Krylov, S. N. Deterministic Error Propagation in ITC: Revealing Multi-Fold Errors in K_d Values Under Standard Conditions. *Biophysical Chemistry* **2025**, *323*, 107455.

10. **Wang, T. Y.**; Dhillon, P.; Krylova, S. M.; Golemi-Kotra, D.; Kohlmann, P.; Krylov, S. N. Introducing Quantitative Assessment of Michaelis Constant (K_m) Accuracy. 2025, *submitted.*

9. **Wang, T.Y.**; Latimer, J.; Rukundo, J.-L.; Kogan, I.; Krylova, S.M.; Schreiber, S.; Kohlmann, P.; Jose, J.; Krylov, S.N. A Practical Approach to Quantitatively Assessing Equilibrium-Constant Accuracy from a Single Binding Isotherm. *Precision Chemistry* **2025,** *3* (2), 89–104.

8. **Wang, T.Y.**; Rukundo, J.-L.; Mao, Z.; Krylov, S.N. Maximizing the Accuracy of Equilibrium Dissociation Constants for Affinity Complexes: From Theory to Practical Recommendations. *ACS Chemical Biology* **2024**, *19* (9), 1852–1867.

7. **Wang, T.Y.**; Ji, W.; Evetron, D.; Le, A.T.H.; Krylova, S.M.; Fournier, R.; Krylov, S.N. Fundamental Determinants of the Accuracy of Equilibrium Constants for Affinity Complexes. *Analytical Chemistry* **2023**, *95*(42), 15826–15832.

6. Teclemichael, E.; Le, A.T.H.; Krylova, S.M.; **Wang, T.Y.**; Krylov, S.N. Bulk Affinity Assays in Aptamer Selection: Challenges, Theory, and Workflow. *Analytical Chemistry* **2022**, *94*(44)*,* 15183–15188.

5. **Wang, T.Y.**; Ivanov, N.A.; Rukundo, J.-L.; Le Blanc, J.C.Y.; Gorin, B.I. Krylov, S.N. Transient Incomplete Separation of Species with Close Diffusivity to Study Stability of Affinity Complexes. *Analytical Chemistry* **2022**, *94*(44), 15415–15422.

4. Le, A.T.H.; **Wang, T.Y.**; Krylova, S.M.; Beloborodov, S.S.; Krylov, S.N. Quantitative Characterization of Partitioning in Selection of DNA Aptamers for Protein Targets by Capillary Electrophoresis. *Analytical Chemistry* **2022**, *94*(5), 2578–2588.

3. Rukundo, J.-L.; Kochmann, S.; **Wang, T.Y.**; Ivanov, N.A.; Le Blanc, J.C.Y.; Gorin, B.I. Krylov, S.N. Template Instrumentation for "Accurate Constant *via* Transient Incomplete Separation" (ACTIS). *Analytical Chemistry* **2021**, *93*(34), 11654–11659.

2. Le, A.T.H.; Krylova, S.M.; Beloborodov, S.S.; **Wang, T.Y.**; Hili, R.; Johnson, P.E.; Li, F.; Veedu, R.N.; Belyanskaya, S.; Krylov, S.N. How to Develop and Prove High-Efficiency Selection of Ligands from Oligonucleotide Libraries: A Universal Framework for Aptamers and DNA-Encoded Small-Molecule Ligands. *Analytical Chemistry* **2021**, *93*(13), 5343–5354.

1. **Wang, T.Y.**; Hu, L.; Krylov, S.N. Empirical Predictor of Conditions which Support Ideal-Filter Capillary Electrophoresis. *Electrophoresis* **2020**, *41*, 1225–1229.

Bibliography

(1) Sposito, G. *Chemical Equilibria and Kinetics in Soils*; Oxford University Press, 1994.

(2) Martell, A. E.; Smith, R. M. *Critical Stability Constants*; Springer, 1974.

(3) Janoš, P. Determination of equilibrium constants from chromatographic and electrophoretic measurements. *Journal of Chromatography A* **2004**, *1037* (1-2), 15–28.

(4) Lew, M. J.; Angus, J. A. Analysis of competitive agonist-antagonist interactions by nonlinear regression. *Trends in Pharmacological Sciences* **1995**, *16* (10), 328–337.

(5) Johnson, K. A. A century of enzyme kinetic analysis, 1913 to 2013. *FEBS letters* **2013**, *587* (17), 2753–2766.

(6) Okpara, O. G.; Ogbeide, O. M.; Ike, O. C.; Menechukwu, K. C.; Ejike, E. C. Optimum isotherm by linear and nonlinear regression methods for lead (II) ions adsorption from aqueous solutions using synthesized coconut shell–activated carbon (SCSAC). *Toxin Reviews* **2021**, *40* (4), 901–914.

(7) Lednev, I. K.; Hester, R. E.; Moore, J. N. A cation-specific, light-controlled transient chromoionophore based on a benzothiazolium styryl azacrown ether dye. *Journal of the American Chemical Society* **1997**, *119* (15), 3456–3461.

(8) Shashilov, V. A.; Ermolenkov, V. V.; Lednev, I. K. Multiple Bicyclic Diamide− Lutetium Complexes in Solution: Chemometric Analysis of Deep-UV Raman Spectroscopic Data. *Inorganic Chemistry* **2006**, *45* (9), 3606–3612.

(9) Alongi, K. S.; Shields, G. C. Theoretical calculations of acid dissociation constants: A review article. *Annual Reports in Computational Chemistry* **2010**, *6*, 113–138.

(10) Wang, T. Y.; Latimer, J.; Rukundo, J. L.; Kogan, I.; Krylova, S. M.; Schreiber, S.; Kohlmann, P.; Jose, J.; Krylov, S. N. A Practical Approach to Quantitatively Assessing Equilibrium-Constant Accuracy from a Single Binding Isotherm. *Precision Chemistry* **2025**, *3* (2), 89–104.

(11) Krylov, S. N. Underestimation of the Complexity of K_d Determination: Causes, Implications, and Ways to Improve. *ACS Measurement Science Au* **2024**, *4*(3), 231–232.

(12) Ma, W.; Yang, L.; He, L. Overview of the detection methods for equilibrium dissociation constant K_D of drug-receptor interaction. *Journal of Pharmaceutical Analysis* **2018**, *8* (3), 147–152.

(13) Baird, J. K. A generalized statement of the law of mass action. *Journal of Chemical Education* **1999**, *76* (8), 1146.

(14) Dais, P. Impact of Gibbs' and Duhem's approaches to thermodynamics on the development of chemical thermodynamics. *Archive for History of Exact Sciences* **2021**, *75*, 175–248.

(15) Pathare, B.; Tambe, V.; Patil, V. A review on various analytical methods used in determination of dissociation constant. *International Journal of Pharmacy and Pharmaceutical Sciences* **2014**, *6* (8), 26–34.

(16) Rosenbaum, M. I.; Clemmensen, L. S.; Bredt, D. S.; Bettler, B.; Strømgaard, K. Targeting receptor complexes: a new dimension in drug discovery. *Nature Reviews Drug Discovery* **2020**, *19* (12), 884–901.

(17) Luttrell, L. M. Minireview: More than just a hammer: ligand "bias" and pharmaceutical discovery. *Molecular Endocrinology* **2014**, *28* (3), 281–294.

(18) Deng, G.; Sanyal, G. Applications of mass spectrometry in early stages of target based drug discovery. *Journal of Pharmaceutical and Biomedical Analysis* **2006**, *40* (3), 528–538.

(19) Nunez, S.; Venhorst, J.; Kruse, C. G. Target–drug interactions: first principles and their application to drug discovery. *Drug Discovery Today* **2012**, *17* (1-2), 10–22.

(20) Jing, M.; Bowser, M. T. Methods for measuring aptamer-protein equilibria: a review. *Analytica Chimica Acta* **2011**, *686* (1-2), 9–18.

(21) Crivianu-Gaita, V.; Thompson, M. Aptamers, antibody scFv, and antibody Fab'fragments: An overview and comparison of three of the most versatile biosensor biorecognition elements. *Biosensors and Bioelectronics* **2016**, *85*, 32–45.

(22) Zhou, R.; Chen, S.; Wu, Q.; Liu, L.; Wang, Y.; Mo, Y.; Zeng, Z.; Zu, X.; Xiong, W.; Wang, F. CD155 and its receptors in cancer immune escape and immunotherapy. *Cancer Letters* **2023**, *573*, 216381.

(23) Lu, Y.; Sega, E.; Leamon, C. P.; Low, P. S. Folate receptor-targeted immunotherapy of cancer: mechanism and therapeutic potential. *Advanced Drug Delivery Reviews* **2004**, *56* (8), 1161–1176.

(24) Eissa, S.; Ng, A.; Siaj, M.; Tavares, A. C.; Zourob, M. Selection and identification of DNA aptamers against okadaic acid for biosensing application. *Analytical Chemistry* **2013**, *85* (24), 11794–11801.

(25) Sudhaharan, T.; Liu, P.; Foo, Y. H.; Bu, W.; Lim, K. B.; Wohland, T.; Ahmed, S. Determination of in vivo dissociation constant, K_D, of Cdc42-effector complexes in live mammalian cells using single wavelength fluorescence cross-correlation spectroscopy. *Journal of Biological Chemistry* **2009**, *284* (20), 13602–13609.

(26) Pérez, I. On the experimental determination of the one-way speed of light. *European Journal of Physics* **2011**, *32* (4), 993.

(27) Guercio, G.; Zanetti, V. Determination of gravitational acceleration using a rubber ball. *American Journal of Physics* **1987**, *55* (1), 59–63.

(28) van Marken Lichtenbelt, W. D.; Daanen, H. A.; Wouters, L.; Fronczek, R.; Raymann, R. J.; Severens, N. M.; Van Someren, E. J. Evaluation of wireless determination of skin temperature using iButtons. *Physiology & Behavior* **2006**, *88* (4-5), 489–497.

(29) Bolt, G. Determination of the charge density of silica sols. *The Journal of Physical Chemistry* **1957**, *61* (9), 1166–1169.

(30) Nagai, H.; Kuwabara, K.; Carta, G. Temperature dependence of the dissociation constants of several amino acids. *Journal of Chemical & Engineering Data* **2008**, *53* (3), 619–627.

(31) Roy, R. N.; Gibbons, J. J.; Baker, G. E. Acid dissociation constants and pH values for standard “bes” and “tricine” buffer solutions in 30, 40, and 50 mass% dimethyl sulfoxide/water between 25 and− 25° C. *Cryobiology* **1985**, *22* (6), 589–600.

(32) Bottari, F.; Daems, E.; de Vries, A.-M.; Van Wielendaele, P.; Trashin, S.; Blust, R.; Sobott, F.; Madder, A.; Martins, J. C.; De Wael, K. Do aptamers always bind? The need for a multifaceted analytical approach when demonstrating binding affinity between aptamer and low molecular weight compounds. *Journal of the American Chemical Society* **2020**, *142* (46), 19622–19630.

(33) Fielding, L. NMR methods for the determination of protein–ligand dissociation constants. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2007**, *51* (4), 219–242.

(34) Jarmoskaite, I.; AlSadhan, I.; Vaidyanathan, P. P.; Herschlag, D. How to measure and evaluate binding affinities. *Elife* **2020**, *9*, e57264.

(35) Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Current Topics in Medicinal Chemistry* **2003**, *3* (1), 39–53.

(36) Stangherlin, S.; Ding, Y.; Liu, J. Dissociation Constant (Kd) Measurement for Small-Molecule Binding Aptamers: Homogeneous Assay Methods and Critical Evaluations. *Small Methods* **2024**, 2401572.

(37) Sun, Y.-S.; Landry, J. P.; Fei, Y.; Zhu, X.; Luo, J.; Wang, X.; Lam, K. Effect of fluorescently labeling protein probes on kinetics of protein−ligand reactions. *Langmuir* **2008**, *24* (23), 13399–13405.

(38) Nimse, S. B.; Song, K.; Sonawane, M. D.; Sayyed, D. R.; Kim, T. Immobilization techniques for microarray: challenges and applications. *Sensors* **2014**, *14* (12), 22208–22229.

(39) Chen, Y.; Ming, H. Review of surface plasmon resonance and localized surface plasmon resonance sensor. *Photonic Sensors* **2012**, *2*, 37–49.

(40) Homola, J. Present and future of surface plasmon resonance biosensors. *Analytical and Bioanalytical Chemistry* **2003**, *377* (3), 528–539.

(41) Schneider, C. S.; Bhargav, A. G.; Perez, J. G.; Wadajkar, A. S.; Winkles, J. A.; Woodworth, G. F.; Kim, A. J. Surface plasmon resonance as a high throughput method to evaluate specific and non-specific binding of nanotherapeutics. *Journal of Controlled Release* **2015**, *219*, 331–344.

(42) Mullett, W. M.; Lai, E. P.; Yeung, J. M. Surface plasmon resonance-based immunoassays. *Methods* **2000**, *22* (1), 77–91.

(43) Lin, S.; Shih-Yuan Lee, A.; Lin, C.-C.; Lee, C.-K. Determination of binding constant and stoichiometry for antibody-antigen interaction with surface plasmon resonance. *Current Proteomics* **2006**, *3* (4), 271–282.

(44) Mrksich, M.; Whitesides, G. M. Using self-assembled monolayers that present oligo (ethylene glycol) groups to control the interactions of proteins with surfaces. In *ACS Symposium Series*, 1997; American Chemical Society: Vol. 680, pp 361–373.

(45) Pasche, S.; Textor, M.; Meagher, L.; Spencer, N. D.; Griesser, H. J. Relationship between interfacial forces measured by colloid-probe atomic force microscopy and protein resistance of poly (ethylene glycol)-grafted poly (L-lysine) adlayers on niobia surfaces. *Langmuir* **2005**, *21* (14), 6508–6520.

(46) Ye, H.; Che, J.; Huang, R.; Qi, W.; He, Z.; Su, R. Zwitterionic peptide enhances protein-resistant performance of hyaluronic acid-modified surfaces. *Langmuir* **2020**, *36* (8), 1923–1929.

(47) Schuck, P.; Zhao, H. The role of mass transport limitation and surface heterogeneity in the biophysical characterization of macromolecular binding processes by SPR biosensing. *Surface Plasmon Resonance: Methods and Protocols* **2010**, 15–54.

(48) Mayer, K. M.; Hafner, J. H. Localized surface plasmon resonance sensors. *Chemical Reviews* **2011**, *111* (6), 3828–3857.

(49) Concepcion, J.; Witte, K.; Wartchow, C.; Choo, S.; Yao, D.; Persson, H.; Wei, J.; Li, P.; Heidecker, B.; Ma, W. Label-free detection of biomolecular interactions using BioLayer interferometry for kinetic characterization. *Combinatorial Chemistry & High Throughput Screening* **2009**, *12* (8), 791–800.

(50) Jug, A.; Bratkovič, T.; Ilaš, J. Biolayer interferometry and its applications in drug discovery and development. *TrAC Trends in Analytical Chemistry* **2024**, 117741.

(51) Auer, S.; Koho, T.; Uusi-Kerttula, H.; Vesikari, T.; Blazevic, V.; Hytönen, V. P. Rapid and sensitive detection of norovirus antibodies in human serum with a biolayer interferometry biosensor. *Sensors and Actuators B: Chemical* **2015**, *221*, 507–514.

(52) Wear, M. A.; Walkinshaw, M. D. Thermodynamics of the cyclophilin-A/cyclosporin-A interaction: a direct comparison of parameters determined by surface plasmon resonance using Biacore T100 and isothermal titration calorimetry. *Analytical Biochemistry* **2006**, *359* (2), 285–287.

(53) Bastos, M.; Abian, O.; Johnson, C. M.; Ferreira-da-Silva, F.; Vega, S.; Jimenez-Alesanco, A.; Ortega-Alarcon, D.; Velazquez-Campoy, A. Isothermal titration calorimetry. *Nature Reviews Methods Primers* **2023**, *3* (1), 17.

(54) Johnson, C. M. Isothermal titration calorimetry. *Protein-Ligand Interactions: Methods and Applications* **2021**, 135–159.

(55) Velázquez-Campoy, A.; Ohtaka, H.; Nezami, A.; Muzammil, S.; Freire, E. Isothermal titration calorimetry. *Current Protocols in Cell Biology* **2004**, *23* (1), 17.18. 11–17.18. 24.

(56) Linkuvienė, V.; Krainer, G.; Chen, W.-Y.; Matulis, D. Isothermal titration calorimetry for drug design: precision of the enthalpy and binding constant measurements and comparison of the instruments. *Analytical Biochemistry* **2016**, *515*, 61–64.

(57) Roselin, L. S.; Lin, M. S.; Lin, P. H.; Chang, Y.; Chen, W. Y. Recent trends and some applications of isothermal titration calorimetry in biotechnology. *Biotechnology Journal: Healthcare Nutrition Technology* **2010**, *5* (1), 85–98.

(58) Liang, Y. Applications of isothermal titration calorimetry in protein science. *Acta Biochimica et Biophysica Sinica* **2008**, *40* (7), 565–576.

(59) Archer, W. R.; Schulz, M. D. Isothermal titration calorimetry: practical approaches and current applications in soft matter. *Soft Matter* **2020**, *16* (38), 8760–8774.

(60) Ward, W. H.; Holdgate, G. A. 7 Isothermal Titration Calorimetry in Drug Discovery. *Progress in Medicinal Chemistry* **2001**, *38*, 309–376.

(61) Krainer, G.; Broecker, J.; Vargas, C.; Fanghänel, J. r.; Keller, S. Quantifying high-affinity binding of hydrophobic ligands by isothermal titration calorimetry. *Analytical Chemistry* **2012**, *84* (24), 10715–10722.

(62) Torres, F. E.; Recht, M. I.; Coyle, J. E.; Bruce, R. H.; Williams, G. Higher throughput calorimetry: opportunities, approaches and challenges. *Current Opinion in Structural Biology* **2010**, *20* (5), 598–605.

(63) Harish, M.; Venkatraman, P. Evolution of biophysical tools for quantitative protein interactions and drug discovery. *Emerging Topics in Life Sciences* **2021**, *5* (1), 1–12.

(64) Wu, D.; Piszczek, G. Measuring the affinity of protein-protein interactions on a single-molecule level by mass photometry. *Analytical Biochemistry* **2020**, *592*, 113575.

(65) Jia, Y.; Zhang, Z.; Su, C.; Lin, Q. Isothermal titration calorimetry in a polymeric microdevice. *Microfluidics and Nanofluidics* **2017**, *21*, 1–9.

(66) Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D. MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure* **2014**, *1077*, 101–113.

(67) Jerabek-Willemsen, M.; Wienken, C. J.; Braun, D.; Baaske, P.; Duhr, S. Molecular interaction studies using microscale thermophoresis. *Assay and Drug Development Technologies* **2011**, *9* (4), 342–353.

(68) El Deeb, S.; Al-Harrasi, A.; Khan, A.; Al-Broumi, M.; Al-Thani, G.; Alomairi, M.; Elumalai, P.; Sayed, R. A.; Ibrahim, A. E. Microscale thermophoresis as a powerful growing analytical technique for the investigation of biomolecular interaction and the determination of binding parameters. *Methods and Applications in Fluorescence* **2022**, *10* (4), 042001.

(69) Asmari, M.; Ratih, R.; Alhazmi, H. A.; El Deeb, S. Thermophoresis for characterizing biomolecular interaction. *Methods* **2018**, *146*, 107–119.

(70) Rainard, J. M.; Pandarakalam, G. C.; McElroy, S. P. Using microscale thermophoresis to characterize hits from high-throughput screening: a European lead factory perspective. *SLAS DISCOVERY: Advancing Life Sciences R&D* **2018**, *23* (3), 225–241.

(71) Podobnik, M.; Kraševec, N.; Zavec, A. B.; Naneh, O.; Flašker, A.; Caserman, S.; Hodnik, V.; Anderluh, G. How to study protein-protein interactions. *Acta Chimica Slovenica* **2016**, *63* (3), 424–439.

(72) Nowak, J. S.; Czarna, A.; Grudnik, P.; Grygier, P.; Pustelny, K.; Langer, A.; Dubin, G. Microscale thermophoresis (MST) and spectral shift (SpS) in drug discovery. *TrAC Trends in Analytical Chemistry* **2024**, 117716.

(73) Scheuermann, T. H.; Padrick, S. B.; Gardner, K. H.; Brautigam, C. A. On the acquisition and analysis of microscale thermophoresis data. *Analytical Biochemistry* **2016**, *496*, 79–93.

(74) Griffey, R. H.; Redfield, A. G. Proton-detected heteronuclear edited and correlated nuclear magnetic resonance and nuclear Overhauser effect in solution. *Quarterly Reviews of Biophysics* **1987**, *19* (1-2), 51–82.

(75) Pellecchia, M. Solution nuclear magnetic resonance spectroscopy techniques for probing intermolecular interactions. *Chemistry & Biology* **2005**, *12* (9), 961–971.

(76) Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2013**, *73*, 1–16.

(77) Shi, L.; Zhang, N. Applications of solution NMR in drug discovery. *Molecules* **2021**, *26* (3), 576.

(78) Eills, J.; Budker, D.; Cavagnero, S.; Chekmenev, E. Y.; Elliott, S. J.; Jannin, S.; Lesage, A.; Matysik, J.; Meersmann, T.; Prisner, T. Spin hyperpolarization in modern magnetic resonance. *Chemical Reviews* **2023**, *123* (4), 1417–1551.

(79) Crook, A. A.; Powers, R. Quantitative NMR-based biomedical metabolomics: Current status and applications. *Molecules* **2020**, *25* (21), 5128.

(80) Fernández, C.; Wider, G. TROSY in NMR studies of the structure and function of large biological macromolecules. *Current Opinion in Structural Biology* **2003**, *13* (5), 570–580.

(81) Erba, E. B.; Zenobi, R. Mass spectrometric studies of dissociation constants of noncovalent complexes. *Annual Reports Section" C" Physical Chemistry* **2011**, *107*, 199–228.

(82) Bennett, J. L.; Nguyen, G. T.; Donald, W. A. Protein–small molecule interactions in native mass spectrometry. *Chemical Reviews* **2021**, *122* (8), 7327–7385.

(83) Boeri Erba, E.; Petosa, C. The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Science* **2015**, *24* (8), 1176–1192.

(84) Masson, G. R.; Jenkins, M. L.; Burke, J. E. An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opinion on Drug Discovery* **2017**, *12* (10), 981–994.

(85) Kim, M. S.; Zhong, J.; Pandey, A. Common errors in mass spectrometry-based analysis of post-translational modifications. *Proteomics* **2016**, *16* (5), 700–714.

(86) Potier, N.; Rogniaux, H.; Chevreux, G.; Van Dorsselaer, A. Ligand–metal ion binding to proteins: investigation by ESI mass spectrometry. *Methods in Enzymology* **2005**, *402*, 361–389.

(87) *Comparison of NMR and MS*. EMBL-EBI, https://www.ebi.ac.uk/training/online/courses/metabolomics-introduction/designing-a-metabolomics-study/comparison-of-nmr-and-ms/ (accessed 2025 04-03).

(88) Dubský, P.; Dvořák, M.; Ansorge, M. Affinity capillary electrophoresis: the theory of electromigration. *Analytical and Bioanalytical Chemistry* **2016**, *408*, 8623–8641.

(89) Heegaard, N. H.; Nilsson, S.; Guzman, N. A. Affinity capillary electrophoresis: important application areas and some recent developments. *Journal of Chromatography B: Biomedical Sciences and Applications* **1998**, *715* (1), 29–54.

(90) Schou, C.; Heegaard, N. H. Recent applications of affinity interactions in capillary electrophoresis. *Electrophoresis* **2006**, *27* (1), 44–59.

(91) Wang, Y.; Adeoye, D. I.; Ogunkunle, E. O.; Wei, I.-A.; Filla, R. T.; Roper, M. G. Affinity capillary electrophoresis: a critical review of the literature from 2018 to 2020. *Analytical Chemistry* **2020**, *93* (1), 295–310.

(92) Østergaard, J.; Heegaard, N. H. Capillary electrophoresis frontal analysis: principles and applications for the study of drug-plasma protein binding. *Electrophoresis* **2003**, *24* (17), 2903–2913.

(93) Vuignier, K.; Schappler, J.; Veuthey, J.-L.; Carrupt, P.-A.; Martel, S. Improvement of a capillary electrophoresis/frontal analysis (CE/FA) method for determining binding constants: discussion on relevant parameters. *Journal of Pharmaceutical And Biomedical Analysis* **2010**, *53* (5), 1288–1297.

(94) Chen, L.; Ren, J. High-throughput DNA analysis by microchip electrophoresis. *Combinatorial Chemistry & High Throughput Screening* **2004**, *7* (1), 29–43.

(95) Dolnik, V.; Liu, S. Applications of capillary electrophoresis on microchip. *Journal of Separation Science* **2005**, *28* (15), 1994–2009.

(96) Berezovski, M.; Krylov, S. N. Nonequilibrium capillary electrophoresis of equilibrium mixtures− a single experiment reveals equilibrium and kinetic parameters of protein−DNA interactions. *Journal of the American Chemical Society* **2002**, *124* (46), 13674–13675.

(97) Berezovski, M.; Nutiu, R.; Li, Y.; Krylov, S. N. Affinity analysis of a Protein−Aptamer complex using nonequilibrium capillary electrophoresis of equilibrium mixtures. *Analytical Chemistry* **2003**, *75* (6), 1382–1386.

(98) Deng, L.; Fu, Q.; Zhang, Y.; Shui, F.; Tang, J.; Wu, J.; Zeng, J. Study of molecular interactions by nonequilibrium capillary electrophoresis of equilibrium mixtures: Originations, developments, and applications. *Electrophoresis* **2023**, *44* (21-22), 1664–1673.

(99) Petrov, A.; Okhonin, V.; Berezovski, M.; Krylov, S. N. Kinetic capillary electrophoresis (KCE): a conceptual platform for kinetic homogeneous affinity methods. *Journal of the American Chemical Society* **2005**, *127* (48), 17104–17110.

(100) Okhonin, V.; Petrov, A. P.; Berezovski, M.; Krylov, S. N. Plug−Plug Kinetic Capillary Electrophoresis: Method for Direct Determination of Rate Constants of Complex Formation and Dissociation. *Analytical Chemistry* **2006**, *78* (14), 4803–4810.

(101) Kanoatov, M.; Galievsky, V. A.; Krylova, S. M.; Cherney, L. T.; Jankowski, H. K.; Krylov, S. N. Using nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) for simultaneous determination of concentration and equilibrium constant. *Analytical Chemistry* **2015**, *87* (5), 3099–3106.

(102) Tanaka, Y.; Terabe, S. Estimation of binding constants by capillary electrophoresis. *Journal of Chromatography B* **2002**, *768* (1), 81–92.

(103) Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T. Quantitative Metabolome Analysis Using Capillary Electrophoresis Mass Spectrometry. *Journal of Proteome Research* **2003**, *2* (5), 488–494.

(104) Wang, X.; Yi, L.; Mukhitov, N.; Schrell, A. M.; Dhumpa, R.; Roper, M. G. Microfluidics-to-mass spectrometry: A review of coupling methods and applications. *Journal of Chromatography A* **2015**, *1382*, 98–116.

(105) FUCHS, H.; GESSNER, R. The result of equilibrium-constant calculations strongly depends on the evaluation method used and on the type of experimental errors. *Biochemical Journal* **2001**, *359* (2), 411–418.

(106) Tellinghuisen, J.; Chodera, J. D. Systematic errors in isothermal titration calorimetry: concentrations and baselines. *Analytical biochemistry* **2011**, *414* (2), 297–299.

(107) Knapp, G.; Schramel, P. Sources of analyte contamination and loss during the analytical process. *Comprehensive Analytical Chemistry* **2003**, *41*, 23–45.

(108) Hibbert, D. B. Systematic errors in analytical measurement results. *Journal of Chromatography A* **2007**, *1158* (1-2), 25–32.

(109) Blum, F. High performance liquid chromatography. *British Journal of Hospital Medicine (2005)* **2014**, *75* (Sup2), C18–C21.

(110) Swartz, M. HPLC detectors: a brief review. *Journal of Liquid Chromatography & Related Technologies* **2010**, *33* (9-12), 1130–1150.

(111) Holm, R.; Elder, D. P. Analytical advances in pharmaceutical impurity profiling. *European Journal of Pharmaceutical Sciences* **2016**, *87*, 118–135.

(112) Geng, X.; Wang, L. Liquid chromatography of recombinant proteins and protein drugs. *Journal of chromatography B* **2008**, *866* (1-2), 133–153.

(113) Guiochon, G. Monolithic columns in high-performance liquid chromatography. *Journal of Chromatography A* **2007**, *1168* (1–2), 101–168.

(114) Yabré, M.; Ferey, L.; Somé, I. T.; Gaudin, K. Greening reversed-phase liquid chromatography methods using alternative solvents for pharmaceutical analysis. *Molecules* **2018**, *23* (5), 1065.

(115) Staub, A.; Guillarme, D.; Schappler, J.; Veuthey, J.-L.; Rudaz, S. Intact protein analysis in the biopharmaceutical field. *Journal of Pharmaceutical and Biomedical Analysis* **2011**, *55* (4), 810–822.

(116) Sasse, J.; Gallagher, S. R. Staining proteins in gels. *Current Protocols in Molecular Biology* **2003**, *63* (1), 10.16. 11–10.16. 25.

(117) Bischoff, K. M.; Shi, L.; Kennelly, P. J. The detection of enzyme activity following sodium dodecyl sulfate–polyacrylamide gel electrophoresis. *Analytical Biochemistry* **1998**, *260* (1), 1–17.

(118) Gräslund, S.; Nordlund, P.; Weigelt, J.; Hallberg, B. M.; Bray, J.; Gileadi, O.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; et al. Protein production and purification. *Nature Methods* **2008**, *5* (2), 135–146.

(119) Hamdan, M.; Righetti, P. G. Modern strategies for protein quantification in proteome analysis: advantages and limitations. *Mass Spectrometry Reviews* **2002**, *21* (4), 287–302.

(120) Gallagher, S. R. One-dimensional SDS gel electrophoresis of proteins. *Current Protocols in Immunology* **2006**, *75* (1), 8.4. 1–8.4. 37.

(121) Singhal, A.; Saini, U.; Chopra, B.; Dhingra, A. K.; Jain, A.; Chaudhary, J. Uv-visible spectroscopy: A review on its pharmaceutical and bio-allied sciences applications. *Current Pharmaceutical Analysis* **2024**, *20* (3), 161–177.

(122) Toptygin, D. Effects of the solvent refractive index and its dispersion on the radiative decay rate and extinction coefficient of a fluorescent solute. *Journal of Fluorescence* **2003**, *13*, 201–219.

(123) Zhou, H.-X.; Pang, X. Electrostatic interactions in protein structure, folding, binding, and condensation. *Chemical Reviews* **2018**, *118* (4), 1691–1741.

(124) Antosiewicz, J. M.; Shugar, D. UV–Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 2: Selected applications. *Biophysical Reviews* **2016**, *8*, 163–177.

(125) Gill, S. C.; Von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* **1989**, *182* (2), 319–326.

(126) Trufelli, H.; Palma, P.; Famiglini, G.; Cappiello, A. An overview of matrix effects in liquid chromatography–mass spectrometry. *Mass Spectrometry Reviews* **2011**, *30* (3), 491–509.

(127) Stutz, H. Advances in the analysis of proteins and peptides by capillary electrophoresis with matrix-assisted laser desorption/ionization and electrospray-mass spectrometry detection. *Electrophoresis* **2005**, *26* (7–8), 1254–1290.

(128) Bharti, S. K.; Roy, R. Quantitative 1H NMR spectroscopy. *TrAC Trends in Analytical Chemistry* **2012**, *35*, 5–26.

(129) Chace, D. H. Mass spectrometry in the clinical laboratory. *Chemical Reviews* **2001**, *101* (2), 445–478.

(130) Giraudeau, P. Challenges and perspectives in quantitative NMR. *Magnetic Resonance in Chemistry* **2017**, *55* (1), 61–69.

(131) Zeleny, R.; Schimmel, H. Influence of the approach to calibration on the accuracy and the traceability of certified values in certified reference materials. *TrAC Trends in Analytical Chemistry* **2012**, *33*, 107–116.

(132) Padariya, C.; Rutkowska, M.; Konieczka, P. The importance and availability of marine certified reference materials. *Critical Reviews in Environmental Science and Technology* **2022**, *52* (18), 3322–3373.

(133) *10 Steps to Improve Pipetting Accuracy*. Thermo Fisher Scientific 2024. https://www.thermofisher.com/ca/en/home/life-science/lab-plasticware-supplies/lab-plasticware-supplies-learning-center/lab-plasticware-supplies-resource-library/fundamentals-of-pipetting/proper-pipetting-techniques/10-steps-to-improve-pipetting-accuracy.html (accessed 2024 03-18).

(134) Protti, M.; Mandrioli, R.; Mercolini, L. Tutorial: Volumetric absorptive microsampling (VAMS). *Analytica Chimica Acta* **2019**, *1046*, 32–47.

(135) Saha, S. *Performing Pipette Calibration Yourself*. Bitesize Bio, 2021. https://bitesizebio.com/40766/performing-pipette-calibration-yourself/ (accessed 2024 03-18).

(136) Kant, R. Laboratory Pipetting Techniques for Accurate Results: Precision in Practice. *Laboratory Techniques*, 47.

(137) Berthier, E.; Warrick, J.; Yu, H.; Beebe, D. J. Managing evaporation for more robust microscale assays Part 1. Volume loss in high throughput assays. *Lab on a Chip* **2008**, *8* (6), 852–859.

(138) Wang, P.; Zhang, X.; Zhou, Y.; Jiang, Q.; Ye, Q.; Chu, Z.; Li, X.; Yang, X.; Yin, Z.; You, J. Solvent-controlled growth of inorganic perovskite films in dry environment for efficient and stable solar cells. *Nature Communications* **2018**, *9* (1), 2225.

(139) Smith, J.; Larmené-Beld, K. Containers. In *Practical Pharmaceutics: An International Guideline for the Preparation, Care and Use of Medicinal Products*, Springer, 2023; pp 169–206.

(140) Kazmierczak, N. P.; Chew, J. A.; Vander Griend, D. A. Bootstrap methods for quantifying the uncertainty of binding constants in the hard modeling of spectrophotometric titration data. *Analytica Chimica Acta* **2022**, *1227*, 339834.

(141) Nguyen, T. H.; Rustenburg, A. S.; Krimmer, S. G.; Zhang, H.; Clark, J. D.; Novick, P. A.; Branson, K.; Pande, V. S.; Chodera, J. D.; Minh, D. D. Bayesian analysis of isothermal titration calorimetry for binding thermodynamics. *PLoS One* **2018**, *13* (9), e0203224.

(142) Feng, F.; Kepler, T. B. Bayesian estimation of the active concentration and affinity constants using surface plasmon resonance technology. *PloS one* **2015**, *10* (6), e0130812.

(143) Peck, C. C.; Sheiner, L. B.; Nichols, A. I. The problem of choosing weights in nonlinear regression analysis of pharmacokinetic data. *Drug Metabolism Reviews* **1984**, *15* (1-2), 133–148.

(144) Romano, J. P.; Wolf, M. Resurrecting weighted least squares. *Journal of Econometrics* **2017**, *197* (1), 1–19.

(145) Flatman, R.; McLAUCHLAN, R. W.; Juge, N.; Furniss, C.; Berrin, J.-G.; Hughes, R. K.; Manzanares, P.; Ladbury, J. E.; O'BRIEN, R.; Williamson, G. Interactions defining the specificity between fungal xylanases and the xylanase-inhibiting protein XIP-I from wheat. *Biochemical Journal* **2002**, *365* (3), 773–781.

(146) Tellinghuisen, J. Weighted least squares in calibration: the problem with using "quality coefficients" to select weighting formulas. *Journal of Chromatography B* **2008**, *872* (1–2), 162–166.

(147) Carroll, R. J.; Wu, C. J.; Ruppert, D. The effect of estimating weights in weighted least squares. *Journal of the American Statistical Association* **1988**, *83* (404), 1045–1054.

(148) Grossfield, A.; Zuckerman, D. M. Quantifying uncertainty and sampling quality in biomolecular simulations. *Annual Reports in Computational Chemistry* **2009**, *5*, 23–48.

(149) Williams, P. J.; Kim, Y. H. Resampling techniques and their application to pharmacometrics. *Pharmacometrics: The Science of Quantitative Pharmacology* **2007**, 401–419.

(150) Schulte, C.; Soldà, A.; Spänig, S.; Adams, N.; Bekić, I.; Streicher, W.; Heider, D.; Strasser, R.; Maric, H. M. Multivalent binding kinetics resolved by fluorescence proximity sensing. *Communications Biology* **2022**, *5* (1), 1070.

(151) Efron, B.; Tibshirani, R. J. *An Introduction to the Bootstrap*; Chapman and Hall/CRC, 1994.

(152) Hong, H.; Carlin, B. P.; Shamliyan, T. A.; Wyman, J. F.; Ramakrishnan, R.; Sainfort, F.; Kane, R. L. Comparing Bayesian and frequentist approaches for multiple outcome mixed treatment comparisons. *Medical Decision Making* **2013**, *33* (5), 702–714.

(153) Gelman, A.; Hwang, J.; Vehtari, A. Understanding predictive information criteria for Bayesian models. *Statistics and Computing* **2014**, *24*, 997–1016.

(154) Mohammad-Djafari, A. Bayesian approach with prior models which enforce sparsity in signal and image processing. *EURASIP Journal on Advances in Signal Processing* **2012**, *2012*, 1–19.

(155) Radivojević, T.; Akhmatskaya, E. Modified hamiltonian monte carlo for bayesian inference. *Statistics and Computing* **2020**, *30* (2), 377–404.

(156) Richey, M. The evolution of Markov chain Monte Carlo methods. *The American Mathematical Monthly* **2010**, *117* (5), 383–413.

(157) Lintusaari, J.; Gutmann, M. U.; Dutta, R.; Kaski, S.; Corander, J. Fundamentals and recent developments in approximate Bayesian computation. *Systematic Biology* **2017**, *66* (1), e66–e82.

(158) Roy, V. Convergence diagnostics for Markov chain Monte Carlo. *Annual Review of Statistics and Its Application* **2020**, *7* (1), 387–412.

(159) Levitus, M. Tutorial: measurement of fluorescence spectra and determination of relative fluorescence quantum yields of transparent samples. *Methods and Applications in Fluorescence* **2020**, *8* (3), 033001.

(160) Werle, P.; Mazzinghi, P.; D'Amato, F.; De Rosa, M.; Maurer, K.; Slemr, F. Signal processing and calibration procedures for in situ diode-laser absorption spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2004**, *60* (8–9), 1685–1705.

(161) Huang, Z.-h.; Hong, X.-y.; Dong, L.-h.; Zhan, S.-y.; Wang, X.-p.; Liu, X. Study on calibration method for the performance index of SPR sensors. *Optoelectronics Letters* **2013**, *9* (5), 329–332.

(162) Schasfoort, R. B. M.; McWhirter, A.; Kooyman, R. P. H.; Corn, R. M.; Wark, A.; Lee, H. J.; Gedig, E.; Engbers, G.; Walstrom, L.; de Mol, N. J.; et al. SPR Instrumentation. In *Handbook of Surface Plasmon Resonance*, Schasfoort, R. B. M., Tudos, A. J. Eds.; The Royal Society of Chemistry, 2008; pp 35–80.

(163) Cheng, W. L.; Markus, C.; Lim, C. Y.; Tan, R. Z.; Sethi, S. K.; Loh, T. P. Calibration practices in clinical mass spectrometry: Review and recommendations. *Annals of Laboratory Medicine* **2023**, *43* (1), 5–18.

(164) Mao, Z. Evaluation and Impact of Measurement Uncertainty in Laboratory Instruments and Analytical Methods: A Comprehensive Study of Volumetric, Chromatographic, and Gravimetric Techniques. *International Journal of Chemistry and Materials Science* **2024**, *1* (1), 10.70088.

(165) Jani, D.; Allinson, J.; Berisha, F.; Cowan, K. J.; Devanarayan, V.; Gleason, C.; Jeromin, A.; Keller, S.; Khan, M. U.; Nowatzke, B. Recommendations for use and fit-for-purpose validation of biomarker multiplex ligand binding assays in drug development. *The AAPS Journal* **2016**, *18*, 1–14.

(166) Demchenko, A. P. Optimization of fluorescence response in the design of molecular biosensors. *Analytical Biochemistry* **2005**, *343* (1), 1–22.

(167) Yanase, Y.; Hiragun, T.; Kaneko, S.; Gould, H. J.; Greaves, M. W.; Hide, M. Detection of refractive index changes in individual living cells by means of surface plasmon resonance imaging. *Biosensors and Bioelectronics* **2010**, *26* (2), 674–681.

(168) Keller, S.; Vargas, C.; Zhao, H.; Piszczek, G.; Brautigam, C. A.; Schuck, P. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Analytical Chemistry* **2012**, *84* (11), 5066–5073.

(169) Olaru, A.; Bala, C.; Jaffrezic-Renault, N.; Aboul-Enein, H. Y. Surface plasmon resonance (SPR) biosensors in pharmaceutical analysis. *Critical Reviews in Analytical Chemistry* **2015**, *45* (2), 97–105.

(170) Patel, K. H.; Evenhuis, C. J.; Cherney, L. T.; Krylov, S. N. Simplified universal method for determining electrolyte temperatures in a capillary electrophoresis instrument with forced-air cooling. *Electrophoresis* **2012**, *33* (6), 1079–1085.

(171) Sisavath, N.; Rukundo, J. L.; Le Blanc, J. Y.; Galievsky, V. A.; Bao, J.; Kochmann, S.; Stasheuski, A. S.; Krylov, S. N. Transient incomplete separation facilitates finding accurate equilibrium dissociation constant of protein–small molecule complex. *Angewandte Chemie* **2019**, *131* (20), 6707–6711.

(172) Rukundo, J.-L.; Le Blanc, J. Y.; Kochmann, S.; Krylov, S. N. Assessing Accuracy of an Analytical Method in silico: Application to "Accurate Constant via Transient Incomplete Separation"(ACTIS). *Analytical Chemistry* **2020**, *92* (17), 11973–11980.

(173) Chamieh, J.; Leclercq, L.; Martin, M.; Slaoui, S.; Jensen, H.; Østergaard, J.; Cottet, H. Limits in size of Taylor dispersion analysis: representation of the different hydrodynamic regimes and application to the size-characterization of cubosomes. *Analytical Chemistry* **2017**, *89* (24), 13487–13493.

(174) Okada, T. Hydrodynamic Chromatography in Narrow and Wide-Bores; Whether Radial Diffusion Is Essential or Not. *Journal of Liquid Chromatography & Related Technologies* **2010**, *33* (9-12), 1116–1129.

(175) Adrover, A.; Cerbelli, S.; Garofalo, F.; Giona, M. Convection-dominated dispersion regime in wide-bore chromatography: A transport-based approach to assess the occurrence of slip flows in microchannels. *Analytical Chemistry* **2009**, *81* (19), 8009–8014.

(176) Duvvuri, H.; Wheeler, L. C.; Harms, M. J. pytc: Open-source python software for global analyses of isothermal titration calorimetry data. *Biochemistry* **2018**, *57* (18), 2578–2583.

(177) Rukundo, J.-L.; Kochmann, S.; Wang, T. Y.; Ivanov, N. A.; Le Blanc, J. Y.; Gorin, B. I.; Krylov, S. N. Template Instrumentation for "Accurate Constant via Transient Incomplete Separation". *Analytical Chemistry* **2021**, *93* (34), 11654–11659.

(178) Wang, T. Y.; Rukundo, J.-L.; Le, A. T.; Ivanov, N. A.; Le Blanc, J. Y.; Gorin, B. I.; Krylov, S. N. Transient Incomplete Separation of Species with Close Diffusivity to Study the Stability of Affinity Complexes. *Analytical Chemistry* **2022**, *94* (44), 15415–15422.

(179) Wang, T. Y.; Ji, H.; Everton, D.; Le, A. T.; Krylova, S. M.; Fournier, R.; Krylov, S. N. Fundamental Determinants of the Accuracy of Equilibrium Constants for Affinity Complexes. *Analytical Chemistry* **2023**, *95* (42), 15826–15832.

(180) Pollard, T. D. A guide to simple and informative binding assays. *Molecular Biology of the Cell* **2010**, *21* (23), 4061–4067.

(181) Hulme, E. C.; Trevethick, M. A. Ligand binding assays at equilibrium: validation and interpretation. *British Journal of Pharmacology* **2010**, *161* (6), 1219–1237.

(182) Wang, T. Y.; Rukundo, J.-L.; Mao, Z.; Krylov, S. N. Maximizing the Accuracy of Equilibrium Dissociation Constants for Affinity Complexes: From Theory to Practical Recommendations. *ACS Chemical Biology* **2024**, *19* (9), 1852–1867.

(183) Wang, T. Y.; Dhillon, P.; Krylova, S. M.; Bijlani, A.; Schreiber, S. G.-K.; Jose, D.; Krylov, S. N. Introducing Quantitative Assessment of Michaelis Constant (Km) Accuracy. 2025, *submitted*.

(184) Callies, O.; Daranas, A. H. Application of isothermal titration calorimetry as a tool to study natural product interactions. *Natural Product Reports* **2016**, *33* (7), 881–904.

(185) Stangherlin, S.; Ding, Y.; Liu, J. Dissociation Constant (Kd) Measurement for Small-Molecule Binding Aptamers: Homogeneous Assay Methods and Critical Evaluations. *Small Methods*, 2401572.

(186) Thomson, J. A.; Ladbury, J. E. Isothermal titration calorimetry: a tutorial. *Biocalorimetry* **2004**, *2*, 37–58.

(187) Wang, T. Y.; Krylov, S. N. Deterministic Error Propagation in ITC: Revealing Multi-Fold Errors in K_d Values Under Standard Conditions. *Biophysical Chemistry* **2025**, *323*, 107455.

(188) Wang, T.Y.; Bijlani, A.; Chao, E. H. P.; Johnson, P.E.; Krylov, S.N. A Browser-Based Tool for Assessing Accuracy of ITC-Derived Parameters: K_d, $\Delta H°$, and *n.* **2025**, *ChemBioChem, 2500194,* published online.

(189) Khavrutskii, L.; Yeh, J.; Timofeeva, O.; Tarasov, S. G.; Pritt, S.; Stefanisko, K.; Tarasova, N. Protein purification-free method of binding affinity determination by microscale thermophoresis. *Journal of Visualized Experiments: JOVE* **2013**, (78), 50541.

(190) Pan, Y.; Sackmann, E. K.; Wypisniak, K.; Hornsby, M.; Datwani, S. S.; Herr, A. E. Determination of equilibrium dissociation constants for recombinant antibodies by high-throughput affinity electrophoresis. *Scientific Reports* **2016**, *6* (1), 39774.

(191) Krylov, S. N. Nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM): A novel method for biomolecular screening. *SLAS Discovery* **2006**, *11* (2), 115–122.

(192) Bahga, S. S.; Bercovici, M.; Santiago, J. G. Robust and high-resolution simulations of nonlinear electrokinetic processes in variable cross-section channels. *Electrophoresis* **2012**, *33* (19-20), 3036–3051.

(193) Takahashi, K.; Arjunan, S. N. V.; Tomita, M. Space in systems biology of signaling pathways–towards intracellular molecular crowding in silico. *FEBS Letters* **2005**, *579* (8), 1783–1788.

(194) Campoy, A. V.; Freire, E. ITC in the post-genomic era…? Priceless. *Biophysical Chemistry* **2005**, *115* (2-3), 115–124.

(195) Nouadje, G.; Nertz, M.; Verdeguer, P.; Couderc, F. Ball-lens laser-induced fluorescence detector as an easy-to-use highly sensitive detector for capillary electrophoresis application to the identification of biogenic amines in dairy products. *Journal of Chromatography A* **1995**, *717* (1–2), 335–343.

(196) Drabovich, A. P.; Berezovski, M.; Okhonin, V.; Krylov, S. N. Selection of smart aptamers by methods of kinetic capillary electrophoresis. *Analytical Chemistry* **2006**, *78* (9), 3171–3178.

(197) Qian, S.; Chang, D.; He, S.; Li, Y. Aptamers from random sequence space: Accomplishments, gaps and future considerations. *Analytica Chimica Acta* **2022**, *1196*, 339511.

(198) Nkodo, A. E.; Garnier, J. M.; Tinland, B.; Ren, H.; Desruisseaux, C.; McCormick, L. C.; Drouin, G.; Slater, G. W. Diffusion coefficient of DNA molecules during free solution electrophoresis. *Electrophoresis* **2001**, *22* (12), 2424–2432.

(199) de Jong, S.; Krylov, S. N. Pressure-based approach for the analysis of protein adsorption in capillary electrophoresis. *Analytical Chemistry* **2012**, *84* (1), 453–458.

(200) Edward, J. T. Molecular volumes and the Stokes-Einstein equation. *Journal of Chemical Education* **1970**, *47* (4), 261.

(201) Tehranirokh, M.; Van den Bronk, M.; Smith, P.; Dai, Z.; Ragunathan, K.; Muscalu, A.; Mills, S.; Breadmore, M. C.; Shellie, R. A. Automated liquid-liquid extraction of organic compounds from aqueous samples using a multifunction autosampler syringe. *Journal of Chromatography A* **2021**, *1642*, 462032.

(202) Berezovski, M.; Krylov, S. N. Nonequilibrium capillary electrophoresis of equilibrium mixtures− a single experiment reveals equilibrium and kinetic parameters of protein−DNA interactions. *Journal of the American Chemical Society* **2002**, *124* (46), 13674–13675.

(203) Berezovski, M.; Drabovich, A.; Krylova, S. M.; Musheev, M.; Okhonin, V.; Petrov, A.; Krylov, S. N. Nonequilibrium capillary electrophoresis of equilibrium mixtures: a universal tool for development of aptamers. *Journal of the American Chemical Society* **2005**, *127* (9), 3165–3171.

(204) Lou, B.; Chen, E.; Zhao, X.; Qu, F.; Yan, J. The application of capillary electrophoresis for assisting whole-cell aptamers selection by characterizing complete ssDNA distribution. *Journal of Chromatography a* **2016**, *1437*, 203–209.

(205) Riley, K. R.; Gagliano, J.; Xiao, J.; Libby, K.; Saito, S.; Yu, G.; Cubicciotti, R.; Macosko, J.; Colyer, C. L.; Guthold, M. Combining capillary electrophoresis and next-generation sequencing for aptamer selection. *Analytical and Bioanalytical Chemistry* **2015**, *407*, 1527–1532.

(206) Berezovski, M.; Krylov, S. N. Using nonequilibrium capillary electrophoresis of equilibrium mixtures for the determination of temperature in capillary electrophoresis. *Analytical Chemistry* **2004**, *76* (23), 7114–7117.

(207) Xuan, X.; Li, D. Analytical study of Joule heating effects on electrokinetic transportation in capillary electrophoresis. *Journal of Chromatography A* **2005**, *1064* (2), 227–237.

(208) Musheev, M. U.; Filiptsev, Y.; Krylov, S. N. Noncooled capillary inlet: a source of systematic errors in capillary-electrophoresis-based affinity analyses. *Analytical Chemistry* **2010**, *82* (20), 8637–8641.

(209) Almehmadi, L. M.; Valsangkar, V. A.; Halvorsen, K.; Zhang, Q.; Sheng, J.; Lednev, I. K. Surface-enhanced Raman spectroscopy for drug discovery: peptide-RNA binding. *Analytical and Bioanalytical Chemistry* **2022**, *414* (20), 6009–6016.

(210) Fu, H.; Qian, C.; Tong, W.; Li, H.; Chen, D. D. Mass spectrometry and affinity capillary electrophoresis for characterization of host-guest interactions. *Journal of Chromatography A* **2019**, *1589*, 182–190.

(211) Arshavsky-Graham, S.; Massad-Ivanir, N.; Paratore, F.; Scheper, T.; Bercovici, M.; Segal, E. On chip protein pre-concentration for enhancing the sensitivity of porous silicon biosensors. *ACS Sensors* **2017**, *2* (12), 1767–1773.

(212) Avila, L. Z.; Chu, Y. H.; Blossey, E. C.; Whitesides, G. M. Use of affinity capillary electrophoresis to determine kinetic and equilibrium constants for binding of arylsulfonamides to bovine carbonic anhydrase. *Journal of Medicinal Chemistry* **1993**, *36* (1), 126–133.

(213) Mwakibete, H.; Cristantino, R.; Bloor, D.; Wyn-Jones, E.; Holzwarth, J. Reliability of the experimental methods to determine equilibrium constants for surfactant/cyclodextrin inclusion complexes. *Langmuir* **1995**, *11* (1), 57–60.

(214) Du, X.; Li, Y.; Xia, Y.-L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q. Insights into protein–ligand interactions: mechanisms, models, and methods. *International Journal of Molecular Sciences* **2016**, *17* (2), 144.

(215) Wätzig, H.; Oltmann-Norden, I.; Steinicke, F.; Alhazmi, H. A.; Nachbar, M.; El-Hady, D. A.; Albishri, H. M.; Baumann, K.; Exner, T.; Böckler, F. M. Data quality in drug discovery: the role of analytical performance in ligand binding assays. *Journal of Computer-Aided Molecular Design* **2015**, *29*, 847–865.

(216) Shalaeva, M.; Kenseth, J.; Lombardo, F.; Bastin, A. Measurement of dissociation constants (pKa values) of organic compounds by multiplexed capillary electrophoresis using aqueous and cosolvent buffers. *Journal of Pharmaceutical Sciences* **2008**, *97* (7), 2581–2606.

(217) Zhang, H.; Li, X.-F.; Le, X. C. Binding-induced DNA assembly and its application to yoctomole detection of proteins. *Analytical Chemistry* **2012**, *84* (2), 877–884.

(218) Nakane, R.; Kurihara, O.; Natsubori, A. Boron Isotope Exchange between Boron Fluoride and Its Alkyl Halide Complexes. I. Relation between Equilibrium Constant of Isotopic Exchange

Reaction, Stability, and Catalytic Activity of Boron Fluoride Complex. *The Journal of Physical Chemistry* **1964**, *68* (10), 2876–2882.

(219) Han, B. A suite of mathematical solutions to describe ternary complex formation and their application to targeted protein degradation by heterobifunctional ligands. *Journal of Biological Chemistry* **2020**, *295* (45), 15280–15291.

(220) Qian, C.; Fu, H.; Kovalchik, K. A.; Li, H.; Chen, D. D. Y. Specific binding constant and stoichiometry determination in free solution by mass spectrometry and capillary electrophoresis frontal analysis. *Analytical Chemistry* **2017**, *89* (17), 9483–9490.

(221) Tang, Z.; Shangguan, D.; Wang, K.; Shi, H.; Sefah, K.; Mallikratchy, P.; Chen, H. W.; Li, Y.; Tan, W. Selection of aptamers for molecular recognition and characterization of cancer cells. *Analytical Chemistry* **2007**, *79* (13), 4900–4907.

(222) Cooper, M. A. Optical biosensors in drug discovery. *Nature Reviews Drug Discovery* **2002**, *1* (7), 515–528.

(223) Shortridge, M. D.; Hage, D. S.; Harbison, G. S.; Powers, R. Estimating protein−ligand binding affinity using high-throughput screening by NMR. *Journal of Combinatorial Chemistry* **2008**, *10* (6), 948–958.

(224) Jusko, W. J.; Molins, E. A.; Ayyar, V. S. Seeking nonspecific binding: assessing the reliability of tissue dilutions for calculating fraction unbound. *Drug Metabolism and Disposition* **2020**, *48* (10), 894–902.

(225) Strohkendl, I.; Saifuddin, F. A.; Rybarski, J. R.; Finkelstein, I. J.; Russell, R. Kinetic basis for DNA target specificity of CRISPR-Cas12a. *Molecular Cell* **2018**, *71* (5), 816–824. e813.

(226) Tellinghuisen, J. Designing isothermal titration calorimetry experiments for the study of 1: 1 binding: problems with the "standard protocol". *Analytical Biochemistry* **2012**, *424* (2), 211–220.

(227) Tellinghuisen, J. Statistical error in isothermal titration calorimetry. In *Methods in Enzymology*, Vol. 383; Elsevier, 2004; pp 245–282.

(228) Rundlett, K. L.; Armstrong, D. W. Examination of the origin, variation, and proper use of expressions for the estimation of association constants by capillary electrophoresis. *Journal of Chromatography A* **1996**, *721* (1), 173–186.

(229) Bowser, M. T.; Chen, D. D. Monte Carlo simulation of error propagation in the determination of binding constants from rectangular hyperbolae. 2. Effect of the maximum-response range. *The Journal of Physical Chemistry A* **1999**, *103* (1), 197–202.

(230) Tellinghuisen, J. A study of statistical error in isothermal titration calorimetry. *Analytical Biochemistry* **2003**, *321* (1), 79–88.

(231) Krylova, S. M.; Karkhanina, A. A.; Musheev, M. U.; Bagg, E. A.; Schofield, C. J.; Krylov, S. N. DNA aptamers for as analytical tools for the quantitative analysis of DNA-dealkylating enzymes. *Analytical Biochemistry* **2011**, *414* (2), 261–265.

(232) Krylov, S. N. NECEEM for development, characterisation and analytical utilisation of aptamers. *LabPlus International* **2005**, 6–9.

(233) Hoare, S. R.; Fleck, B. A.; Williams, J. P.; Grigoriadis, D. E. The importance of target binding kinetics for measuring target binding affinity in drug discovery: a case study from a CRF1 receptor antagonist program. *Drug Discovery Today* **2020**, *25* (1), 7–14.

(234) Archakov, A. I.; Govorun, V. M.; Dubanov, A. V.; Ivanov, Y. D.; Veselovsky, A. V.; Lewi, P.; Janssen, P. Protein-protein interactions as a target for drugs in proteomics. *Proteomics* **2003**, *3* (4), 380–391.

(235) Strange, P. G. Antipsychotic drugs: importance of dopamine receptors for mechanisms of therapeutic actions and side effects. *Pharmacological Reviews* **2001**, *53* (1), 119–134.

(236) Tracy, T. S.; Hummel, M. A. Modeling kinetic data from in vitro drug metabolism enzyme experiments. *Drug Metabolism Reviews* **2004**, *36* (2), 231–242.

(237) Halford, S. E. An end to 40 years of mistakes in DNA–protein association kinetics? *Biochemical Society Transactions* **2009**, *37* (2), 343–348.

(238) Sims, R. J.; Reinberg, D. Histone H3 Lys 4 methylation: caught in a bind? *Genes & Development* **2006**, *20* (20), 2779–2786.

(239) Rusling, J. F. Computerized method for mechanistic classification of one-electron potentiostatic current-potential curves. *Analytical Chemistry* **1983**, *55* (11), 1713–1718.

(240) Johnson, M. L.; Correia, J. J.; Yphantis, D. A.; Halvorson, H. R. Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophysical Journal* **1981**, *36* (3), 575–588.

(241) Wells, J. A. Binding in the growth hormone receptor complex. *Proceedings of the National Academy of Sciences* **1996**, *93* (1), 1–6.

(242) Levengood, J. D.; Tolbert, B. S. Idiosyncrasies of hnRNP A1-RNA recognition: Can binding mode influence function. In *Seminars in Cell & Developmental Biology*, 2019; Elsevier: Vol. 86, pp 150–161.

(243) Teclemichael, E.; Le, A. T.; Krylova, S. M.; Wang, T. Y.; Krylov, S. N. Bulk Affinity Assays in Aptamer Selection: Challenges, Theory, and Workflow. *Analytical Chemistry* **2022**, *94* (44), 15183–15188.

(244) Le, A. T.; Krylova, S. M.; Krylov, S. N. Determination of the Equilibrium Constant and Rate Constant of Protein–Oligonucleotide Complex Dissociation under the Conditions of Ideal-Filter Capillary Electrophoresis. *Analytical Chemistry* **2019**, *91* (13), 8532–8539.

(245) Piñeiro, L.; Novo, M.; Al-Soufi, W. Fluorescence emission of pyrene in surfactant solutions. *Advances in Colloid and Interface Science* **2015**, *215*, 1–12.

(246) Zhao, Q.; Tao, J.; Feng, W.; Uppal, J. S.; Peng, H.; Le, X. C. Aptamer binding assays and molecular interaction studies using fluorescence anisotropy-a review. *Analytica Chimica Acta* **2020**, *1125*, 267–278.

(247) Krylova, S. M.; Dove, P. M.; Kanoatov, M.; Krylov, S. N. Slow-dissociation and slow-recombination assumptions in nonequilibrium capillary electrophoresis of equilibrium mixtures. *Analytical Chemistry* **2011**, *83* (19), 7582–7585.

(248) Okhonin, V.; Berezovski, M. V.; Krylov, S. N. MASKE: Macroscopic approach to studying kinetics at equilibrium. *Journal of the American Chemical Society* **2010**, *132* (20), 7062–7068.

(249) Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; West, G. M.; Griffin, P. R. Differential hydrogen/deuterium exchange mass spectrometry analysis of protein–ligand interactions. *Expert Review of Proteomics* **2011**, *8* (1), 43–59.

(250) Kanoatov, M.; Mehrabanfar, S.; Krylov, S. N. Systematic approach to optimization of experimental conditions in nonequilibrium capillary electrophoresis of equilibrium mixtures. *Analytical Chemistry* **2016**, *88* (18), 9300–9308.

(251) Poole, S. K.; Patel, S.; Dehring, K.; Workman, H.; Poole, C. F. Determination of acid dissociation constants by capillary electrophoresis. *Journal of Chromatography A* **2004**, *1037* (1–2), 445–454.

(252) Pesce, A. J.; Michael, J. G. Artifacts and limitations of enzyme immunoassay. *Journal of Immunological Methods* **1992**, *150* (1–2), 111–119.

(253) Nistor, C.; Emnéus, J. Immunoassay: potentials and limitations. *Comprehensive Analytical Chemistry* **2005**, *44*, 375–427.

(254) Tighe, P. J.; Ryder, R. R.; Todd, I.; Fairclough, L. C. ELISA in the multiplex era: potentials and pitfalls. *PROTEOMICS–Clinical Applications* **2015**, *9* (3–4), 406–422.

(255) Nguyen, B.; Tanious, F. A.; Wilson, W. D. Biosensor-surface plasmon resonance: quantitative analysis of small molecule–nucleic acid interactions. *Methods* **2007**, *42* (2), 150–161.

(256) Unser, S.; Bruzas, I.; He, J.; Sagle, L. Localized surface plasmon resonance biosensing: current challenges and approaches. *Sensors* **2015**, *15* (7), 15684–15716.

(257) Petersen, R. L. Strategies using bio-layer interferometry biosensor technology for vaccine research and development. *Biosensors* **2017**, *7* (4), 49.

(258) Renaud, J.-P.; Chung, C.-w.; Danielson, U. H.; Egner, U.; Hennig, M.; Hubbard, R. E.; Nar, H. Biophysics in drug discovery: impact, challenges and opportunities. *Nature reviews Drug Discovery* **2016**, *15* (10), 679–698.

(259) Lee, N.; Wiegand, S. Thermophoretic micron-scale devices: practical approach and review. *Entropy* **2020**, *22* (9), 950.

(260) Magnez, R.; Bailly, C.; Thuru, X. Microscale thermophoresis as a tool to study protein interactions and their implication in human diseases. *International Journal of Molecular Sciences* **2022**, *23* (14), 7672.

(261) Asmari, M.; Michalcová, L.; Ibrahim, A. E.; Glatz, Z.; Wätzig, H.; El Deeb, S. Studying molecular interactions via capillary electrophoresis and microscale thermophoresis: A review. *Electrophoresis* **2023**, *44* (13–14), 1114–1142.

(262) Leavitt, S.; Freire, E. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Current Opinion in Structural Biology* **2001**, *11* (5), 560–566.

(263) Prozeller, D.; Morsbach, S.; Landfester, K. Isothermal titration calorimetry as a complementary method for investigating nanoparticle–protein interactions. *Nanoscale* **2019**, *11* (41), 19265–19273.

(264) Dalvit, C.; Gmür, I.; Rößler, P.; Gossert, A. D. Affinity measurement of strong ligands with NMR spectroscopy: Limitations and ways to overcome them. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2023**, 138–139, 52–69.

(265) Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T. L.; Homans, S. W.; Kessler, H.; Luchinat, C. Perspectives on NMR in drug discovery: a technique comes of age. *Nature Reviews Drug Discovery* **2008**, *7* (9), 738–745.

(266) Dalvit, C. NMR methods in fragment screening: theory and a comparison with other biophysical techniques. *Drug Discovery Today* **2009**, *14* (21–22), 1051–1057.

(267) Krylov, S. N. Kinetic CE: foundation for homogeneous kinetic affinity methods. *Electrophoresis* **2007**, *28* (1–2), 69–88.

(268) Chu, Y. H.; Whitesides, G. M. Affinity Capillary Electrophoresis Can Simultaneously Measure Binding Constants of Multiple Peptides to Vancomycin. *Journal of Organic Chemistry* **1992**, *57* (13), 3524–3525.

(269) Gomez, F. A.; Avila, L. Z.; Chu, Y. H.; Whitesides, G. M. Determination of Binding Constants of Ligands to Proteins by Affinity Capillary Electrophoresis: Compensation for Electroosmotic Flow. *Analytical Chemistry* **1994**, *66* (11), 1785–1791.

(270) Jameson, D. M.; Ross, J. A. Fluorescence polarization/anisotropy in diagnostics and imaging. *Chemical Reviews* **2010**, *110* (5), 2685–2708.

(271) Wolff, M. Polarization of light reflected from rough planetary surface. *Applied Optics* **1975**, *14* (6), 1395–1405.

(272) Jablonski, A. On the notion of emission anisotropy. *Bulletin of the Polish Academy of Sciences Technical Sciences* **1960**, *8*, 259–264.

(273) Gradinaru, C. C.; Marushchak, D. O.; Samim, M.; Krull, U. J. Fluorescence anisotropy: from single molecules to live cells. *Analyst* **2010**, *135* (3), 452–459.

(274) Weber, G. Polarization of the fluorescence of macromolecules. I. Theory and experimental method. *Biochem J* **1952**, *51* (2), 145–155.

(275) *Synergy H1 Operator's Manual*. BioTek® Instruments, Inc., 2012. https://www.lebonheur.org/research/investigator-resources/manuals/BioTek%20Synergy%20H1%20Hybrid%20Multi-Mode%20Plate%20Reader%20manual.pdf (accessed 2024 03-09).

(276) *Omega Series Operating Manual*. BMG LABTECH, 2017. https://med-fom-avgaylab.sites.olt.ubc.ca/files/2018/04/0415B0001K-Operating-Manual-Omega.pdf (accessed 2024 03-09).

(277) Clark, S. M.; Konermann, L. Determination of ligand− protein dissociation constants by electrospray mass spectrometry-based diffusion measurements. *Analytical Chemistry* **2004**, *76* (23), 7077–7083.

(278) Jensen, H.; Østergaard, J. Flow induced dispersion analysis quantifies noncovalent interactions in nanoliter samples. *Journal of the American Chemical Society* **2010**, *132* (12), 4070–4071.

(279) Bielejewska, A.; Bylina, A.; Duszczyk, K.; Fiałkowski, M.; Hołyst, R. Evaluation of ligand-selector interaction from effective diffusion coefficient. *Analytical Chemistry* **2010**, *82* (13), 5463–5469.

(280) Clark, S. M.; Leaist, D. G.; Konermann, L. Taylor dispersion monitored by electrospray mass spectrometry: a novel approach for studying diffusion in solution. *Rapid Communications in Mass Spectrometry* **2002**, *16* (15), 1454–1462.

(281) Griffiths, A. On the movement of a coloured index along a capillary tube, and its application to the measurement of the circulation of water in a closed circuit. *Proceedings of the Physical Society of London* **1910**, *23* (1), 190.

(282) Taylor, G. I. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1953**, *219* (1137), 186–203.

(283) Trumbore, C. N.; Grehlinger, M.; Stowe, M.; Keleher, F. M. Further experiments on a new, fast method for determining molecular weights of diffusing species in a liquid phase. *Journal of Chromatography A* **1985**, *322*, 443–454.

(284) Schulte, J.; Tants, J.-N.; von Ehr, J.; Schlundt, A.; Morgner, N. Determination of dissociation constants via quantitative mass spectrometry. *Frontiers in Analytical Science* **2023**, *3*, 1119489.

(285) Kempen, E. C.; Brodbelt, J. S. A method for the determination of binding constants by electrospray ionization mass spectrometry. *Analytical Chemistry* **2000**, *72* (21), 5411–5416.

(286) Rinken, T.; Rinken, A.; Tenno, T.; Järv, J. Calibration of glucose biosensors by using pre-steady state kinetic data. *Biosensors and Bioelectronics* **1998**, *13* (7–8), 801–807.

(287) Canbay, H. S. Spectrophotometric determination of acid dissociation constants of some arylpropionic acids and arylacetic acids in acetonitrile-water binary mixtures at 25ºC. *Brazilian Journal of Pharmaceutical Sciences* **2023**, *58*, e20740.

(288) Van de Weert, M.; Stella, L. Fluorescence quenching and ligand binding: A critical discussion of a popular methodology. *Journal of Molecular Structure* **2011**, *998* (1–3), 144–150.

(289) Andrus, A.; Bloch, W. HPLC of oligonucleotides and polynucleotides. *HPLC Macromol* **1998**, *872*, 141.

(290) Raynal, B.; Lenormand, P.; Baron, B.; Hoos, S.; England, P. Quality assessment and optimization of purified protein samples: why and how? *Microbial Cell Factories* **2014**, *13*, 1–10.

(291) Bange, A.; Halsall, H. B.; Heineman, W. R. Microfluidic immunosensor systems. *Biosensors and Bioelectronics* **2005**, *20* (12), 2488–2503.

(292) Chen, T.-S. Determination of the capacitance, inductance, and characteristic impedance of rectangular lines. *IRE Transactions on Microwave Theory and Techniques* **1960**, *8* (5), 510–519.

(293) O'brien, R.; Cannon, D.; Rowlands, W. Electroacoustic determination of particle size and zeta potential. *Journal of Colloid and Interface Science* **1995**, *173* (2), 406–418.

(294) Domich, P. D.; Baker-Jarvis, J.; Geyer, R. G. Optimization techniques for permittivity and permeability determination. *Journal of Research of the National Institute of Standards and Technology* **1991**, *96* (5), 565.

(295) Paketurytė, V.; Petrauskas, V.; Zubrienė, A.; Abian, O.; Bastos, M.; Chen, W.-Y.; Moreno, M. J.; Krainer, G.; Linkuvienė, V.; Sedivy, A. Uncertainty in protein–ligand binding constants: asymmetric confidence intervals versus standard errors. *European Biophysics Journal* **2021**, *50* (3), 661–670.

(296) Sohn, R.; Menke, W. Application of maximum likelihood and bootstrap methods to nonlinear curve-fit problems in geochemistry. *Geochemistry, Geophysics, Geosystems* **2002**, *3* (7), 1–17.

(297) Sullivan, S. G.; Greenland, S. Bayesian regression in SAS software. *International Journal of Epidemiology* **2013**, *42* (1), 308–317.

(298) Gunn-Sandell, L. B.; Bedrick, E. J.; Hutchins, J. L.; Berg, A. A.; Kaizer, A. M.; Carlson, N. E. A practical guide to adopting Bayesian analyses in clinical research. *Journal of Clinical and Translational Science* **2024**, *8* (1), e3.

(299) Klakamp, S. L.; Drake, A. W. Miscalculating equilibrium constants for multivalent protein complexes: site-binding concentration or protein molecular concentration? *Journal of Pharmaceutical Sciences* **2021**, *110* (7), 2585–2589.

(300) Wienken, C. J.; Baaske, P.; Rothbauer, U.; Braun, D.; Duhr, S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nature Communications* **2010**, *1* (1), 100.

(301) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable determinations of protein–ligand interactions by direct ESI-MS measurements. Are we there yet? *Journal of The American Society for Mass Spectrometry* **2012**, *23* (3), 431–441.

(302) Mokhov, V.; Aliukov, S.; Alabugin, A.; Osintsev, K. A Review of Mathematical Models of Macroeconomics, Microeconomics, and Government Regulation of the Economy. *Mathematics* **2023**, *11* (14), 3246.

(303) Rust, R. T.; Metters, R. Mathematical models of service. *European Journal of Operational Research* **1996**, *91* (3), 427–439.

(304) Mellaku, M. T.; Sebsibe, A. S. Potential of mathematical model-based decision making to promote sustainable performance of agriculture in developing countries: A review article. *Heliyon* **2022**, *8* (2), e08968.

(305) Brown, R. J. Systematic error arising from "sequential" standard addition calibrations. 2. Determination of analyte mass fraction in blank solutions. *Analytica Chimica Acta* **2009**, *648* (2), 153–156.

(306) Morley, S. K. Alternatives to accuracy and bias metrics based on percentage errors for radiation belt modeling applications. **2016**, https://doi.org/10.2172/1260362.

(307) Zhu, D.; Stumpf, C. R.; Krahn, J. M.; Wickens, M.; Hall, T. M. T. A 5′ cytosine binding pocket in Puf3p specifies regulation of mitochondrial mRNAs. *Proceedings of the National Academy of Sciences* **2009**, *106* (48), 20192–20197.

(308) Wang, Y.; Opperman, L.; Wickens, M.; Hall, T. M. T. Structural basis for specific recognition of multiple mRNA targets by a PUF regulatory protein. *Proceedings of the National Academy of Sciences* **2009**, *106* (48), 20186–20191.

(309) Ruckstuhl, A. Introduction to nonlinear regression. *IDP Institut fur Datenanalyse und Prozessdesign, Zurcher Hochschule fur Angewandte Wissenschaften* **2010**, 365.

(310) Wang, N. X.; von Recum, H. A. Affinity-based drug delivery. *Macromolecular Bioscience* **2011**, *11* (3), 321–332.

(311) Zheng, C.; Han, L.; Yap, C.; Ji, Z.; Cao, Z.; Chen, Y. Therapeutic targets: progress of their exploration and investigation of their characteristics. *Pharmacological Reviews* **2006**, *58* (2), 259–279.

(312) Szwajkajzer, D.; Carey, J. Molecular and biological constraints on ligand-binding affinity and specificity. *Biopolymers: Original Research on Biomolecules* **1997**, *44* (2), 181–198.

(313) Paketurytė, V.; Linkuvienė, V.; Krainer, G.; Chen, W.-Y.; Matulis, D. Repeatability, precision, and accuracy of the enthalpies and Gibbs energies of a protein–ligand binding reaction measured by isothermal titration calorimetry. *European Biophysics Journal* **2019**, *48*, 139–152.

(314) Torsi, L.; Magliulo, M.; Manoli, K.; Palazzo, G. Organic field-effect transistor sensors: a tutorial review. *Chemical Society Reviews* **2013**, *42* (22), 8612–8628.

(315) Reigner, B.; Blesch, K.; Weidekamm, E. Clinical pharmacokinetics of capecitabine. *Clinical Pharmacokinetics* **2001**, *40*, 85–104.

(316) Yan, F.; Chu, Y.; Zhang, K.; Zhang, F.; Bhandari, N.; Ruan, G.; Dai, Z.; Liu, Y.; Zhang, Z.; Kan, A. T. Determination of adsorption isotherm parameters with correlated errors by measurement error models. *Chemical Engineering Journal* **2015**, *281*, 921–930.

(317) Stingeder, G. Optimization of secondary ion mass spectrometry for quantitative trace analysis. *Analytica Chimica Acta* **1994**, *297* (1), 231–251.

(318) Ritzmann, D.; Wright, P. S.; Holderbaum, W.; Potter, B. A method for accurate transmission line impedance parameter estimation. *IEEE Transactions on instrumentation and measurement* **2016**, *65* (10), 2204–2213.

(319) Sursyakova, V. V.; Levdansky, V. A.; Rubaylo, A. I. Determination of binding constants for strong complexation by affinity capillary electrophoresis: the example of complexes of ester betulin derivatives with (2-hydroxypropyl)-γ-cyclodextrin. *Analytical and Bioanalytical Chemistry* **2020**, *412*, 5615–5625.

(320) Kohlmann, P.; Krylov, S. N.; Marchand, P.; Jose, J. FRET Assays for the Identification of C. albicans HSP90-Sba1 and Human HSP90α-p23 Binding Inhibitors. *Pharmaceuticals* **2024**, *17* (4), 516.

(321) Andersson, L. O.; Rehnström, A.; Eaker, D. L. Studies on "nonspecific" binding: The nature of the binding of fluorescein to bovine serum albumin. *European Journal of Biochemistry* **1971**, *20* (3), 371–380.

(322) Pinger, C. W.; Castiaux, A.; Speed, S.; Spence, D. M. Plate reader compatible 3D-printed device for teaching equilibrium dialysis binding assays. *Journal of Chemical Education* **2018** *95* (9), 1662–1667.

(323) Munson, M. S.; Meacham, J. M.; Locascio, L. E.; Ross, D. Counterflow rejection of adsorbing proteins for characterization of biomolecular interactions by temperature gradient focusing. *Analytical Chemistry* **2008**, *80* (1), 172–178.

(324) Pearl, L. H. Hsp90 and Cdc37–a chaperone cancer conspiracy. *Current Opinion in Genetics & Development* **2005**, *15* (1), 55–61.

(325) Wang, L.; Zhang, Q.; You, Q. Targeting the HSP90–CDC37–kinase chaperone cycle: A promising therapeutic strategy for cancer. *Medicinal Research Reviews* **2022**, *42* (1), 156–182.

(326) Miyata, Y.; Nakamoto, H.; Neckers, L. The therapeutic target Hsp90 and cancer hallmarks. *Current Pharmaceutical Design* **2013**, *19* (3), 347–365.

(327) Chapman, J.; Ismail, A. E.; Dinu, C. Z. Industrial applications of enzymes: Recent advances, techniques, and outlooks. *Catalysts* **2018**, *8* (6), 238.

(328) Kalp, M.; Sheri, A.; Buynak, J. D.; Bethel, C. R.; Bonomo, R. A.; Carey, P. R. Efficient inhibition of class A and class D β-lactamases by Michaelis complexes. *Journal of Biological Chemistry* **2007**, *282* (30), 21588–21591.

(329) Badarau, A.; Shi, Q.; Chow, J. W.; Zajicek, J.; Mobashery, S.; Vakulenko, S. Aminoglycoside 2 ″-phosphotransferase type IIIa from Enterococcus. *Journal of Biological Chemistry* **2008**, *283* (12), 7638–7647.

(330) Simeth, N. A.; Kinateder, T.; Rajendran, C.; Nazet, J.; Merkl, R.; Sterner, R.; König, B.; Kneuttinger, A. C. Towards photochromic azobenzene-based inhibitors for tryptophan synthase. *Chemistry–A European Journal* **2021**, *27* (7), 2439–2451.

(331) Balachandran, N.; Heimhalt, M.; Liuni, P.; To, F.; Wilson, D. J.; Junop, M. S.; Berti, P. J. Potent inhibition of 3-deoxy-d-arabinoheptulosonate-7-phosphate (DAHP) synthase by DAHP oxime, a phosphate group mimic. *Biochemistry* **2016**, *55* (48), 6617–6629.

(332) Panchagnula, R.; Thomas, N. S. Biopharmaceutics and pharmacokinetics in drug research. *International Journal of Pharmaceutics* **2000**, *201* (2), 131–150.

(333) Mendes, P.; Kell, D. Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics (Oxford, England)* **1998**, *14* (10), 869–883.

(334) Govind, R.; Gao, C.; Laai, L.; Yan, X.; Pfanstiel, S.; Tabak, H. Development of methodology to determine the bioavailability and biodegradation kinetics of toxic organic pollutant compounds in soil. *Applied Biotechnology for Site Remediation* **1994**, 229–239.

(335) Chandrasekaran, M. *Enzymes in food and beverage processing*; CRC Press Boca Raton, FL, USA, 2016.

(336) Csöoregi, E.; Gáspñr, S.; Niculescu, M.; Mattiasson, B.; Schuhmann, W. Amperometric enzyme-based biosensors for application in food and beverage industry. *Physics and Chemistry Basis of Biotechnology* **2002**, 105–129.

(337) Atkins, G.; Nimmo, I. Current trends in the estimation of Michaelis-Menten parameters. *Analytical Biochemistry* **1980**, *104* (1), 1–9.

(338) Sagnella, G. Model fitting, parameter estimation, linear and non-linear regression. *Trends in Biochemical Sciences* **1985**, *10* (3), 100–103.

(339) Hill, C. M.; Waightm, R. D.; Bardsley, W. G. Does any enzyme follow the Michaelis—Menten equation? *Molecular and Cellular Biochemistry* **1977**, *15*, 173–178.

(340) Cho, Y.-S.; Lim, H.-S. Comparison of various estimation methods for the parameters of Michaelis-Menten equation based on in vitro elimination kinetic simulation data. *Translational and Clinical Pharmacology* **2018**, *26* (1), 39.

(341) Buchholz, P. C.; Ohs, R.; Spiess, A. C.; Pleiss, J. Progress curve analysis within BioCatNet: Comparing kinetic models for enzyme-catalyzed self-ligation. *Biotechnology Journal* **2019**, *14* (3), 1800183.

(342) Schnell, S. Validity of the Michaelis–Menten equation–steady-state or reactant stationary assumption: that is the question. *The FEBS Journal* **2014**, *281* (2), 464–472.

(343) Bentz, J.; Tran, T. T.; Polli, J. W.; Ayrton, A.; Ellens, H. The steady-state Michaelis–Menten analysis of P-glycoprotein mediated transport through a confluent cell monolayer cannot predict the correct Michaelis constant Km. *Pharmacological Research* **2005**, *22*, 1667–1677.

(344) Davidian, M.; Giltinan, D. M. Nonlinear models for repeated measurement data: an overview and update. *Journal of Agricultural, Biological, and Environmental Statistics* **2003**, *8*, 387–419.

(345) Kleinbaum, D. G.; Kupper, L. L.; Muller, K. E.; Nizam, A. *Applied Regression Analysis and Other Multivariable Methods*; Duxbury press Belmont, CA, 1988.

(346) Udema, I. I. Rate constants are determinable outside the original Michaelis–Menten mathematical formalism wherein the substrate concentration range is approx. 1.6 to 4.8 times enzyme concentration: A pre-steady-state scenario and beyond. *World Journal of Advanced Research and Reviews* **2022**, *16* (1), 350–367.

(347) Mathis, J. R.; Back, K.; Starks, C.; Noel, J.; Poulter, C. D.; Chappell, J. Pre-steady-state study of recombinant sesquiterpene cyclases. *Biochemistry* **1997**, *36* (27), 8340–8348.

(348) Tzafriri, A. R. Michaelis-Menten kinetics at high enzyme concentrations. *Bulletin of Mathematical Biology* **2003**, *65* (6), 1111–1129.

(349) Cha, S. Kinetic behavior at high enzyme concentrations: magnitude of errors of Michaelis-Menten and other approximations. *Journal of Biological Chemistry* **1970**, *245* (18), 4814–4818.

(350) Schnell, S.; Maini, P. K. Enzyme kinetics at high enzyme concentration. *Bulletin of Mathematical Biology* **2000**, *62*, 483–499.

(351) Little, T. Method validation essentials, limit of blank, limit of detection, and limit of quantitation. *BioPharm International* **2015**, *28* (4).

(352) Currie, D. J. Estimating Michaelis-Menten parameters: bias, variance and experimental design. *Biometrics* **1982**, 907–919.

(353) Matthews, J.; Allcock, G. Optimal designs for Michaelis–Menten kinetic studies. *Statistics in Medicine* **2004**, *23* (3), 477–491.

(354) Yun, K.-I.; Han, T.-S. Relationship between enzyme concentration and Michaelis constant in enzyme assays. *Biochimie* **2020**, *176*, 12–20.

(355) Segel, L. A. On the validity of the steady state assumption of enzyme kinetics. *Bulletin of Mathematical Biology* **1988**, *50*, 579–593.

(356) Yin, W.-B.; Grundmann, A.; Cheng, J.; Li, S.-M. Acetylaszonalenin biosynthesis in Neosartorya fischeri: Identification of the biosynthetic gene cluster by genomic mining and functional proof of the genes by biochemical investigation. *Journal of Biological Chemistry* **2009**, *284* (1), 100–109.

(357) Passonneau, J. V.; Lowry, O. H. *Enzymatic analysis: a practical guide*; Springer Science & Business Media, 2008.

(358) Tiffany, T.; Thayer, P.; Coelho, C.; Manning, G. A propagation of error analysis of the enzyme activity expression. A model for determining the total system random error of a kinetic enzyme analyzer. *Clinical Chemistry* **1976**, *22* (9), 1438–1450.

(359) Lyn, J. A.; Ramsey, M. H.; Fussell, R. J.; Wood, R. Measurement uncertainty from physical sample preparation: estimation including systematic error. *Analyst* **2003**, *128* (11), 1391–1398.

(360) Schnell, S.; Hanson, S. M. A test for measuring the effects of enzyme inactivation. *Biophysical Chemistry* **2007**, *125* (2–3), 269–274.

(361) Kruger, N. J. Errors and artifacts in coupled spectrophotometric assays of enzyme activity. *Phytochemistry* **1995**, *38* (5), 1065–1071.

(362) Helmerhorst, E.; Chandler, D. J.; Nussio, M.; Mamotte, C. D. Real-time and label-free bio-sensing of molecular interactions by surface plasmon resonance: a laboratory medicine perspective. *The Clinical Biochemist Reviews* **2012**, *33* (4), 161.

(363) Peltomaa, R.; Glahn-Martínez, B.; Benito-Peña, E.; Moreno-Bondi, M. C. Optical biosensors for label-free detection of small molecules. *Sensors* **2018**, *18* (12), 4126.

(364) Plach, M.; Schubert, T. Biophysical characterization of aptamer-target interactions. *Aptamers in Biotechnology* **2020**, 1–15.

(365) Kawski, A. Fluorescence anisotropy: theory and applications of rotational depolarization. *Critical Reviews in Analytical Chemistry* **1993**, *23* (6), 459–529.

(366) Barbero, N.; Napione, L.; Quagliotto, P.; Pavan, S.; Barolo, C.; Barni, E.; Bussolino, F.; Viscardi, G. Fluorescence anisotropy analysis of protein–antibody interaction. *Dyes and Pigments* **2009**, *83* (2), 225–229.

(367) Houtman, J. C.; Brown, P. H.; Bowden, B.; Yamaguchi, H.; Appella, E.; Samelson, L. E.; Schuck, P. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Science* **2007**, *16* (1), 30–42.

(368) Kantonen, S. A.; Henriksen, N. M.; Gilson, M. K. Evaluation and Minimization of Uncertainty in ITC Binding Measurements: Heat Error, Concentration Error, Saturation, and Stoichiometry. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2017**, *1861* (2), 485–498.

(369) Zhao, H.; Piszczek, G.; Schuck, P. SEDPHAT–a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* **2015**, *76*, 137–148.

(370) Poon, G. M. Explicit formulation of titration models for isothermal titration calorimetry. *Analytical Biochemistry* **2010**, *400* (2), 229–236.

(371) Guide, T. ITC Data Analysis in Origin®. https://bif.wiscweb.wisc.edu/wp-content/uploads/sites/389/2017/11/ITC_Data_Analysis_in_Origin.pdf

(372) Grüner, S.; Neeb, M.; Barandun, L. J.; Sielaff, F.; Hohn, C.; Kojima, S.; Steinmetzer, T.; Diederich, F.; Klebe, G. Impact of protein and ligand impurities on ITC-derived protein–ligand thermodynamics. *Biochimica et Biophysica Acta (BBA)-General Subjects* **2014**, *1840* (9), 2843–2850.

(373) Broecker, J.; Vargas, C.; Keller, S. Revisiting the optimal c value for isothermal titration calorimetry. *Analytical Biochemistry* **2011**, *418* (2), 307–309.

(374) Perozzo, R.; Folkers, G.; Scapozza, L. Thermodynamics of protein–ligand interactions: history, presence, and future aspects. *Journal of Receptors and Signal Transduction* **2004**, *24* (1–2), 1–52.

(375) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry* **1989**, *179* (1), 131–137.

(376) Dutta, A. K.; Rösgen, J.; Rajarathnam, K. Using isothermal titration calorimetry to determine thermodynamic parameters of protein–glycosaminoglycan interactions. In *Glycosaminoglycans*, Springer, 2015; pp 315–324.

(377) Matsuyama, B. Y.; Krasteva, P. V.; Navarro, M. V. Isothermal titration calorimetry to determine apparent dissociation constants (K_d) and stoichiometry of interaction (n) of C-di-GMP binding proteins. *c-di-GMP Signaling: Methods and Protocols* **2017**, 403–416.

(378) Tellinghuisen, J. Optimizing isothermal titration calorimetry protocols for the study of 1: 1 binding: Keeping it simple. *Biochimica et Biophysica Acta (BBA)-General Subjects* **2016**, *1860* (5), 861–867.

(379) Hansen, L. D.; Quinn, C. Obtaining precise and accurate results by ITC. *European Biophysics Journal* **2019**, *48*, 825–835.

(380) Broecker, J.; Vargas, C.; Keller, S. Revisiting the optimal cvalue for isothermal titration calorimetry. *Analytical Biochemistry* **2011**, *418* (2), 307–309.

(381) Tellinghuisen, J. Isothermal titration calorimetry at very low c. *Analytical Biochemistry* **2008**, *373* (2), 395–397.

(382) Hansen, L. D.; Fellingham, G. W.; Russell, D. J. Simultaneous determination of equilibrium constants and enthalpy changes by titration calorimetry: Methods, instruments, and uncertainties. *Analytical Biochemistry* **2011**, *409* (2), 220–229.

(383) Le, V. H.; Buscaglia, R.; Chaires, J. B.; Lewis, E. A. Modeling complex equilibria in ITC experiments: thermodynamic parameters estimation for a three binding site model. *Analytical Biochemistry* **2012**, *434* (2), 233.

(384) Bastos, M.; Velazquez-Campoy, A. Isothermal titration calorimetry (ITC): a standard operating procedure (SOP). *European Biophysics Journal* **2021**, *50*, 363–371.

(385) Duff, M. R.; Grubbs, J.; Howell, E. E. Isothermal titration calorimetry for measuring macromolecule-ligand affinity. *Journal of Visualized Experiments: JoVE* **2011**, (55), 2796.

(386) Matilla, M. A.; Martín-Mora, D.; Krell, T. The use of isothermal titration calorimetry to unravel chemotactic signalling mechanisms. *Environmental Microbiology* **2020**, *22* (8), 3005–3019.

(387) Liu, X.; Hou, Y.; Chen, S.; Liu, J. Controlling dopamine binding by the new aptamer for a FRET-based biosensor. *Biosensors and Bioelectronics* **2021**, *173*, 112798.

(388) Huang, P.-J. J.; Liu, J. A DNA aptamer for theophylline with ultrahigh selectivity reminiscent of the classic RNA aptamer. *ACS Chemical Biology* **2022**, *17* (8), 2121–2129.

(389) Burnouf, D.; Ennifar, E.; Guedich, S.; Puffer, B.; Hoffmann, G.; Bec, G.; Disdier, F. o.; Baltzinger, M.; Dumas, P. kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *Journal of the American Chemical Society* **2012**, *134* (1), 559–565.

(390) Schwarz, F. P.; Reinisch, T.; Hinz, H.-J.; Surolia, A. Recommendations on measurement and analysis of results obtained on biological substances using isothermal titration calorimetry (IUPAC Technical Report). *Pure and Appllied Chemistry* **2008**, *80* (9), 2025–2040.

(391) Krishnamoorthy, J.; Mohanty, S. Open-ITC: an alternate computational approach to analyze the isothermal titration calorimetry data of complex binding mechanisms. *Journal of Molecular Recognition* **2011**, *24* (6), 1056–1066.

(392) Tellinghuisen, J. Stupid statistics! *Methods in Cell Biology* **2008**, *84*, 737–780.

(393) Tellinghuisen, J. Calibration in isothermal titration calorimetry: Heat and cell volume from heat of dilution of NaCl(aq). *Analytical Biochemistry* **2007**, *360* (1), 47–55.

(394) Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angewandte Chemie International Edition* **2002**, *41* (15), 2644–2676.

(395) Kaiyum, Y. A.; Hoi Pui Chao, E.; Dhar, L.; Shoara, A. A.; Nguyen, M. D.; Mackereth, C. D.; Dauphin-Ducharme, P.; Johnson, P. E. Ligand-Induced Folding in a Dopamine-Binding DNA Aptamer. *ChemBioChem* **2024**, *25* (23), e202400493.

Appendices

Appendix A. Supplementary Information for Chapter 2

Appendix A1. Supplementary Figures for "Template Instrumentation for 'Accurate Constant via Transient Incomplete Separation' (ACTIS)"

Figure A1.1: Determination of K_d Value of the BSA–Fluorescein Pair on Day 2

For the parallel experiments, identical samples containing 100 nM fluorescein and different concentrations of BSA were used. The left panels show representative separagrams for each concentration of BSA, and the right panels show the corresponding binding isotherms obtained for all of the separagrams. *R* is the fraction of unbound fluorescein.

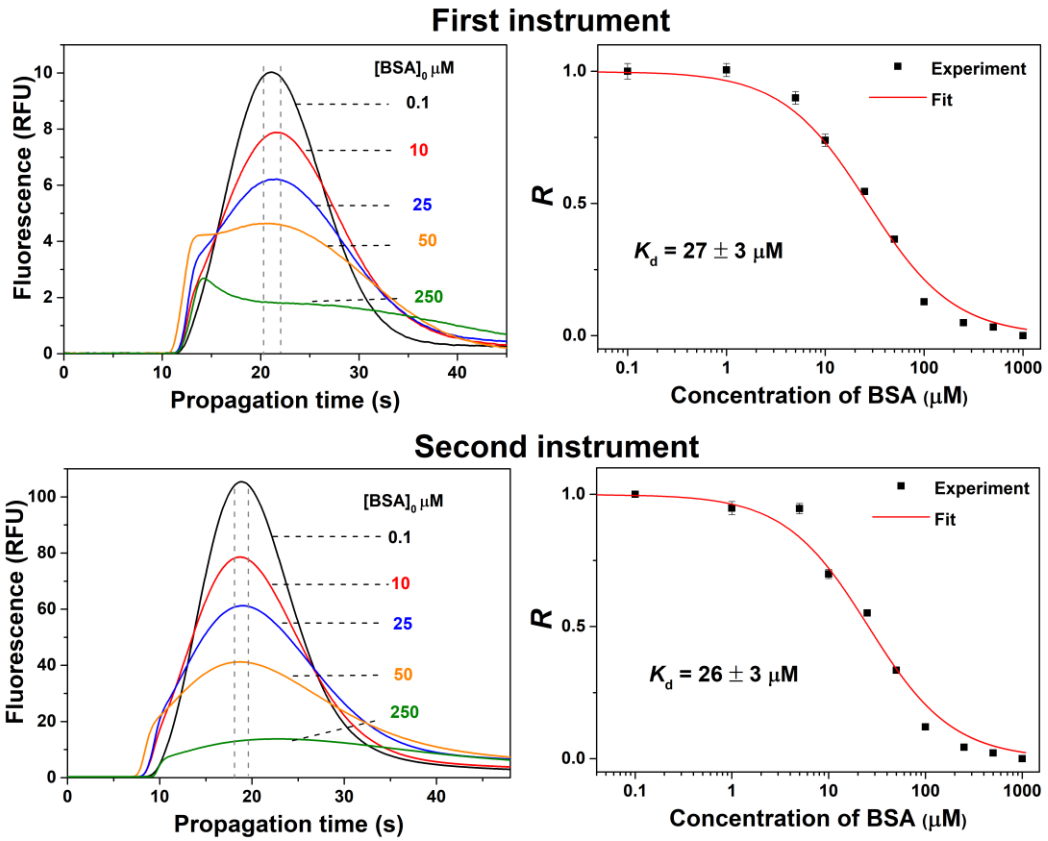


Figure A1.2: Determination of K_d Value of the BSA–Fluorescein Pair on Day 3

For the parallel experiments, identical samples containing 100 nM fluorescein and different concentrations of BSA were used. The left panels show representative separagrams for each concentration of BSA, and the right panels show the corresponding binding isotherms obtained for all of the separagrams. *R* is the fraction of unbound fluorescein.

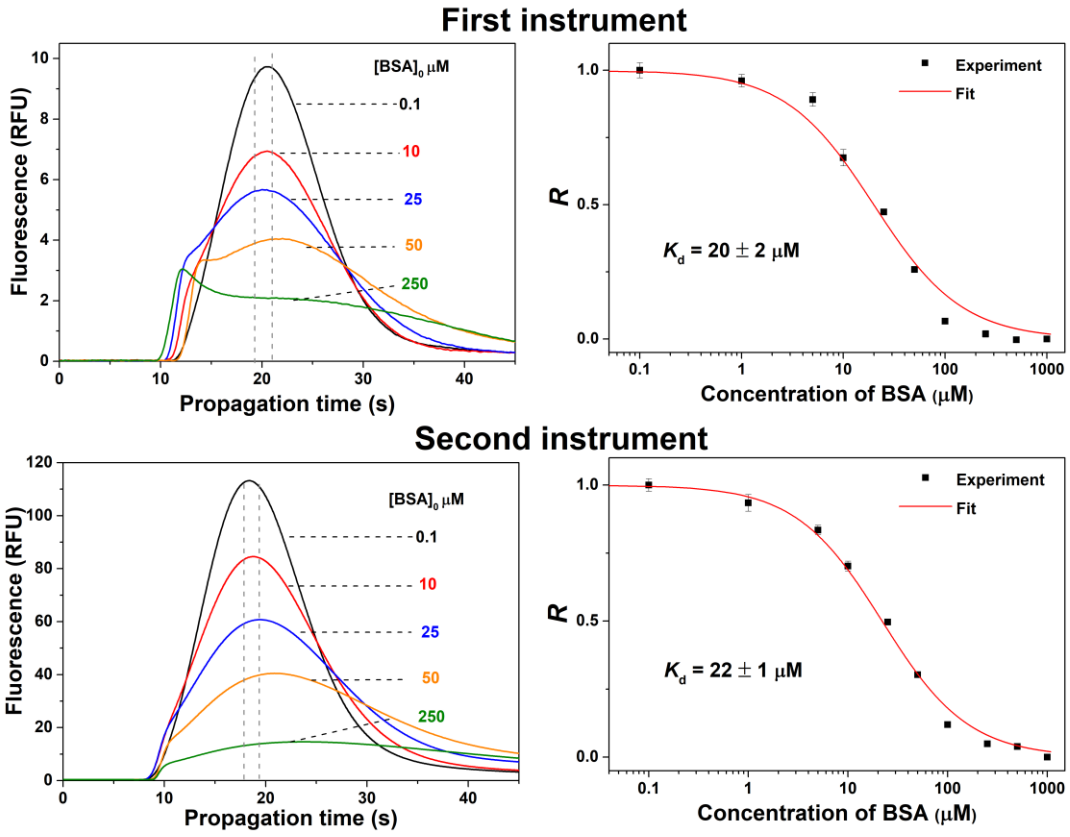


Figure A1.3: Determination of K_d Value of the BSA–Fluorescein Pair on Day 4

For the parallel experiments, identical samples containing 200 nM fluorescein and different concentrations of BSA were used. The left panels show representative separagrams for each concentration of BSA, and the right panels show the corresponding binding isotherms obtained for all of the separagrams. *R* is the fraction of unbound fluorescein.

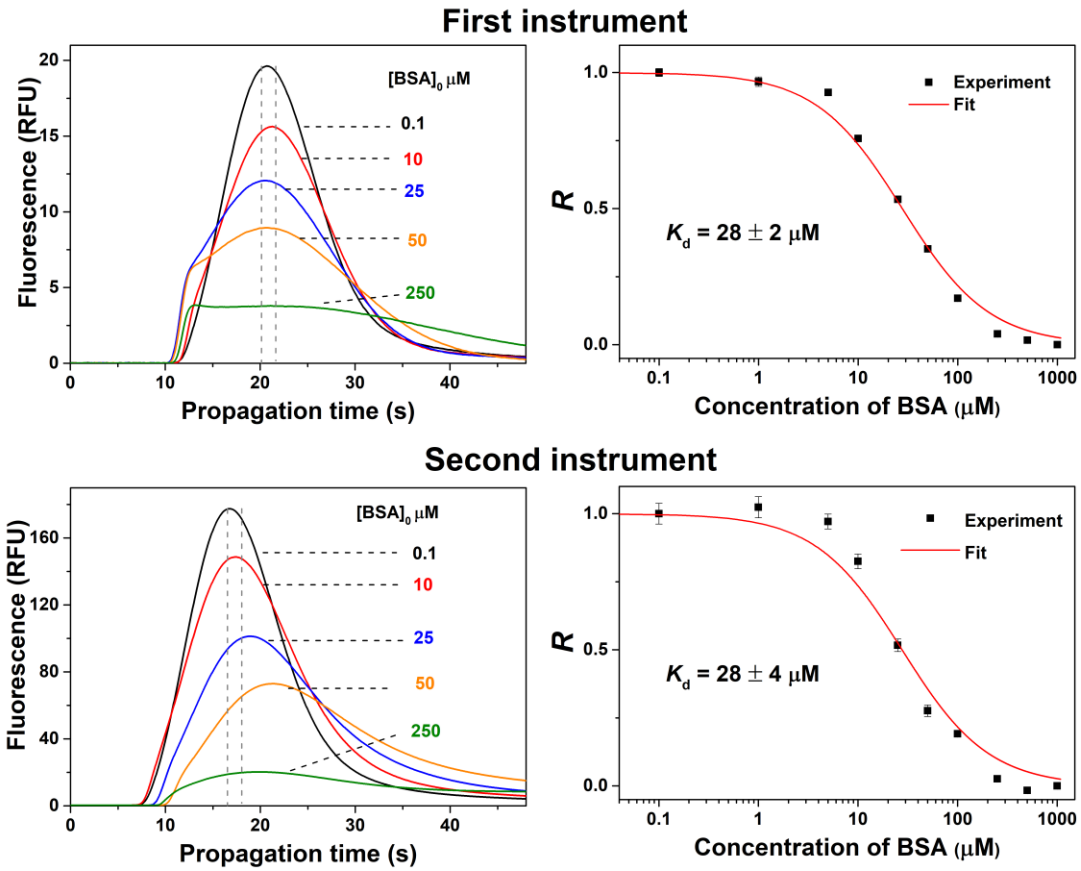


Figure A1.4: Determination of K_d Value of the BSA–Fluorescein Pair on Day 5

For the parallel experiments, identical samples containing 100 nM fluorescein and different concentrations of BSA were used. The left panels show representative separagrams for each concentration of BSA, and the right panels show the corresponding binding isotherms obtained for all of the separagrams. *R* is the fraction of unbound fluorescein. A new batch of BSA was used for this experiment, while all other experiments were performed with a single other batch.

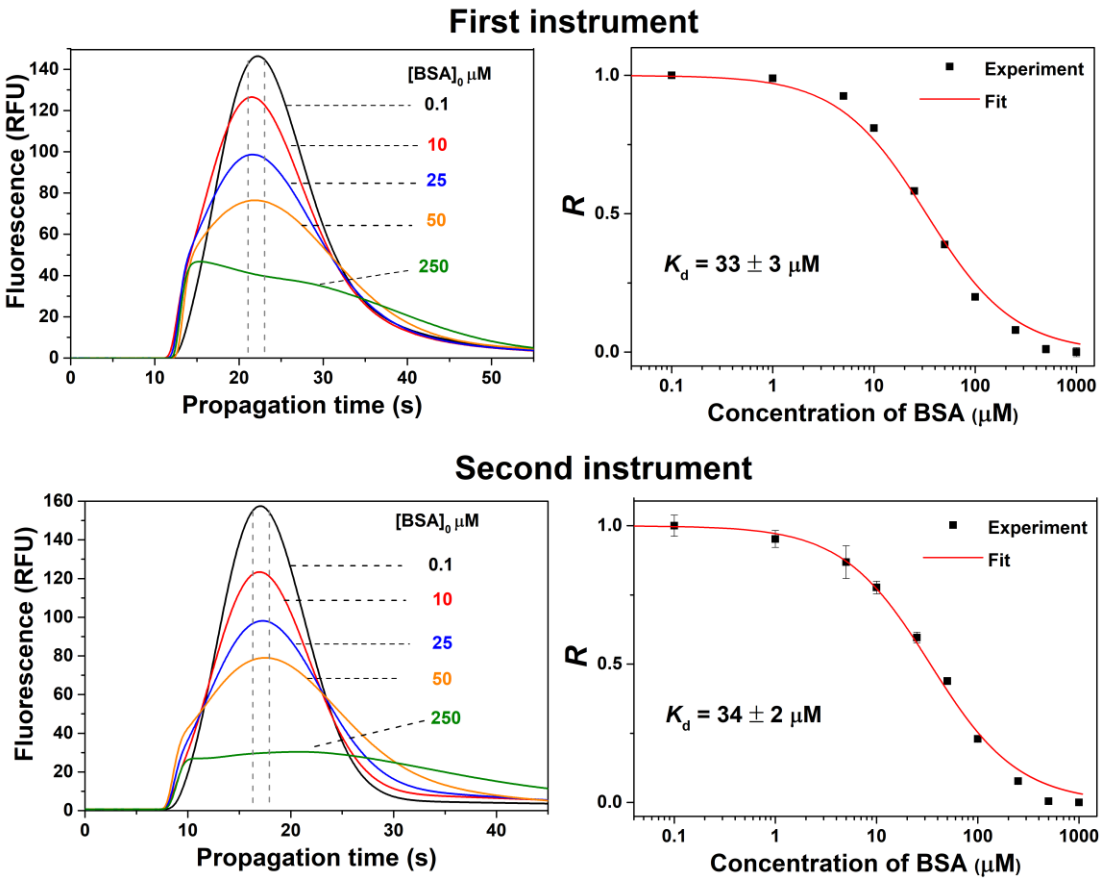


Figure A1.5: Repeatability of the ACTIS Separagrams for the BSA-Fluorescein Complex, Day 1, Instrument 1

For all the equilibrium mixtures, the total concentration of fluorescein was 100 nM, while the total concentrations of BSA varied from 0.1 to 1,000 µM as shown in the figure.

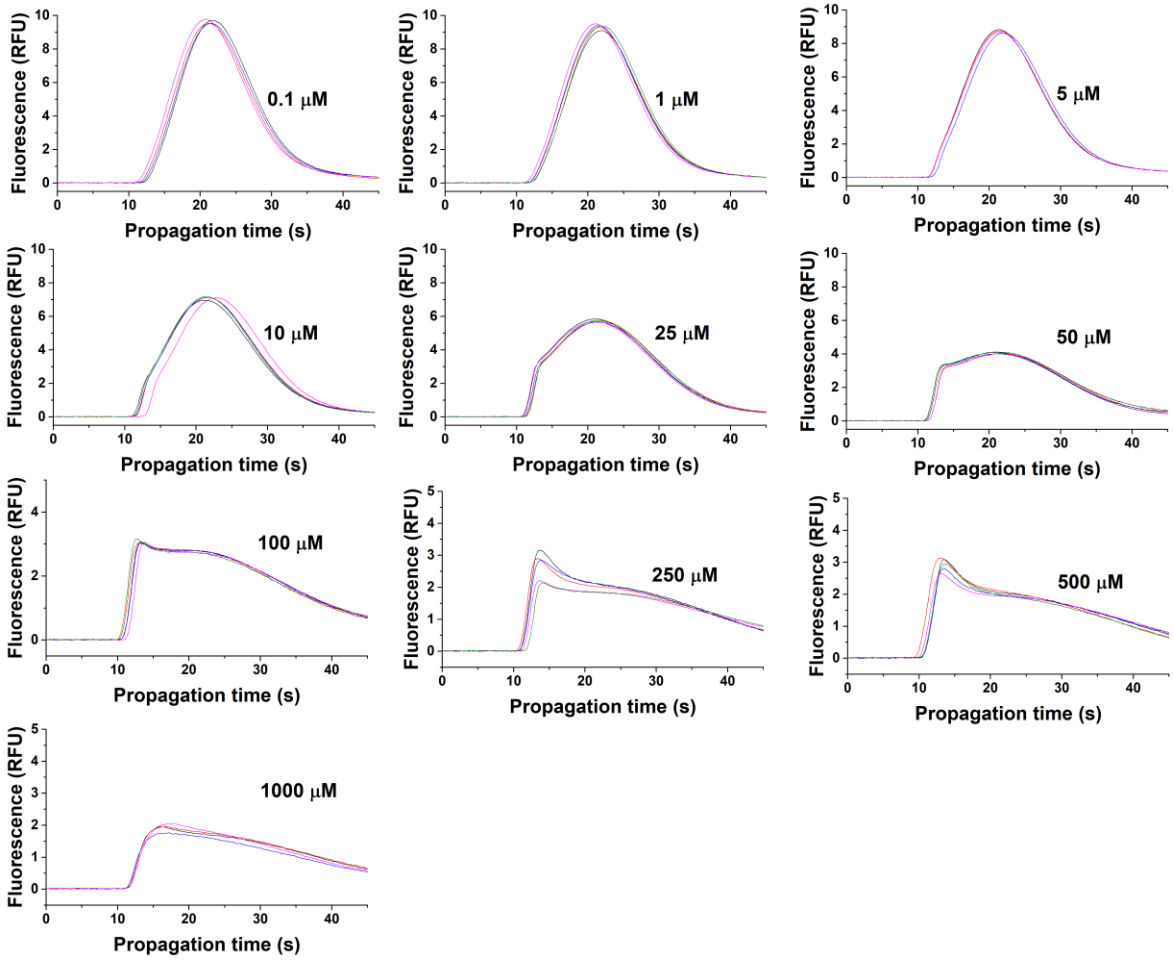
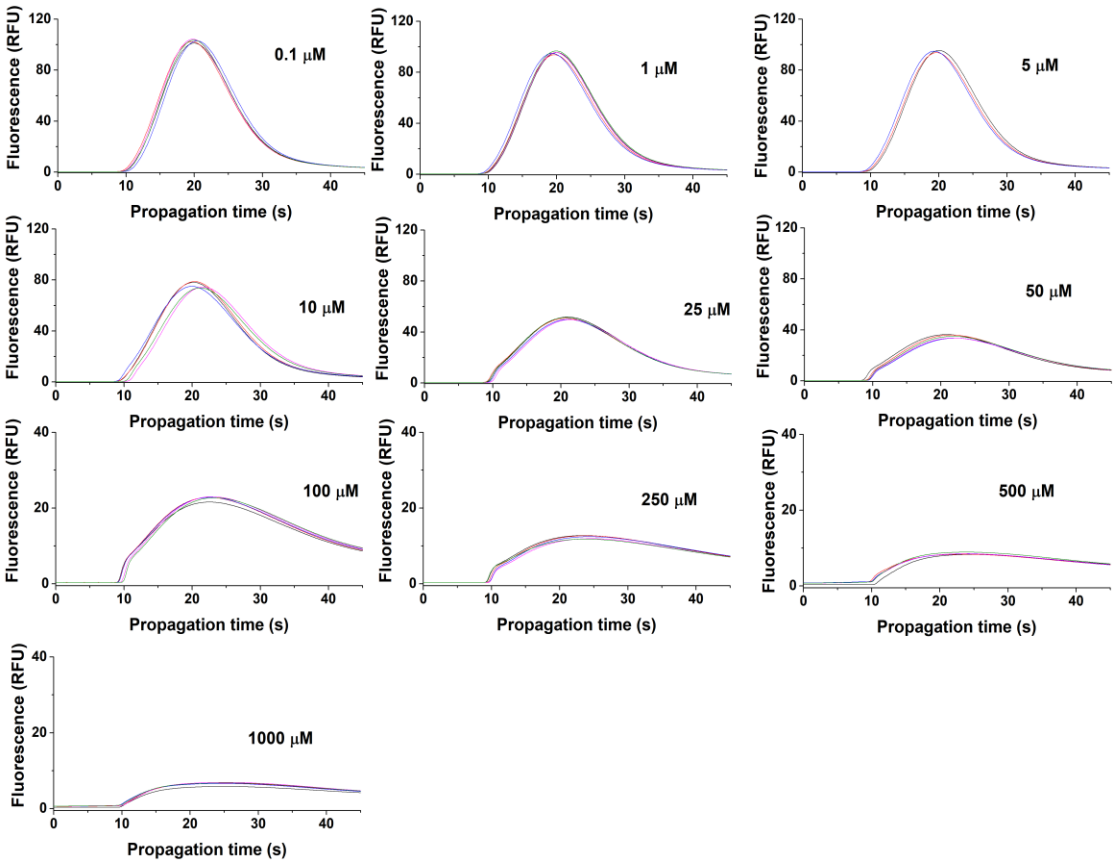


Figure A1.6: Repeatability of the ACTIS Separagrams for the BSA-Fluorescein Complex, Day 1, Instrument 2

For all the equilibrium mixtures, the total concentration of fluorescein was 100 nM, while the total concentrations of BSA varied from 0.1 to 1,000 µM as shown in the figure.



Appendix A2. Supporting Information for "Transient Incomplete Separation of Species with Close Diffusivity to Study Stability of Affinity Complexes"

Note A2.1: Detailed Settings for COMSOL Simulations

All virtual experiments were simulated with COMSOL Multiphysics software version 5.6a. "Transport of diluted species and Laminar Flow" modules were used, and the simulation files can be found in the Supporting files (**COMSOL.zip** saved on Figshare: https://doi.org/10.6084/m9.figshare.20173283). The geometry of the virtual ACTIS instrument is depicted in Scheme A2.1. In all virtual experiments, we considered a ligand L with typical diffusion coefficient of μ_L = 100 $\mu m^2/s$ and a target T with diffusion coefficient ranging from μ_T = 50 to 95 $\mu m^2/s$. The remaining parameters were: $k_{on,inp} = 10^5\ M^{-1}s^{-1}$, $k_{off,inp} = 5 \times 10^{-5}\ s^{-1}$, $K_d = k_{off,inp} / k_{on,inp}$ = 0.5 nM, L_0 = 0.02 nM, T = 300 K. T_0 was varied from 0.01 to 50 nM using 11 different non-zero concentrations plus the zero concentration. The geometry used in the experiments was a mimic of a real experimental setup. The optimal volumetric flow rate Q_{TIS} was chosen so that the detection time (t_{det}) (defined as the migration time of the diffusive peak maximum on the separagram) was approximately equal to the characteristic diffusion time (τ_L) of L:

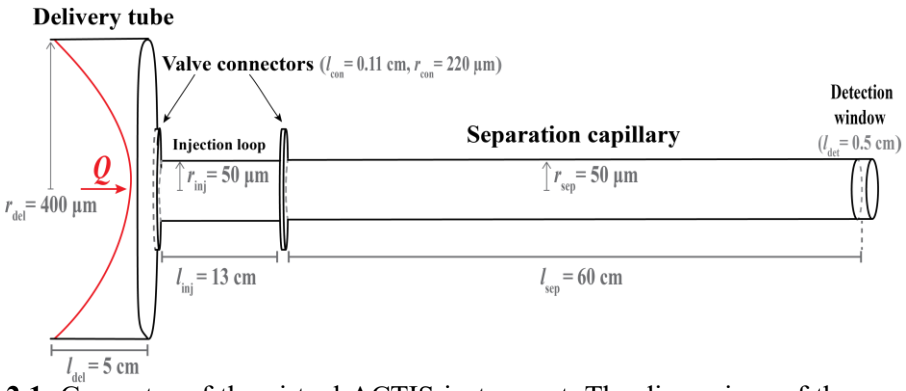
$$t_{det} \approx \tau_L \tag{A1}$$

The characteristic diffusion time is defined as:

$$t_L = a^2 / \mu_L \tag{A2}$$

where a is the inner radius of the separation capillary and m_L is the diffusion coefficient of L. Also, based on the definition of volumetric flow rate, we know:

$$Q_{TIS} = \pi\, a^2\, l / t_{det} \tag{A3}$$



Scheme A2.1. Geometry of the virtual ACTIS instrument. The dimensions of the components are described in the schematic.

where l is the distance between the position of the sample plug after sample transfer into the separation capillary and the detection window on this capillary.

The above three equations result in the following ratio:

$$l/Q_{TIS} \approx 1/(\pi\mu_L) \qquad \text{(A4)}$$

Given a ligand diffusion coefficient of $\mu_L \approx 100$ µm^2/s, with the separation capillary with an inner radius of a = 50 µm. According to **Eqs (A1)** and **(A2)**, the detection time will be:

$$t_{det} = (50\ \mu m)^2/(100\ \mu m^2/s) = 25\ s \qquad \text{(A5)}$$

After the process of sample transfer, the distance between the centre of sample plug and the centre of detector is l = 40 cm; therefore, based on **Eq (A4)**, the flow rate was calculated and set to be:

$$Q_{TIS} \approx \pi\mu_L l = 8\ \mu L/min \qquad \text{(A6)}$$

Note that, ACTIS is very robust to variations in the flow rate, which allows one to use rough approximations for μ_L and l.

Note A2.2: Procedure for Preparation of a Series of Equilibrium Mixtures

To guarantee that L_0 is highly consistent in all equilibrium mixtures (EMs), we introduced a new workflow of sample preparation for each ACTIS titration experiment. The general idea of this workflow is to use the identical ligand solution to dilute the EM with the highest protein concentration in series. By discarding the EM with the highest protein concentration, the remaining EMs are used as the samples. Here, as an example, we show the procedure of preparing 120 µL EMs of 0.05 nM aptamer with the addition of 0–10 nM MutS.

Buffer: 50 mM Tris-acetate pH 8.2 with 0.05% Triton™ X-100

Stock solutions:

DNA stock solution: 1.0 µM Alexa488-labelled aptamer

Protein stock solution: 10.8 µM His-tagged MutS

Steps:

1. Prepare 350 µL of 0.1 nM aptamer and 1,300 µL of 0.05 nM aptamer:
    a. 16 µL buffer + 4 µL of 1.0 µM aptamer = 20 µL of 200 nM aptamer, then annealing it with a Thermocycler (melted at 95 ºC for 2 min and gradually cooled down to 20 ºC at a rate of 0.5 ºC/s).
    b. 95 µL buffer + 5 µL of 200 nM aptamer = 100 µL of 10 nM aptamer.
    c. 990 µL buffer + 10 µL of 10 nM aptamer = 1000 µL of 0.1 nM aptamer.
    d. 650 µL buffer + 650 µL of 0.1 nM aptamer = 1300 µL of 0.05 nM aptamer.
2. Prepare 120 µL of 0.05 nM aptamer + 0, 0.010, 0.039, 0.156, 0.625, 2.5, and 10 nM MutS:
    a. 119 µL buffer + 135 µL of 0.1 nM aptamer + 16 µL of 10.8 µM MutS = 270 µL of 0.05 nM aptamer + 640 nM MutS.

b. In 9 sample tubes, add 120 µL of 0.05 nM aptamer to each tube. Then, dilute the protein concentration 4 times in series by adding 40 µL of sample from the previous tube. After the dilutions, we get 120 µL of 0.05 nM aptamer + 0, 0.039, 0.156, 0.625, 2.5, 10, 40, and 160 nM MutS. We also get 160 µL of 0.05 nM aptamer + 0.010 nM MutS, and 230 µL of 0.05 nM aptamer + 640 nM MutS.

c. Pipet 40 µL solution from the "0.05 nM aptamer + 0.010 nM MutS" sample and discard it to make the sample volume at each protein concentration to be consistent to avoid the concentration bias caused by evaporation and sample adsorption to the sample vials.

3. Incubate the samples in a dark box for 2 h.

With this new workflow, we conducted a mock titration experiment with 0.05 nM aptamer at 0–10 nM mock protein concentration (see **Figures A2.4** and **A2.5** for the results).

Note A2.3: Calculation of Weighted Average K_d Value of The Muts–Aptamer Complex

In the three ACTIS titration experiments, the K_d values for the MutS–aptamer complex were determined to be $K_{d,1} = 0.17 \pm 0.02$ nM, $K_{d,2} = 0.19 \pm 0.05$, and $K_{d,3} = 0.11 \pm 0.02$ nM.

The weighted average value $K_{d,ave}$ and its uncertainty $\delta K_{d,ave}$ were calculated as follows:

$$K_{\mathrm{d,ave}} = \sum_{i=1}^{3} \frac{K_{\mathrm{d},i}}{\delta K_{\mathrm{d},i}^{\,2}} \Bigg/ \sum_{i=1}^{3} \frac{1}{\delta K_{\mathrm{d},i}^{\,2}} = 0.14 \text{ nM} \tag{A7}$$

$$\delta K_{\mathrm{d,ave}} = \sqrt{\frac{1}{\sum_{i=1}^{3} \delta K_{\mathrm{d},i}^{\,-2}}} = 0.01 \text{ nM} \tag{A8}$$

Therefore, the weighted average value $K_{d,ave}$ of the MutS–aptamer complex is 0.14 ± 0.01 nM.

Figure A2.1: Optimized ACTIS Instrumentation and Experimental Procedure

ACTIS instrumentation setup and the four steps of a single run in an ACTIS experiment. The red arrows show the directions of the flows. The orange arrow shows the running direction of the syringe pump. See the main text for more details.

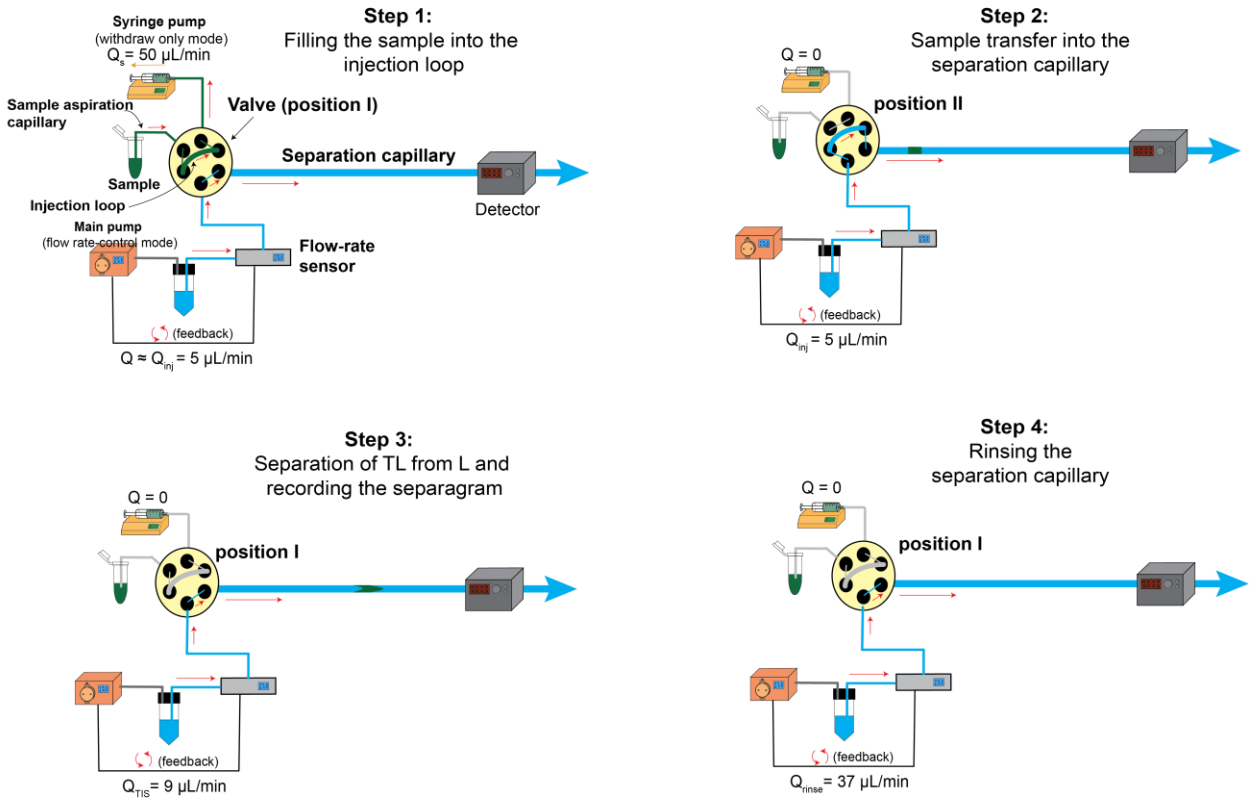


Figure A2.2: Sample Calculation for Dynamic Range

The concentration profiles for different μ_{TL}/μ_L values were produced by COMSOL. Then, the dynamic range corresponding to certain μ_{TL}/μ_L value (taking μ_{TL}/μ_L=0.5 as an example) was calculated by the following procedure: (1) Measure the maximum signal $S_{1,max}$ of the diffusive peak for μ_{TL}/μ_L=1; (2) Measure the maximum signal $S_{0.5,max}$ of the diffusive peak for μ_{TL}/μ_L=0.5; (3) Dynamic range = $[(S_{1,max} - S_{0.5,max})/ S_{1,max}] \times 100\%$. For μ_{TL}/μ_L=0.5, Dynamic range = = [(5.16 - 3.84)/ 5.16] ×100% = 25.6%.

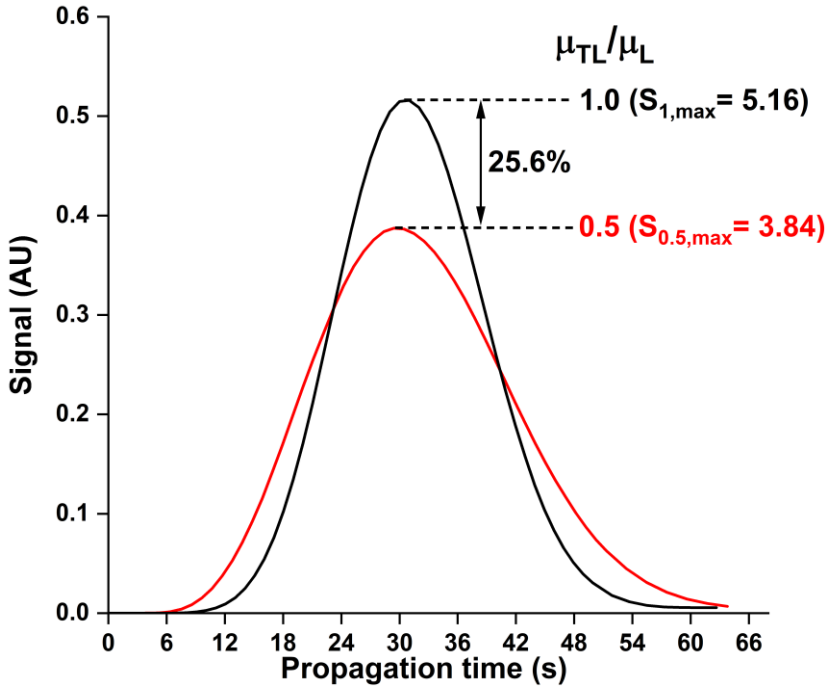
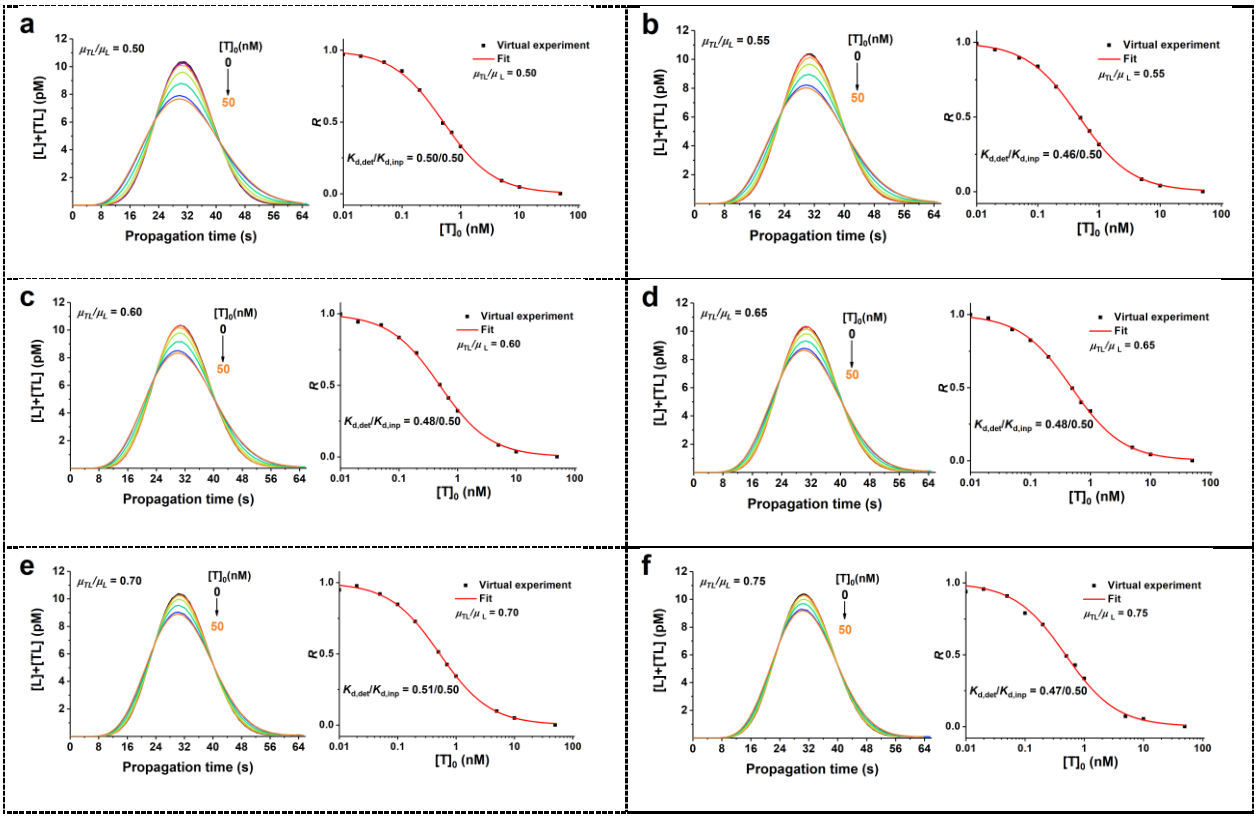


Figure A2.3: The Virtual Experiments with Varied μ_{TL}/μ_L without Noise

Separagrams and their respective binding isotherms for the virtual ACTIS experiments with varying μ_{TL}/μ_L. Panels **(a)**–**(j)** show the representative separagrams (left sub-panel of each panel) and their respective binding isotherm (right sub-panel of each panel) for the virtual ACTIS experiments with μ_{TL}/μ_L increasing from 0.5 to 0.95. The virtual experiments were simulated by COMSOL Multiphysics with the setup stated in Note A2.1: Detailed setup for COMSOL simulations. For all simulations, the mesh size setting of "Finer" was chosen. With μ_{TL}/μ_L increasing from 0.5 to 0.95, the percentage dynamic range of the separagrams decreased from 25.6% to 1.9%, and the ratio of the determined K_d value to the input K_d value varied between 0.93 and 1.18 (see **Figure 2.4** in the main text for more details).



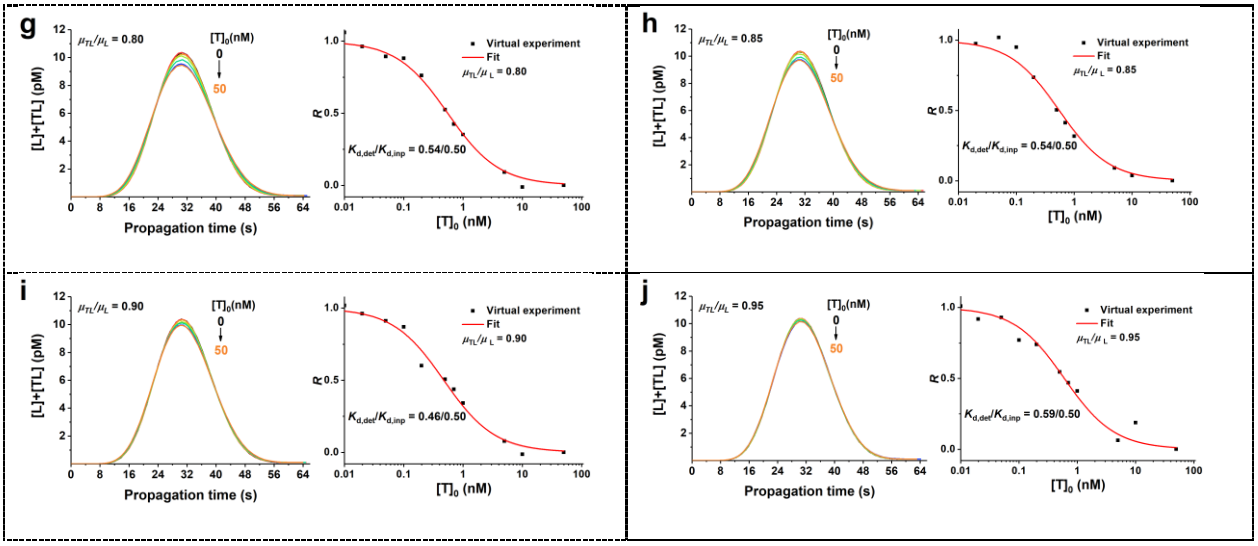

g
μ_{TL}/μ_L = 0.80
[T]$_0$(nM)
0
50
[L]+[TL] (pM)
Propagation time (s)
Virtual experiment
Fit
μ_{TL}/μ_L = 0.80
$K_{d,def}/K_{d,inp}$ = 0.54/0.50
R
[T]$_0$ (nM)
h
μ_{TL}/μ_L = 0.85
μ_{TL}/μ_L = 0.85
$K_{d,def}/K_{d,inp}$ = 0.54/0.50
i
μ_{TL}/μ_L = 0.90
μ_{TL}/μ_L = 0.90
$K_{d,def}/K_{d,inp}$ = 0.46/0.50
j
μ_{TL}/μ_L = 0.95
μ_{TL}/μ_L = 0.95
$K_{d,def}/K_{d,inp}$ = 0.59/0.50

Figure A2.4: The Virtual Experiment with Reduced Simulation Mesh Size

Separagrams **(a)** and the binding isotherm **(b)** for the virtual ACTIS experiment with a reduced mesh size**.** The virtual experiment for μ_{TL}/μ_L of 0.95 was re-conducted by changing the mesh size setting from "Finer" to "Extra Fine". All other settings used in the COMSOL virtual experiment are stated in Note A2.1: Detailed setup for COMSOL simulations. Compared to the result for μ_{TL}/μ_L = 0.95 in 2, with reduced simulation mesh size (from "Finer" to "Extra Fine"), the percentage difference between the determined and input K_d values (*i.e.*, K_d error = $|K_{d,det} - K_{d,inp}|/K_{d,inp}$) was reduced from 18% to 10%. According to **Figure A2.2** and **Figure 2.4** in the main text, μ_{TL}/μ_L = 0.95 produced the smallest dynamic range (*i.e.*, 1.9%) and the K_d error had negative correlation to the dynamic range. Hence, we can conclude that, with the decreased mesh size, the percentage difference between the determined and input K_d values for all other μ_{TL}/μ_L values (that were studied in **Figure A2.4**) would be within 10%.

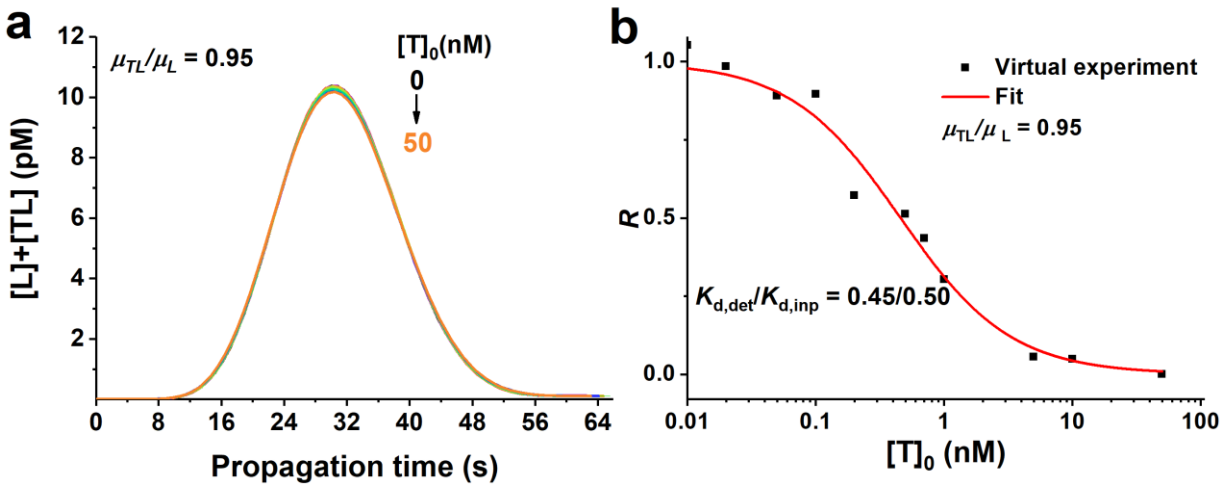


Figure A2.5: Repeatability of the ACTIS Separagrams for the Mock Titration Experiment

Repeatability of ACTIS separagrams for the 0.05 nM aptamer at different mock protein concentrations (0–10 nM). In the mock titration experiment, at step 2a of Note A2.2, the "16 µL of 10.8 µM MutS" was replaced by "16 µL buffer". The experimental procedure was the same to the one used in real ACTIS titration experiments. Experimental conditions were: internal capillary radius a = 50 µm (I.D. = 100 µm) for both injection loop and separation capillary, injection loop's length = 13 cm (injection loop's volume ≈ 1 µL), inlet-to-detector distance l = 60 cm (full separation capillary length = 79 cm), sample loading into the injection loop at a flow rate of 50 µL/min for 25 s, sample-plug transfer from the injection loop to the separation capillary at a flow rate of 5 µL/min during 24 s (plug length ≈ 13 cm, plug-end distance from the capillary inlet ≈ 13 cm), TIS flow rate Q_{TIS} ≈ 9 µL/min, mock protein concentration ranged from 0 to 10 nM. The buffer was 50 mM Tris-acetate pH 8.2 with addition of 0.05% Triton™ X-100.

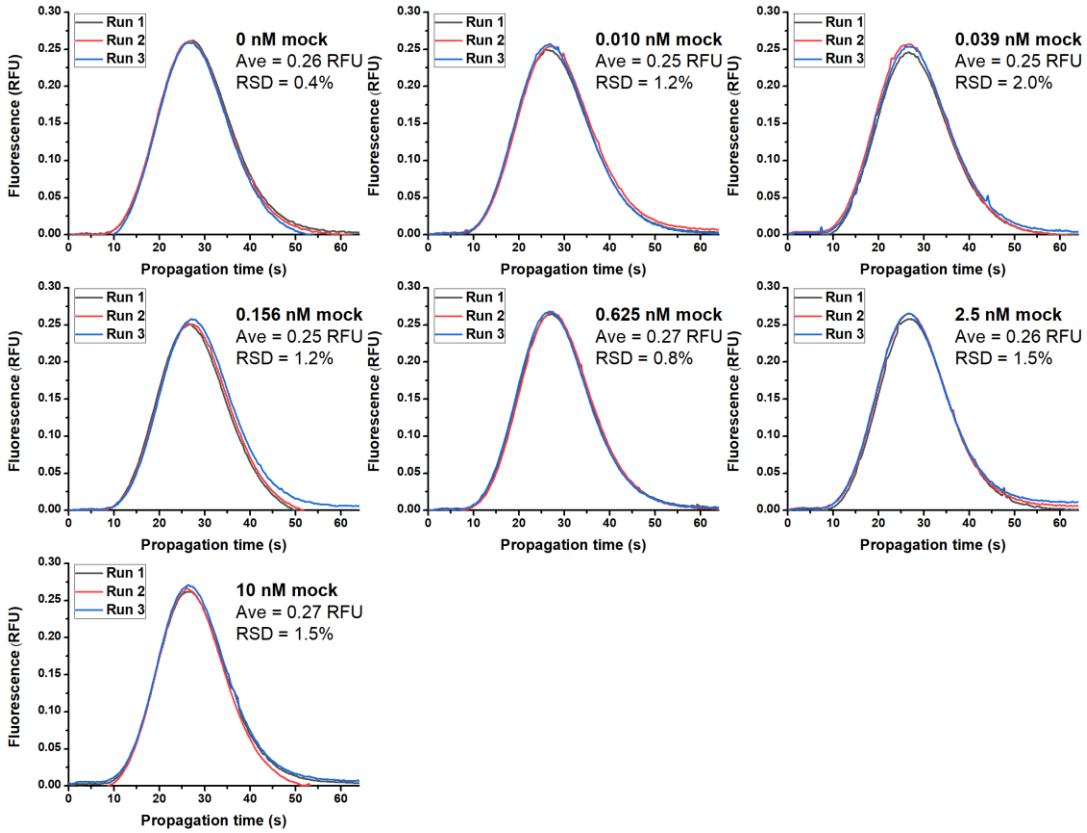


Figure A2.6: Analysis of the Results of the ACTIS-Based Mock Titration Experiment

(a) Averaged separagrams for 0.05 nM aptamer with addition of 0–10 nM mock protein concentration; **(b)** peak intensities of the averaged separagrams with the uncertainties determined in **Figure A2.3.** According to the results, the sample-preparation method does not contribute detectable signal variation on the background of variations caused by all the other sources (*i.e.*, 2%).

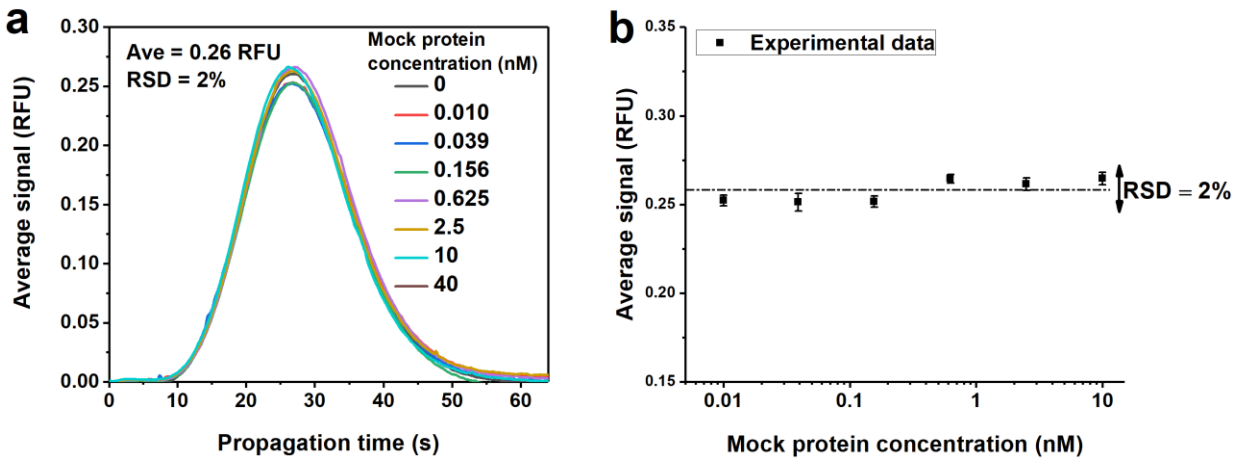


Figure A2.7: The Virtual Experiments with Varied μ_{TL}/μ_L with Added Experimental Noises

Separagrams and their respective binding isotherms for the virtual ACTIS experiments with added experimental noises**.** Panels **(a)**–**(j)** show the representative separagrams (left sub-panel of each panel) and their respective binding isotherm (right sub-panel of each panel) for the virtual ACTIS experiments with μ_{TL}/μ_L increasing from 0.5 to 0.95 with adding experimental noises. The added noises include 1% error for the flow rates and 0.1% random noise for the signal. All other settings used in the COMSOL virtual experiment are stated in Note A2.1: Detailed setup for COMSOL simulations. For all simulations, the mesh size setting of "Finer" was chosen. With μ_{TL}/μ_L increasing from 0.5 to 0.95, the ratio of the determined K_d value to the input K_d value varies between 0.68 and 2.6 (see **Figure 2.5** in the main text for more details). In the

binding isotherms, since the error of a data point is inversely proportional to its respective dynamic range, the general size of the error bars increases with the decrease of dynamic range (*i.e.*, the increase of μ_{TL}/μ_L).

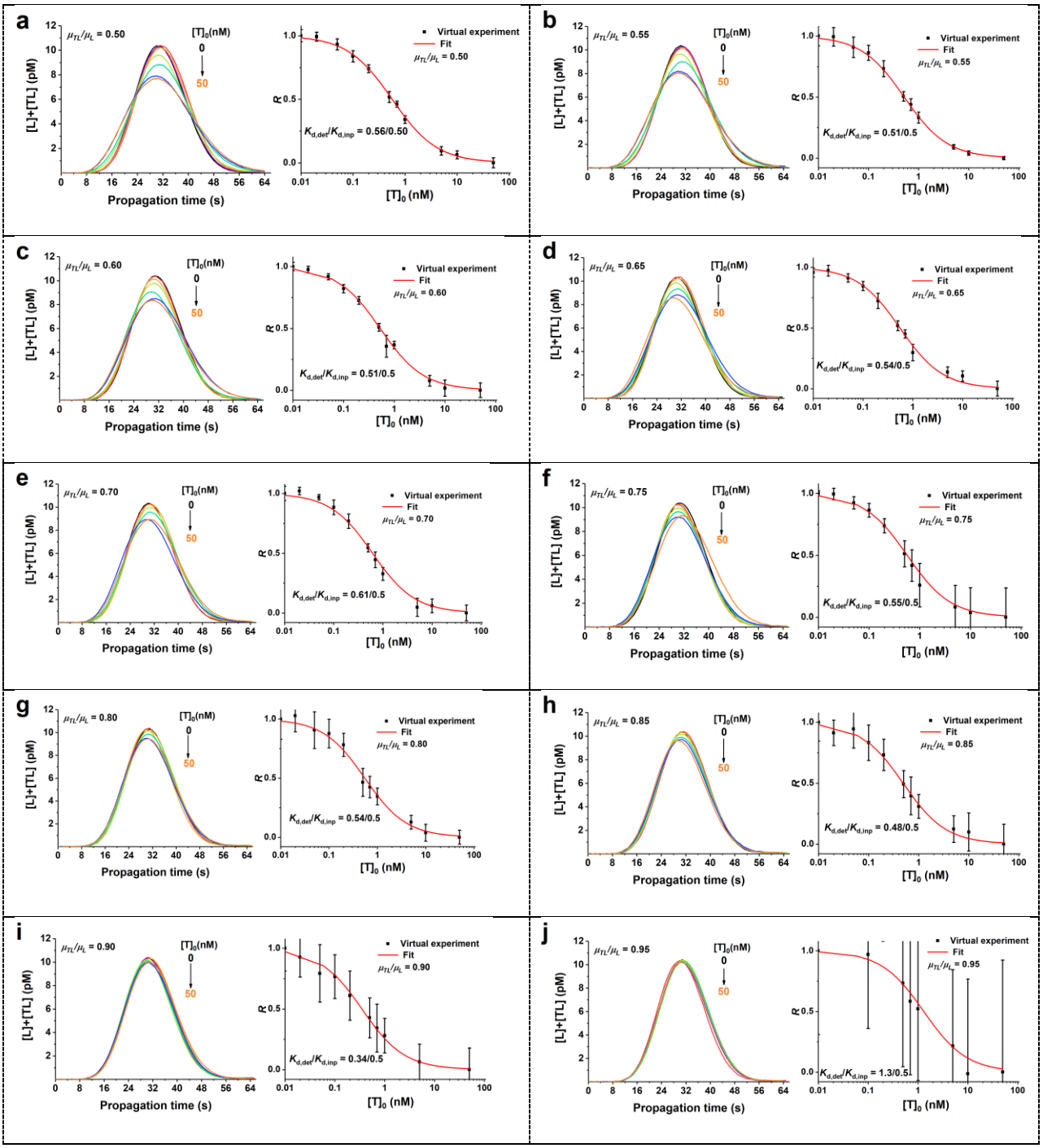


Figure A2.8: Repeatability of the ACTIS Separagrams for the Muts–Aptamer Complex, Day 1

Repeatability of day-1 ACTIS separagrams for MutS−aptamer with three repetitions at each total protein concentration. Experimental conditions were: internal capillary radius a = 50 µm (I.D. = 100 µm) for both injection loop and separation capillary, injection loop's length = 13 cm (injection loop's volume ≈ 1 µL), inlet-to-detector distance l = 60 cm (full separation capillary length = 79 cm), sample loading into the injection loop at a flow rate of 50 µL/min for 25 s, sample-plug transfer from the injection loop to the separation capillary at a flow rate of 5 µL/min during 24 s (plug length ≈ 13 cm, plug-end distance from the capillary inlet ≈ 13 cm), TIS flow rate Q_{TIS} ≈ 9 µL/min, mock protein concentration ranged from 0 to 10 nM. The buffer was 50 mM Tris-acetate pH 8.2 with addition of 0.05% Triton^{TM} X-100. The representative separagrams and binding isotherm for the day-1 ACTIS titration experiment are shown in **Figure 2.7** in the main text.

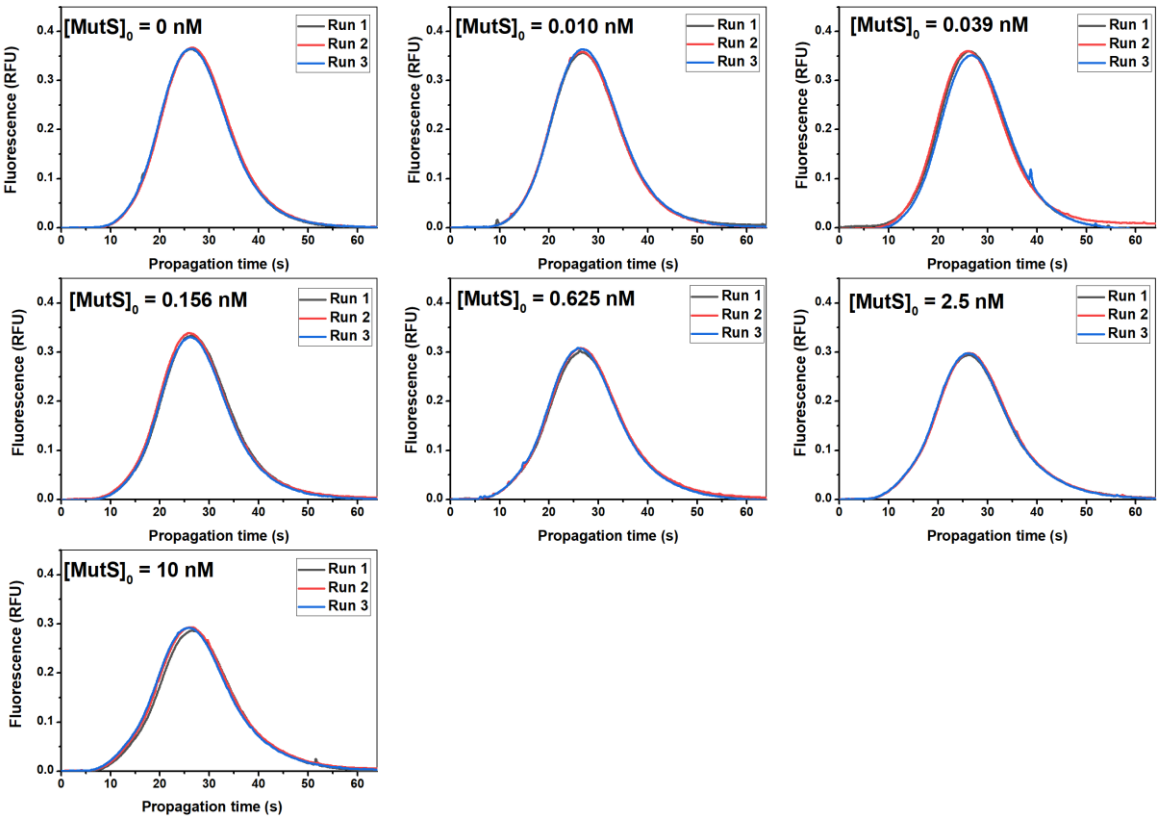


Figure A2.9: Repeatability of the ACTIS Separagrams for the Muts–Aptamer Complex, Day 2

Repeatability of day-2 ACTIS separagrams for MutS−aptamer with three repetitions at each total protein concentration. Experimental conditions were: internal capillary radius a = 50 µm (I.D. = 100 µm) for both injection loop and separation capillary, injection loop's length = 13 cm (injection loop's volume ≈ 1 µL), inlet-to-detector distance l = 60 cm (full separation capillary length = 79 cm), sample loading into the injection loop at a flow rate of 50 µL/min for 25 s, sample-plug transfer from the injection loop to the separation capillary at a flow rate of 5 µL/min during 24 s (plug length ≈ 13 cm, plug-end distance from the capillary inlet ≈ 13 cm), TIS flow rate Q_{TIS} ≈ 9 µL/min, mock protein concentration ranged from 0 to 10 nM. The buffer was 50 mM Tris-acetate pH 8.2 with addition of 0.05% Triton™ X-100

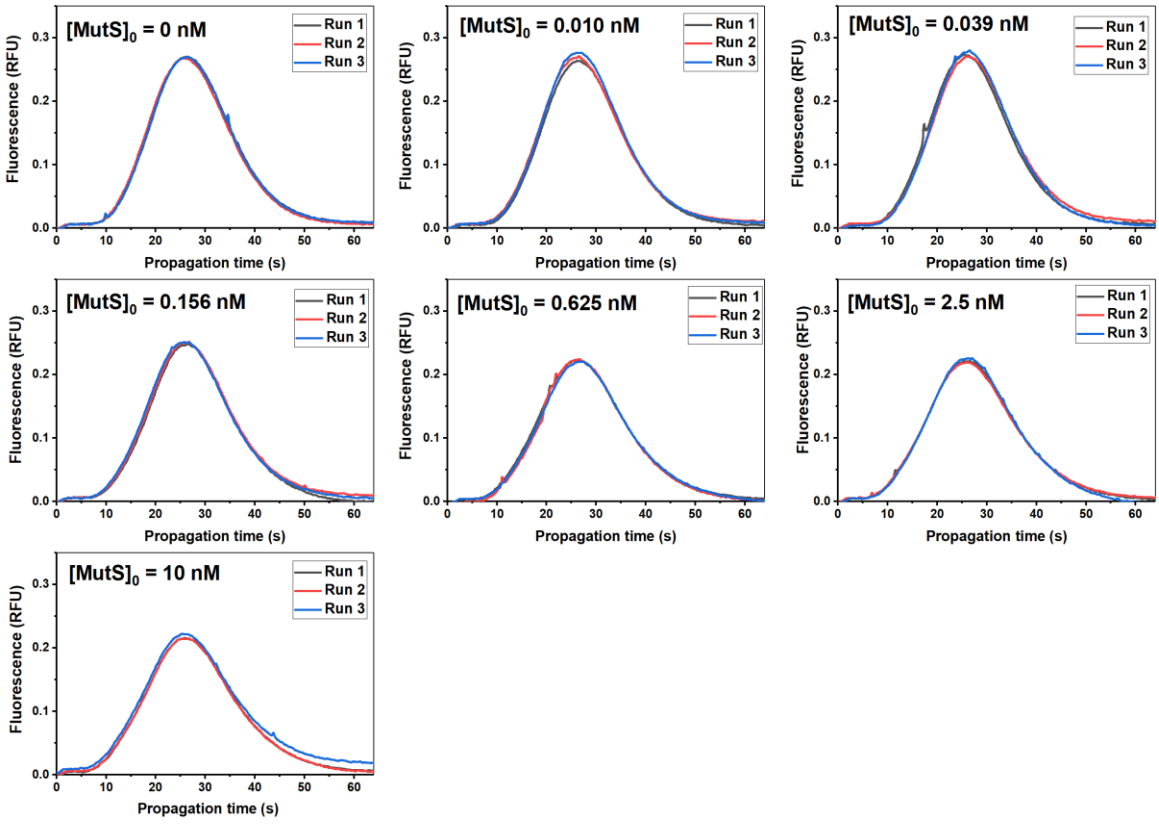


Figure A2.10: ACTIS-Based K_d Determination for the MutS–Aptamer Complex, Day 3

(a) Representative separagrams at different protein concentrations; **(b)** binding isotherm and curve fitting (using nonlinear regression) to determine K_d. The aptamer (ligand) concentration was kept at 0.05 nM, and the MutS (target) concentration was varied from 0 to 10 nM.

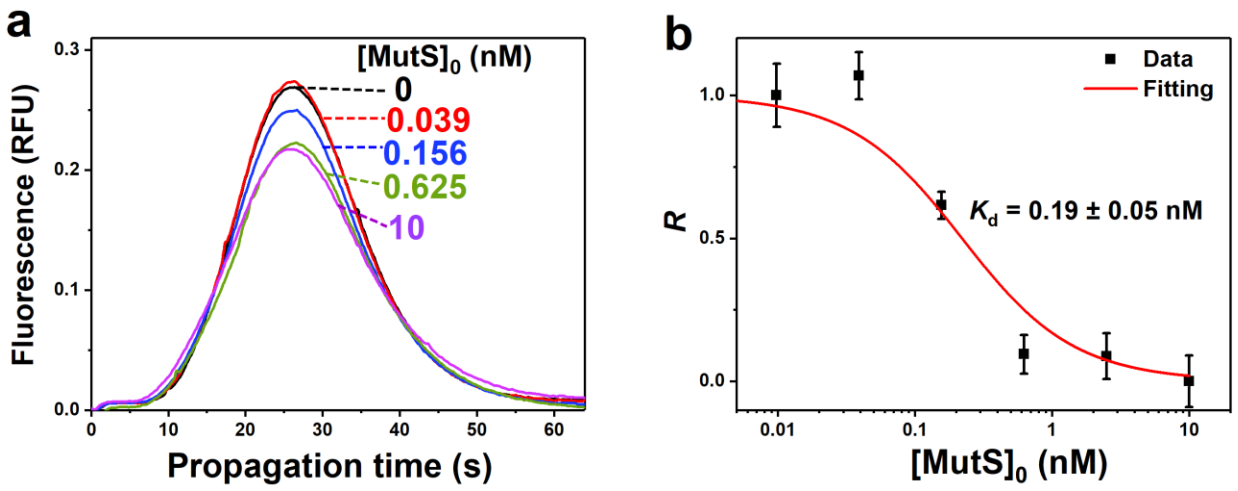
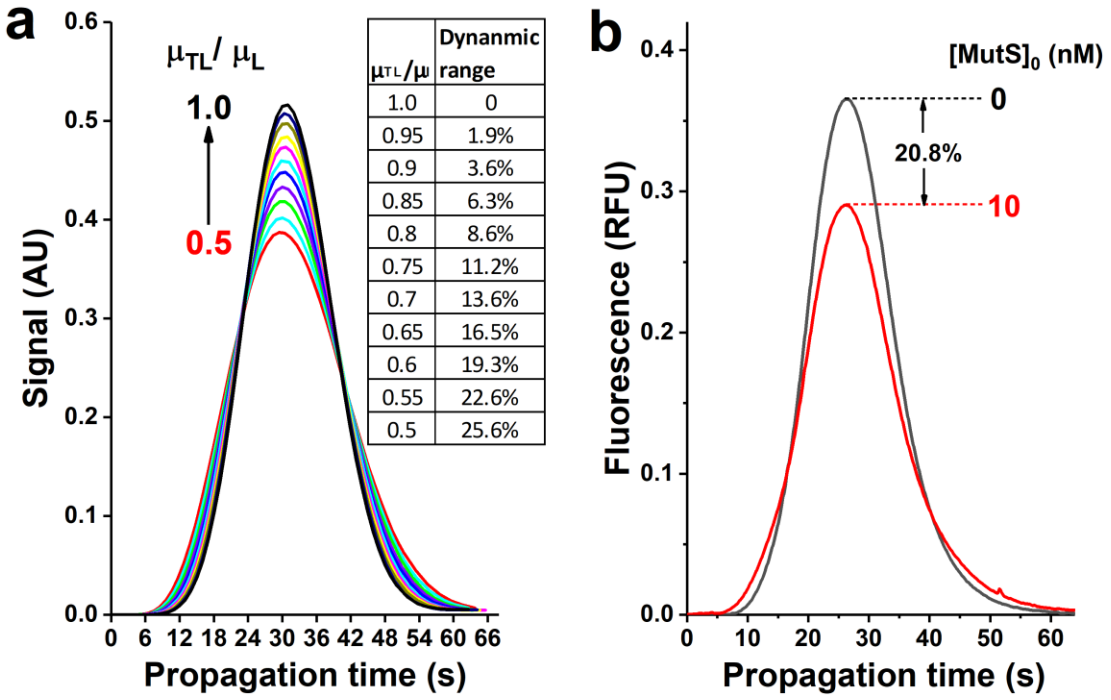


Figure A2.11: Estimation of μ_{TL}/μ_L for the MutS–Aptamer

Estimating μ_{TL}/μ_L for the MutS–aptamer complex by comparing the experimental results to the simulated data. To estimate μ_{TL}/μ_L for MutS–aptamer pair, separagrams of **Figure 2.4a** in the main text **(a)** and **Figure 2.7a** in the main text (**b**, for $[MutS]_0$ = 0 and 10 nM) were re-formatted and compared. By comparing the dynamic range of the experimental separagrams (*i.e.*, 20.8%) to the dynamic ranges obtained from the virtual experiments, the μ_{TL}/μ_L value for the MutS–aptamer complex was estimated to be 0.6.



μ_{TL}/μ	Dynanmic range
1.0	0
0.95	1.9%
0.9	3.6%
0.85	6.3%
0.8	8.6%
0.75	11.2%
0.7	13.6%
0.65	16.5%
0.6	19.3%
0.55	22.6%
0.5	25.6%

Figure A2.12: Repeatability of the ACTIS Separagrams for the MutS–Aptamer Complex in the ACTIS/NECEEM Parallel Experiment

Repeatability of ACTIS separagrams experiment for MutS−aptamer with three repetitions at each total protein concentration (in the ACTIS/NECEEM parallel titration). Experimental conditions were: internal capillary radius a = 50 µm (I.D. = 100 µm) for both injection loop and separation capillary, injection loop's length = 13 cm (injection loop's volume ≈ 1 µL), inlet-to-detector distance l = 60 cm (full separation capillary length = 79 cm), sample loading into the injection loop at a flow rate of 50 µL/min for 25 s, sample-plug transfer from the injection loop to the separation capillary at a flow rate of 5 µL/min during 24 s (plug length ≈ 13 cm, plug-end distance from the capillary inlet ≈ 13 cm), TIS flow rate Q_{TIS} ≈ 9 µL/min, mock protein

concentration ranged from 0 to 10 nM. The buffer was 50 mM Tris-acetate pH 8.2 with addition of 0.05% Triton^{TM} X-100.

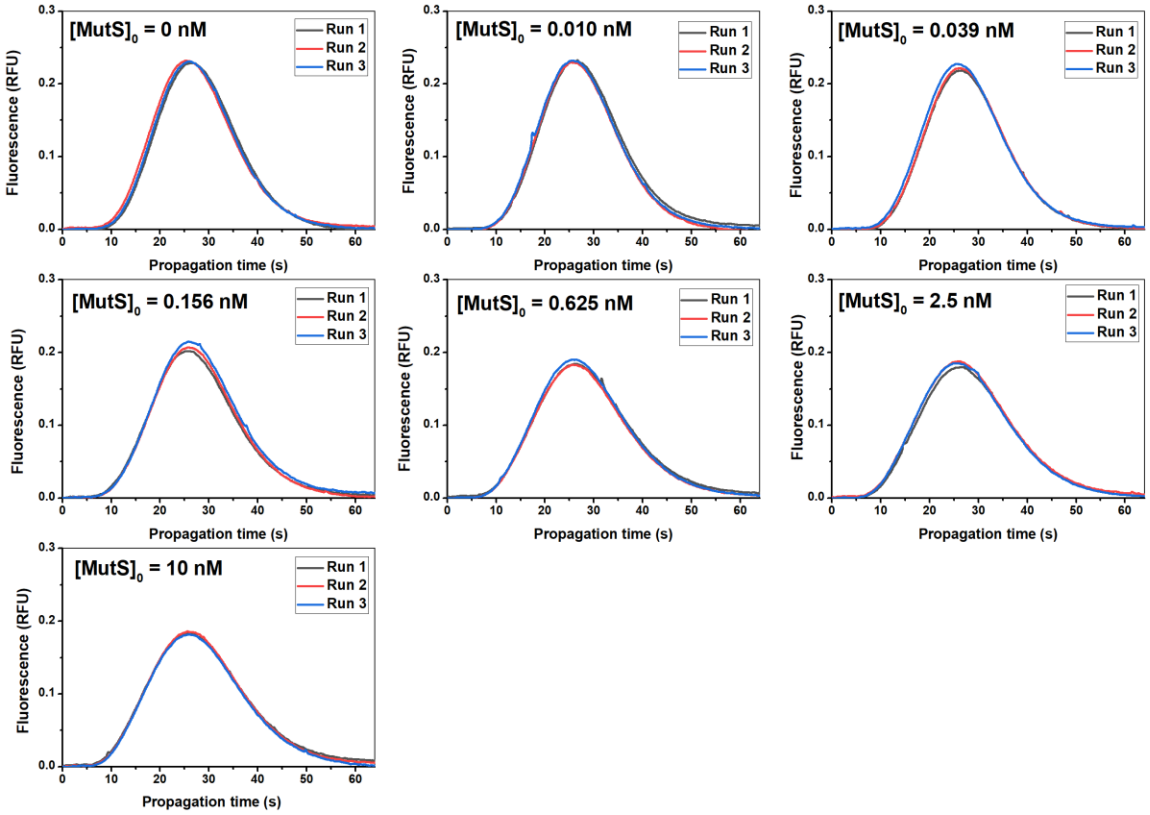


Figure A2.13: Representative ACTIS Separagrams Obtained in the ACTIS/NECEEM Parallel Experiment

The aptamer (ligand) concentration was kept at 0.05 nM, and the MutS (target) concentration was varied from 0 to 10 nM. The binding isotherm for this ACTIS titration experiment is shown in **Figure 2.8a** in the main text.

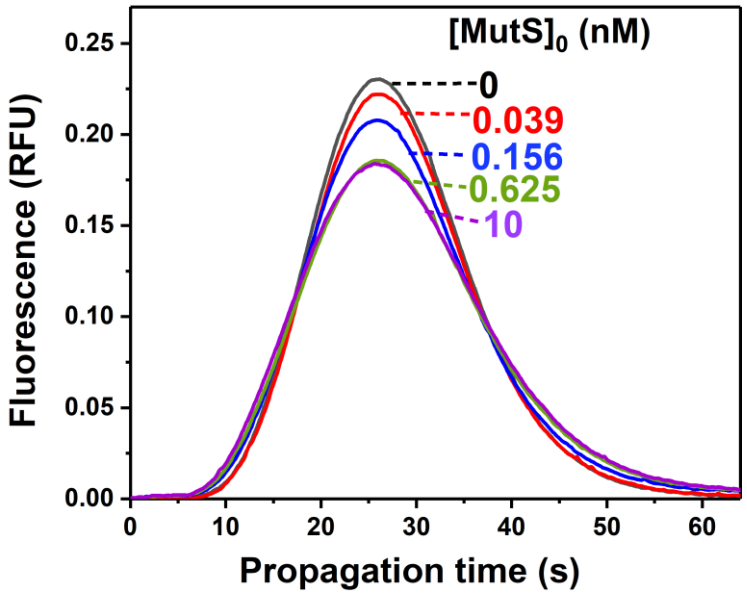


Figure A2.14: Electrophoretograms of the NECEEM Experiments in the ACTIS/NECEEM Parallel Experiment

Experimental conditions were: internal capillary radius a = 50 µm (I.D. = 100 µm) for the separation capillary, inlet-to-detector distance was 39.8 cm (full separation capillary length = 50 cm), sample was injected into the capillary at a pressure of 0.5 psi for 10 s (sample plug length ≈ 1.9 cm). The sample plug was propagated at a pressure of 0.3 psi for 50 s (propagation distance ≈ 5.7 cm). The total separation distance was 33.2 cm (from the center of the sample plug to the detector). The voltage separation was carried out at an electric field of 500 V/cm (25 kV over 50 cm) with a reversed polarity (anode at the outlet). The buffer was 50 mM Tris-acetate pH 8.2 with addition of 0.05% Triton^{TM} X-100. The equilibrium mixtures (samples) contained 0.05 nM aptamer with addition of 0−2.5 nM MutS. Note: In the NECEEM experiment, samples identical to those of the ACTIS experiment (*i.e.*, 0.05 nM aptamer + 0−10 nM MutS)

were studied. However, the NECEEM measurement for the sample of "0.05 nM aptamer + 10 nM MutS" was not completed due to an instrumental technical issue.

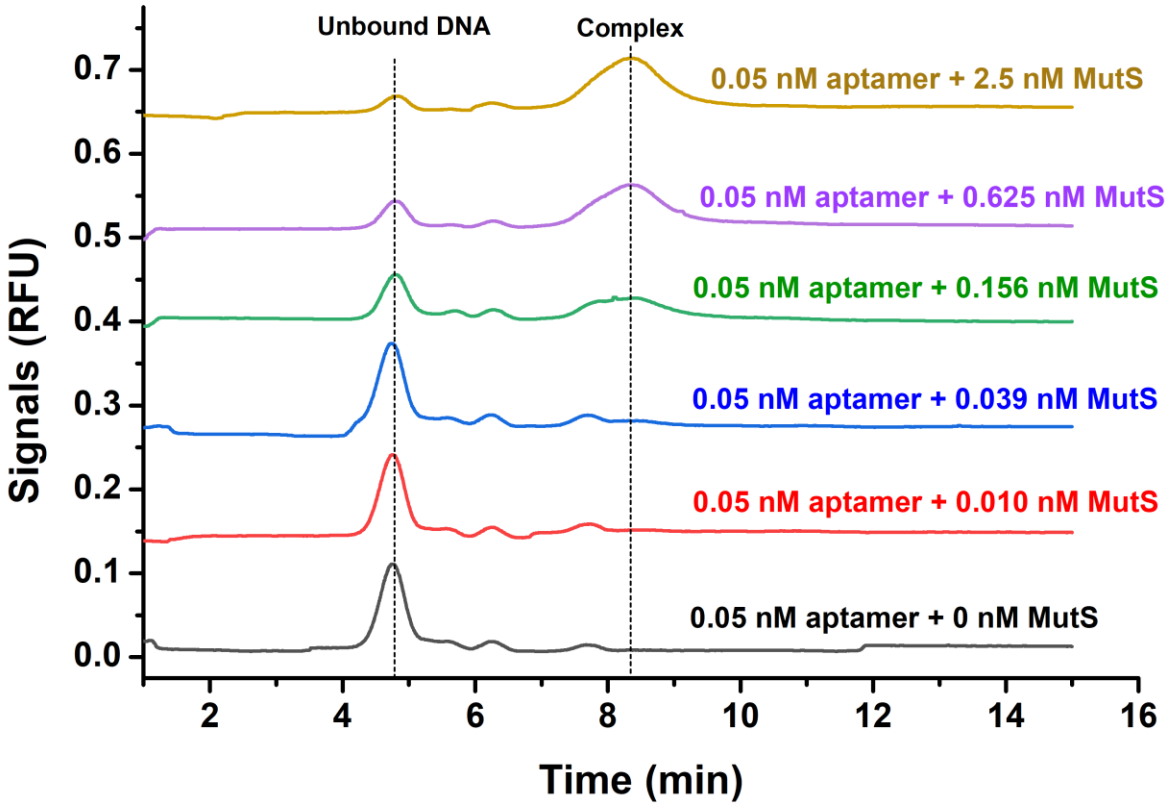


Figure A2.15: Two-Day ACTIS-Based K_d Determination for the SSB–22-nt ssDNA Complex

(a) Representative separagrams at different protein concentrations (Day 1); **(b)** binding isotherm and curve fitting (using nonlinear regression) to determine K_d (Day 1). **(c)** Representative separagrams at different protein concentrations (Day 2); **(d)** binding isotherm and curve fitting (using nonlinear regression) to determine K_d (Day 2). The ssDNA concentration was kept at 10 nM, and the SSB (target) concentration was varied from 0 to 512 nM.

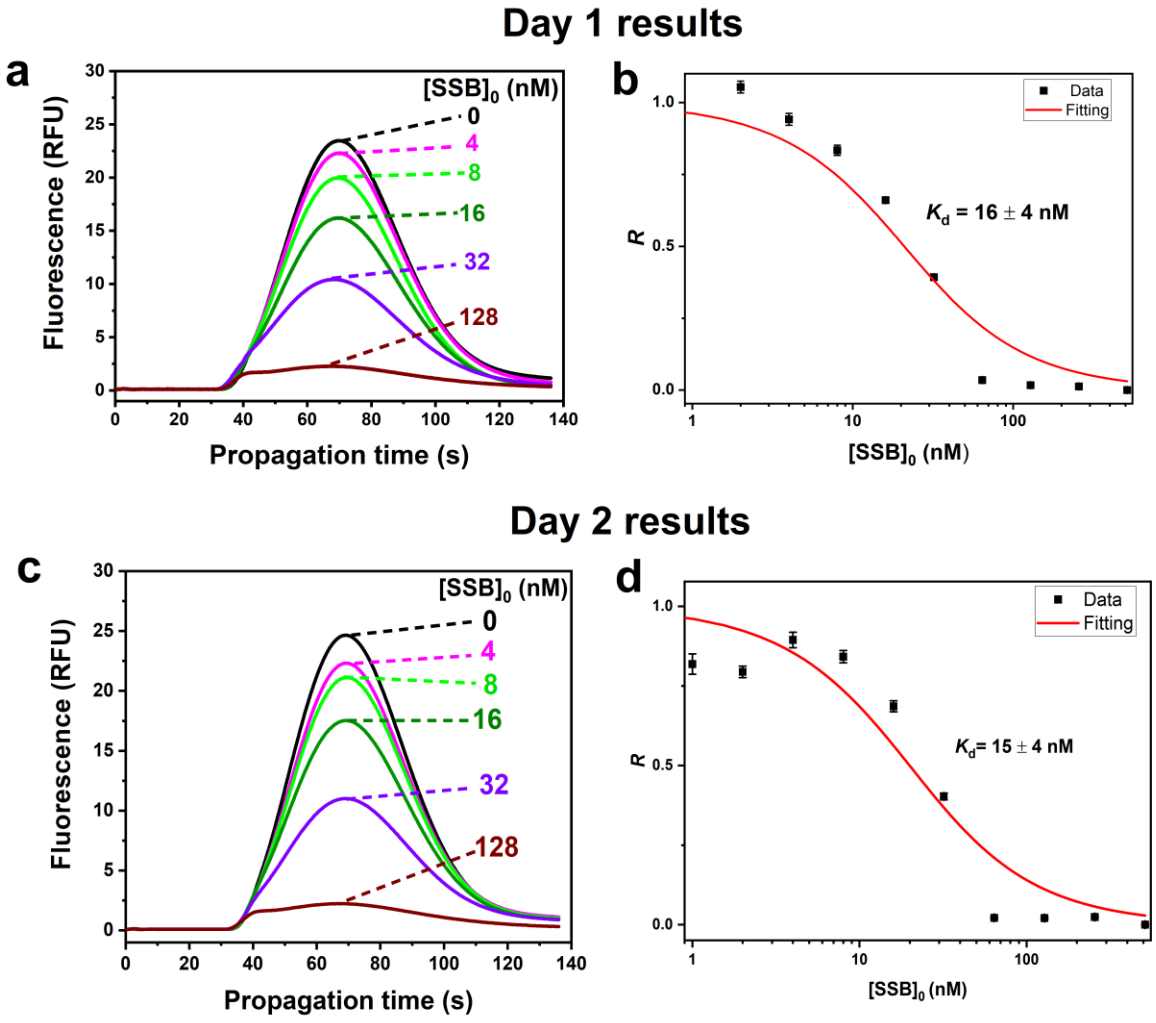


Table A2.1: Calculation of R Values for the NECEEM Titration Experiment

Based on the electrophoretograms in **Figure A2.13**, the areas of the peaks for unbound DNA and complex were measured with Origin Pro. software. Since $[MutS]_0$ = 2.5 nM was much greater than the expected K_d value range of 0.1−0.2 nM, we presumed that 2.5 nM MutS saturated the binding to 0.05 nM aptamer, and the remaining unbound DNA was the impurity (*i.e.*, non-binders) in the aptamer. Hence, from each unbound DNA area, we subtracted the unbound DNA

area of "0.05 nM aptamer + 2.5 nM MutS" to obtain the actual unbound aptamer areas. Then, the areas were corrected by dividing them by their migration times.[2] Based on the corrected areas, the actual fractions of unbound aptamer, *R* values, at different $[MutS]_0$ were calculated. The binding isotherm for this NECEEM titration experiment is shown in **Figure 2.8b** in the main text.

$[MutS]_0$ (nM)	**Area of unbound DNA**	**Actual area of unbound aptamer**	**Migration time of unbound aptamer (min)**	**Corrected area of unbound aptamer**	**Area of complex**	**Migration time of complex (min)**	**Corrected area of complex**	**Corrected total area (RSD ≈ 5%)**	***R***
0.000	0.0516	0.0424	4.75	0.00893	0.00000	8.36	0.000000	0.00893	1.000
0.010	0.0501	0.0409	4.75	0.00861	0.00299	8.36	0.000358	0.00897	0.960
0.039	0.0511	0.0419	4.73	0.00886	0.00808	8.28	0.000976	0.00983	0.901
0.156	0.0274	0.0182	4.79	0.00380	0.03753	8.32	0.004511	0.00831	0.457
0.065	0.0162	0.0070	4.80	0.00146	0.06368	8.33	0.007645	0.00910	0.160
2.500	0.0092	0.0000	4.83	0.00000	0.07433	8.36	0.008891	0.00889	0.000

Appendix B. Supplementary Information for Chapter 3

Appendix B1. Supporting Information for "Fundamental Determinants of the Accuracy of Equilibrium Constants for Affinity Complexes"

Note B1.1: Derivation of the Theoretical Dependence of Fraction of Unbound Ligand (R) on Total Target Concentration (T_0)

The definition of the equilibrium dissociation constant K_d and mass balance for the target and ligand show:

$$K_d = \frac{TL}{C} \tag{B1}$$

$$T_0 = T + C \Rightarrow T = T_0 - C \tag{B2}$$

$$L_0 = L + C \Rightarrow C = L_0 - L \tag{B3}$$

where T, L, and C are equilibrium concentrations of the target, ligand and complex, respectively; T_0 and L_0 are the total concentrations of target and ligand, respectively. The definition of fraction of unbound ligand R ($0 \le R \le 1$) is:

$$R = \frac{L}{L_0} \Rightarrow L = RL_0 \tag{B4}$$

By inserting **Eq (B4)** into **Eq (B3)**, we have:

$$C = L_0 - RL_0 = L_0(1 - R) \tag{B5}$$

By replacing C in **Eq (B2)** with **Eq (B5)**, we obtain:

$$T = T_0 - L_0(1 - R) \tag{B6}$$

With inserting **Eqs (B4)**, **(B5)** and **(B6)** into **Eq (B1)**, we obtain:

$$
\begin{aligned}
K_{\mathrm{d}} &= \frac{\left[T_0 - L_0(1-R)\right]RL_0}{(1-R)L_0} \\
&= \frac{\left[T_0 - L_0(1-R)\right]R}{(1-R)} \\
&= \frac{T_0 - L_0(1-R)}{(1/R-1)}
\end{aligned}
\tag{B7}
$$

By rearranging **Eq (B7)**, we obtain a quadratic equation for R:

$$
L_0R^2 + \left(K_{\mathrm{d}} + T_0 - L_0\right)R - K_{\mathrm{d}} = 0 \tag{B8}
$$

With the quadratic formula, the expression of R is solved as

$$
R = \frac{-(K_{\mathrm{d}} + T_0 - L_0) \pm \sqrt{(K_{\mathrm{d}} + T_0 - L_0)^2 + 4L_0K_{\mathrm{d}}}}{2L_0} \tag{B9}
$$

with two solutions. Now, let's discuss which solution should be picked to satisfy $0 \le R \le 1$.

When $(K_{\mathrm{d}} + T_0 - L_0) \ge 0$, the only solution that satisfies $R \ge 0$ is:

$$
R = \frac{-(K_{\mathrm{d}} + T_0 - L_0) + \sqrt{(K_{\mathrm{d}} + T_0 - L_0)^2 + 4L_0K_{\mathrm{d}}}}{2L_0} \tag{B10}
$$

since the other solution leads to negative values for R.

When $(K_{\mathrm{d}} + T_0 - L_0) < 0$,

$$
-(K_{\mathrm{d}} + T_0 - L_0) = \left|K_{\mathrm{d}} + T_0 - L_0\right| \tag{B11}
$$

$$
\sqrt{(K_{\mathrm{d}} + T_0 - L_0)^2 + 4L_0K_{\mathrm{d}}} \ge \sqrt{(K_{\mathrm{d}} + T_0 - L_0)^2} = \left|K_{\mathrm{d}} + T_0 - L_0\right|, \text{ since } L_0 \ge 0 \text{ and } K_{\mathrm{d}} > 0 \tag{B12}
$$

Therefore, the only solution that satisfies $R \ge 0$ is also **Eq (B10)**:

$$
R = \frac{-(K_{\mathrm{d}} + T_0 - L_0) + \sqrt{(K_{\mathrm{d}} + T_0 - L_0)^2 + 4L_0K_{\mathrm{d}}}}{2L_0}
$$

since the other solution leads to $R < 0$.

To sum up, the theoretical dependence of R on T_0 is **Eq (B10)** or rearranged to:

$$R = -\frac{K_\mathrm{d} + T_0 - L_0}{2L_0} + \sqrt{\left(\frac{K_\mathrm{d} + T_0 - L_0}{2L_0}\right)^2 + \frac{K_\mathrm{d}}{L_0}} \tag{B13}$$

which is **Eq (2.2)** in the main text.

Note B1.2: Derivation of the Dependence of Relative Systematic Error of K_d ($\Delta K_d/K_d$) on L_0/K_d Ratio

Eq (B7) shows:

$$K_\mathrm{d} = \frac{T_0 - L_0(1-R)}{(1/R-1)}$$

The systematic error of K_d can be calculated with the error propagation rule:

$$|\Delta K_\mathrm{d}| = \left|\left(\frac{\partial K_\mathrm{d}}{\partial T_0}\right)\Delta T_0\right| + \left|\left(\frac{\partial K_\mathrm{d}}{\partial L_0}\right)\Delta L_0\right| + \left|\left(\frac{\partial K_\mathrm{d}}{\partial R}\right)\Delta R\right| \tag{B14}$$

which gives,

$$\begin{aligned} |\Delta K_\mathrm{d}| &= \left|\frac{R}{1-R}\Delta T_0\right| + |-R\Delta L_0| + \left|\left[-L_0 + \frac{T_0}{(1-R)^2}\right]\Delta R\right| \\ &= \left|\frac{R}{1-R}\right|\left|\Delta T_0\right| + |R||\Delta L_0| + \left|-L_0 + \frac{T_0}{(1-R)^2}\right||\Delta R| \end{aligned} \tag{B15}$$

Apparently, both $R/(1-R)$ and R are greater than zero since $R \in (0, 1)$ when we use **Eq (B7)** to determine K_d. Now, to expand and rearrange **Eq (B15)**, let's determine the sign of "$-L_0 + T_0/(1-R)^2$":

From **Eq (B7)**, we obtain,

$$T_0 = (1/R - 1)K_\mathrm{d} + (1-R)L_0 \tag{B16}$$

Then,

$$
\begin{aligned}
-L_0 + \frac{T_0}{(1-R)^2} &= \left[\frac{1/R - 1}{(1-R)^2}\right] K_d + \frac{(1-R)}{(1-R)^2} L_0 - L_0 \\
&= \left[\frac{1-R}{R(1-R)^2}\right] K_d + \frac{1}{1-R} L_0 - L_0 \\
&= \frac{K_d}{R(1-R)} + \frac{R}{1-R} L_0
\end{aligned}
\tag{B17}
$$

Eq (B17) clearly shows that the term "$-L_0 + T_0/(1-R)^2$" is greater than zero since both K_d and L_0 are greater than zero and $R \in (0, 1)$. Hence, we rewrite **Eq (B15)** as:

$$
|\Delta K_d| = \frac{R}{1-R}|\Delta T_0| + R|\Delta L_0| + \left[-L_0 + \frac{T_0}{(1-R)^2}\right]|\Delta R| \tag{B18}
$$

To simplify the analysis, we only focus on the scenario of determined R equal 0.5, which is close to the least erroneous case. Then, the actual R is:

$$
R = 0.5 - \Delta R, \quad -0.5 < \Delta R < 0.5 \tag{B19}
$$

Insert **Eq (B19)** into **Eq (B16)**, we can express T_0 with L_0 and ΔR,

$$
T_0 = \frac{(0.5 + \Delta R) K_d}{0.5 - \Delta R} + (0.5 + \Delta R) L_0 \tag{B20}
$$

Now, divide K_d for both sides of **Eq (B18)** to analyze the relative systematic error of K_d:

$$
\left|\frac{\Delta K_d}{K_d}\right| = \frac{R}{1-R}\left|\frac{\Delta T_0}{K_d}\right| + R\left|\frac{\Delta L_0}{K_d}\right| + \left[-\frac{L_0}{K_d} + \frac{1}{(1-R)^2}\frac{T_0}{K_d}\right]|\Delta R| \tag{B21}
$$

Or rewrite as:

$$
\left|\frac{\Delta K_d}{K_d}\right| = \frac{R}{1-R}\left|\frac{\Delta T_0}{T_0}\right|\left(\frac{T_0}{K_d}\right) + R\left|\frac{\Delta L_0}{L_0}\right|\left(\frac{L_0}{K_d}\right) + \left[-\frac{L_0}{K_d} + \frac{1}{(1-R)^2}\frac{T_0}{K_d}\right]|\Delta R| \tag{B22}
$$

By inserting **Eq (B19)** and **Eq (B20)** into **Eq (B22)**, we have:

$$\left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = \frac{0.5-\Delta R}{0.5+\Delta R}\left|\frac{\Delta T_0}{T_0}\right|\left[\frac{0.5+\Delta R}{0.5-\Delta R}+\left(\frac{L_0}{K_{\mathrm{d}}}\right)(0.5+\Delta R)\right]+(0.5-\Delta R)\left|\frac{\Delta L_0}{L_0}\right|\left(\frac{L_0}{K_{\mathrm{d}}}\right)$$
$$+\left[-\frac{L_0}{K_{\mathrm{d}}}+\frac{1}{(0.5+\Delta R)^2}\left(\frac{0.5+\Delta R}{0.5-\Delta R}+\left(\frac{L_0}{K_{\mathrm{d}}}\right)(0.5+\Delta R)\right)\right]|\Delta R| \tag{B23}$$

If we set the absolute value of relative error of R, $|\Delta R/R| = r$ at $R = 0.5 - \Delta R$ (**Eq (B19)**), then:

When $0 \le \Delta R < 0.5$,

$$\begin{aligned}\Delta R &= Rr = (0.5-\Delta R)r\\ &\Rightarrow \Delta R = 0.5r - r\Delta R\\ &\Rightarrow (1+r)\Delta R = 0.5r\\ &\Rightarrow \Delta R = \frac{0.5r}{1+r}\end{aligned} \tag{B24}$$

When $-0.5 < \Delta R < 0$,

$$\begin{aligned}\Delta R &= -Rr = (\Delta R-0.5)r\\ &\Rightarrow \Delta R = r\Delta R - 0.5r\\ &\Rightarrow (r-1)\Delta R = 0.5r\\ &\Rightarrow \Delta R = \frac{0.5r}{r-1}\end{aligned} \tag{B25}$$

If we also set the absolute values of relative errors of T_0 and L_0 as $|\Delta T_0/T_0| = t$ and $|\Delta L_0/L_0| = l$, respectively, **Eq (B23)** is rearranged to (we omit some intermediate steps here):

When $0 \le \Delta R < 0.5$,

$$\left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = \frac{0.5t}{0.5+r}\left[\frac{0.5+r}{0.5}+\frac{0.5+r}{1+r}\frac{L_0}{K_{\mathrm{d}}}\right]+\frac{0.5l}{1+r}\frac{L_0}{K_{\mathrm{d}}}$$
$$+\left[-\frac{L_0}{K_{\mathrm{d}}}+\left(\frac{1+r}{0.5+r}\right)^2\left(\frac{0.5+r}{0.5}+\frac{0.5+r}{1+r}\frac{L_0}{K_{\mathrm{d}}}\right)\right]\frac{0.5r}{1+r} \tag{B26}$$

$$\Rightarrow \left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = \left[t+\frac{r(1+r)}{0.5+r}\right]+\left[\frac{0.5(t+l-r)}{1+r}+\frac{0.5r}{0.5+r}\right]\frac{L_0}{K_{\mathrm{d}}} \tag{B27}$$

When $-0.5 < \Delta R < 0$,

$$\left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = \frac{0.5t}{0.5-r}\left[\frac{0.5-r}{0.5}+\frac{r-0.5}{r-1}\frac{L_0}{K_{\mathrm{d}}}\right]+\frac{0.5l}{1-r}\frac{L_0}{K_{\mathrm{d}}} + \left[-\frac{L_0}{K_{\mathrm{d}}}+\left(\frac{r-1}{r-0.5}\right)^2\left(\frac{0.5-r}{0.5}+\frac{r-0.5}{r-1}\frac{L_0}{K_{\mathrm{d}}}\right)\right]\frac{0.5r}{1-r} \tag{B28}$$

$$\Rightarrow \left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = \left[t+\frac{r(1-r)}{0.5-r}\right]+\left[\frac{0.5(t+l-r)}{1-r}+\frac{0.5r}{0.5-r}\right]\frac{L_0}{K_{\mathrm{d}}} \tag{B29}$$

Or we can write **Eq (B27)** and **Eq (B29)** as:

$$\begin{aligned}&\left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = a+b\frac{L_0}{K_{\mathrm{d}}},\\ &a = t+\frac{r(1+r)}{0.5+r},\ b=\frac{0.5(t+l-r)}{1+r}+\frac{0.5r}{0.5+r}\ \text{ when } 0\le\Delta R<0.5\\ &a = t+\frac{r(1-r)}{0.5-r},\ b=\frac{0.5(t+l-r)}{1-r}+\frac{0.5r}{0.5-r}\ \text{ when } -0.5<\Delta R<0\end{aligned} \tag{B30}$$

which is **Eq (3.3)** in the main text.

Note B1.3: Derivation of the Theoretical Dependence of Determined K_d ($K_{d,Det}$) on L_0

The definition of ΔK_{d} is (with $K_{\mathrm{d,det}}$ representing the determined K_{d} value),

$$\Delta K_{\mathrm{d}} = K_{\mathrm{d,det}} - K_{\mathrm{d}} \tag{B31}$$

By inserting **Eq (B31)** into **Eq (B30)**, we get:

$$\left|\frac{K_{\mathrm{d,det}}-K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = a+b\frac{L_0}{K_{\mathrm{d}}} \tag{B32}$$

By multiplying K_{d} on both sides of **Eq (B32)**, we have:

$$\left|K_{\mathrm{d,det}}-K_{\mathrm{d}}\right| = aK_{\mathrm{d}}+bL_0 \tag{B33}$$

When $\Delta K_{\mathrm{d}} \ge 0$,

$$K_{\mathrm{d,det}} - K_{\mathrm{d}} = aK_{\mathrm{d}} + bL_0$$
$$\Rightarrow K_{\mathrm{d,det}} = (a+1)K_{\mathrm{d}} + bL_0 \quad \text{(B34)}$$

When $\Delta K_{\mathrm{d}} < 0$,

$$K_{\mathrm{d}} - K_{\mathrm{d,det}} = aK_{\mathrm{d}} + bL_0$$
$$\Rightarrow K_{\mathrm{d,det}} = (1-a)K_{\mathrm{d}} - bL_0 \quad \text{(B35)}$$

Or we can combine **Eq (B34)** and **Eq (B35)** to be:

$$K_{\mathrm{d,det}} = c + dL_0,\ c = (a+1)K_{\mathrm{d}},\ d = b \text{ when } \Delta K_{\mathrm{d}} \geq 0$$
$$c = (1-a)K_{\mathrm{d}},\ d = -b \text{ when } \Delta K_{\mathrm{d}} < 0 \quad \text{(B36)}$$

Eq (B36) (**Eq (3.6)** in the main text) suggests that the dependence of $K_{\mathrm{d,det}}$ on L_0 is linear in the whole range of L_0.

Note B1.4: Derivation of the Dependence of Relative Systematic Error of K_d ($|\Delta K_d/K_d|$) on L_0/K_d Ratio for Uncorrelated (Orthogonal) Error Sources

When the error sources of T_0, L_0 and R are uncorrelated, the systematic error of K_{d} can be calculated with the error propagation rule:

$$|\Delta K_{\mathrm{d}}| = \sqrt{\left(\frac{\partial K_{\mathrm{d}}}{\partial T_0}\right)^2 \Delta {T_0}^2 + \left(\frac{\partial K_{\mathrm{d}}}{\partial L_0}\right)^2 \Delta {L_0}^2 + \left(\frac{\partial K_{\mathrm{d}}}{\partial R}\right)^2 \Delta R^2} \quad \text{(B37)}$$

By applying **Eq (B37)** to **Eq (B7)** and square both sides, we have,

$$\Delta {K_{\mathrm{d}}}^2 = \left(\frac{R}{1-R}\right)^2 \Delta {T_0}^2 + R^2 \Delta {L_0}^2 + \left[-L_0 + \frac{T_0}{(1-R)^2}\right]^2 \Delta R^2 \quad \text{(B38)}$$

By replacing T_0 with **Eq (B16)**, we obtain:

$$\Delta {K_{\mathrm{d}}}^2 = \left(\frac{R}{1-R}\right)^2 \Delta {T_0}^2 + R^2 \Delta {L_0}^2 + \left[\left(\frac{1}{1-R}\right)\left(\frac{K_{\mathrm{d}}}{R} + RL_0\right)\right]^2 \Delta R^2 \quad \text{(B39)}$$

Now, divide K_d^2 for both sides of **Eq (B39)** to analyze the relative systematic error of K_d:

$$\begin{aligned}\left(\frac{\Delta K_d}{K_d}\right)^2 &= \left(\frac{R}{1-R}\right)^2\left(\frac{\Delta T_0}{K_d}\right)^2 + R^2\left(\frac{\Delta L_0}{K_d}\right)^2 + \left[\left(\frac{1}{1-R}\right)\left(\frac{1}{R}+R\frac{L_0}{K_d}\right)\right]^2\Delta R^2 \\ &= \left(\frac{R}{1-R}\right)^2\left(\frac{\Delta T_0}{T_0}\right)^2\left(\frac{T_0}{K_d}\right)^2 + R^2\left(\frac{\Delta L_0}{L_0}\right)^2\left(\frac{L_0}{K_d}\right)^2 + \left[\left(\frac{1}{1-R}\right)\left(\frac{1}{R}+R\frac{L_0}{K_d}\right)\right]^2\Delta R^2 \\ &= \left(\frac{R}{1-R}\right)^2\left(\frac{\Delta T_0}{T_0}\right)^2\left[\frac{(1/R-1)K_d+(1-R)L_0}{K_d}\right]^2 + R^2\left(\frac{\Delta L_0}{L_0}\right)^2\left(\frac{L_0}{K_d}\right)^2 + \left[\left(\frac{1}{1-R}\right)\left(\frac{1}{R}+R\frac{L_0}{K_d}\right)\right]^2\Delta R^2 \\ &= \left(1+R\frac{L_0}{K_d}\right)^2\left(\frac{\Delta T_0}{T_0}\right)^2 + R^2\left(\frac{\Delta L_0}{L_0}\right)^2\left(\frac{L_0}{K_d}\right)^2 + \left[\left(\frac{1}{1-R}\right)\left(\frac{1}{R}+R\frac{L_0}{K_d}\right)\right]^2\Delta R^2\end{aligned} \quad \text{(B40)}$$

After setting $|\Delta T_0/T_0| = t$ and $|\Delta L_0/L_0| = l$, **Eq (B40)** can be expanded and rearranged to be:

$$\left(\frac{\Delta K_d}{K_d}\right)^2 = \left[t^2+\left(\frac{\Delta R}{R(1-R)}\right)^2\right] + \left[2Rt^2+2\left(\frac{\Delta R}{1-R}\right)^2\right]\frac{L_0}{K_d} + \left[R^2t^2+R^2l^2+\left(\frac{R\Delta R}{1-R}\right)^2\right]\left(\frac{L_0}{K_d}\right)^2 \quad \text{(B41)}$$

When $0 \le \Delta R < 0.5$, by inserting **Eqs (B19)** and **(B24)** into **Eq (B41)**, we obtain:

$$\left(\frac{\Delta K_d}{K_d}\right)^2 = \left[t^2+\left(\frac{r(1+r)}{0.5+r}\right)^2\right] + \left[\frac{t^2}{1+r}+2\left(\frac{0.5r}{0.5+r}\right)^2\right]\frac{L_0}{K_d} + \left[\left(\frac{0.5}{1+r}\right)^2\left(t^2+l^2+\left(\frac{0.5r}{0.5+r}\right)^2\right)\right]\left(\frac{L_0}{K_d}\right)^2 \quad \text{(B42)}$$

or

$$\left|\frac{\Delta K_d}{K_d}\right| = \sqrt{\left[t^2+\left(\frac{r(1+r)}{0.5+r}\right)^2\right] + \left[\frac{t^2}{1+r}+2\left(\frac{0.5r}{0.5+r}\right)^2\right]\frac{L_0}{K_d} + \left[\left(\frac{0.5}{1+r}\right)^2\left(t^2+l^2+\left(\frac{0.5r}{0.5+r}\right)^2\right)\right]\left(\frac{L_0}{K_d}\right)^2} \quad \text{(B43)}$$

When $-0.5 < \Delta R < 0$, by inserting **Eqs B19** and **(B25)** into **Eq (B41)**, we obtain:

$$\left(\frac{\Delta K_d}{K_d}\right)^2 = \left[t^2+\left(\frac{r(1-r)}{r-0.5}\right)^2\right] + \left(\frac{t^2}{1-r}+2\left(\frac{0.5r}{r-0.5}\right)^2\right)\frac{L_0}{K_d} + \left[\left(\frac{0.5}{1-r}\right)^2\left(t^2+l^2+\left(\frac{0.5r}{r-0.5}\right)^2\right)\right]\left(\frac{L_0}{K_d}\right)^2 \quad \text{(B44)}$$

or

$$\left|\frac{\Delta K_d}{K_d}\right| = \sqrt{\left[t^2+\left(\frac{r(1-r)}{r-0.5}\right)^2\right] + \left(\frac{t^2}{1-r}+2\left(\frac{0.5r}{r-0.5}\right)^2\right)\frac{L_0}{K_d} + \left[\left(\frac{0.5}{1-r}\right)^2\left(t^2+l^2+\left(\frac{0.5r}{r-0.5}\right)^2\right)\right]\left(\frac{L_0}{K_d}\right)^2} \quad \text{(B45)}$$

We can combine **Eqs (B43)** and **(B45)** as:

$$\left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = \sqrt{\alpha^2 + \lambda\frac{L_0}{K_{\mathrm{d}}} + \beta^2\left(\frac{L_0}{K_{\mathrm{d}}}\right)^2},$$

$$\alpha = \sqrt{t^2 + \left(\frac{r + r^2}{0.5 + r}\right)^2},\ \beta = \frac{0.5}{1+r}\sqrt{t^2 + l^2 + \left(\frac{0.5r}{0.5+r}\right)^2},\ \lambda = \frac{t^2}{1+r} + 2\left(\frac{0.5r}{0.5+r}\right)^2 \text{ when } 0 \le \Delta R < 0.5$$

$$\alpha = \sqrt{t^2 + \left(\frac{r - r^2}{r - 0.5}\right)^2},\ \beta = \frac{0.5}{1-r}\sqrt{t^2 + l^2 + \left(\frac{0.5r}{r-0.5}\right)^2},\ \lambda = \frac{t^2}{1-r} + 2\left(\frac{0.5r}{r-0.5}\right)^2 \text{ when } -0.5 < \Delta R < 0 \quad \text{(B46)}$$

Note B1.5: Derivation of the Theoretical Dependence of Determined K_d ($K_{d,det}$) on L_0 for Uncorrelated (Orthogonal) Error Sources

By inserting **Eq (B31)** into **Eq (B46)**, we get:

$$\left|\frac{K_{\mathrm{d,det}} - K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = \sqrt{\alpha^2 + \lambda\frac{L_0}{K_{\mathrm{d}}} + \beta^2\left(\frac{L_0}{K_{\mathrm{d}}}\right)^2} \quad \text{(B47)}$$

By multiplying K_{d} on both sides of **Eq (B47)**, we have:

$$\left|K_{\mathrm{d,det}} - K_{\mathrm{d}}\right| = \sqrt{\alpha^2 {K_{\mathrm{d}}}^2 + \lambda K_{\mathrm{d}} L_0 + \beta^2 {L_0}^2} \quad \text{(B48)}$$

When $\Delta K_{\mathrm{d}} \ge 0$,

$$\begin{aligned} K_{\mathrm{d,det}} - K_{\mathrm{d}} &= \sqrt{\alpha^2 {K_{\mathrm{d}}}^2 + \lambda K_{\mathrm{d}} L_0 + \beta^2 {L_0}^2} \\ \Rightarrow K_{\mathrm{d,det}} &= K_{\mathrm{d}} + \sqrt{\alpha^2 {K_{\mathrm{d}}}^2 + \lambda K_{\mathrm{d}} L_0 + \beta^2 {L_0}^2} \end{aligned} \quad \text{(B49)}$$

From **Eq (B49)**, we can infer that, at low L_0 range, $K_{\mathrm{d,det}}$ approaches a constant:

$$K_{\mathrm{d,det}} = K_{\mathrm{d}} + \alpha K_{\mathrm{d}} = (1+\alpha)K_{\mathrm{d}} \quad \text{(B50)}$$

Or we can write,

$$K_{\mathrm{d,det}} = \gamma,\ \gamma = (1+\alpha)K_{\mathrm{d}} \quad \text{(B51)}$$

At high L_0 range, the dependency of $K_{\mathrm{d,det}}$ on L_0 approaches a linear equation:

$$K_{\mathrm{d,det}} = \beta L_0 \tag{B52}$$

When $\Delta K_\mathrm{d} < 0$,

$$\begin{aligned} K_\mathrm{d} - K_{\mathrm{d,det}} &= \sqrt{\alpha^2 {K_\mathrm{d}}^2 + \lambda K_\mathrm{d} L_0 + \beta^2 {L_0}^2} \\ \Rightarrow K_{\mathrm{d,det}} &= K_\mathrm{d} - \sqrt{\alpha^2 {K_\mathrm{d}}^2 + \lambda K_\mathrm{d} L_0 + \beta^2 {L_0}^2} \end{aligned} \tag{B53}$$

From **Eq (B53)**, we can infer that, at low L_0 range, $K_{\mathrm{d,det}}$ approaches a constant:

$$K_{\mathrm{d,det}} = K_\mathrm{d} - \alpha K_\mathrm{d} = (1-\alpha) K_\mathrm{d} \tag{B54}$$

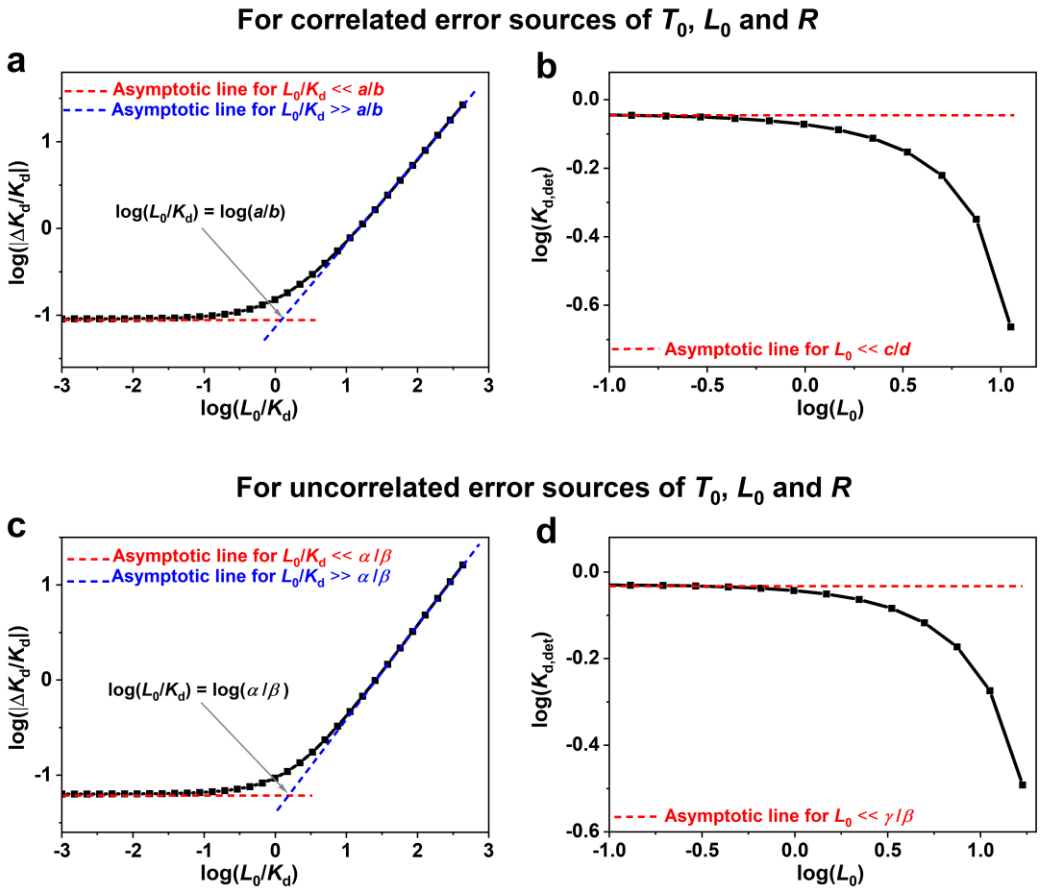
Or we use the same parameter γ to express **Eq (B54)** as,

$$K_{\mathrm{d,det}} = \gamma,\ \gamma = (1-\alpha) K_\mathrm{d} \tag{B55}$$

The dependency of $K_{\mathrm{d,det}}$ on L_0 at high L_0 range is unavailable because $K_{\mathrm{d,det}}$ is required to be greater than zero.

Figure B1.1: General Trends in Dependencies of Log($\lvert\Delta K_d/K_d\rvert$) on Log($L_0/K_d$) and Log($K_{d,det}$) on Log($L_0$)

Panels **(a)** and **(c)** show the general trends in dependency of $\log(\lvert\Delta K_\mathrm{d}/K_\mathrm{d}\rvert)$ on $\log(L_0/K_\mathrm{d})$; panels **(b)** and **(d)** show the general trends in dependency of $\log(K_{\mathrm{d,det}})$ on $\log(L_0)$. In all panels, reasonable relative errors of L_0, T_0 and R were used: $\lvert\Delta T_0/T_0\rvert = \lvert\Delta L_0/L_0\rvert = 0.05$ and $\Delta R/R = -0.02$. In panels **(b)** and **(d)**, we set $K_\mathrm{d} = 1$ (unitless) and $\Delta K_\mathrm{d} < 0$. With these conditions, the parameters a, b, c and d were calculated to be 0.011, 0.040, 0.99 and –0.040, respectively; α, β and γ were calculated to be 0.064, 0.037 and 0.94, respectively.



With **Eqs (B30), (B36), (B46)** and (**B51**), we produced representative graphs to show the dependence of $|\Delta K_d/K_d|$ on L_0/K_d and the dependence of $K_{d,det}$ on L_0 in double-log scale. The graphs corresponding to $0 \leq \Delta R < 0.5$ and $\Delta K_d \geq 0$ are shown in **Figure 3.2** in the main text, and the graphs corresponding to $-0.5 < \Delta R < 0$ and $\Delta K_d < 0$ are shown in **Figure B1.1**. **Figures 3.2a, 3.2c** (in main text) and **Figures B1.1a**, **B1.1c** show the same general trends in dependency of $\log(|\Delta K_d/K_d|)$ on $\log(L_0/K_d)$. Although **Figures B1.1b** and **B1.1d** cannot show the part at high L_0 range (where $K_{d,det}$ equals physically unfeasible negative values), the results in **Figure B1.1b** and **B1.1d** suggest a dependency of $K_{d,det}$ on L_0 that is similar to the trend shown in **Figures 3.2b** and **3.2d** (in main text): $K_{d,det}$ is insensitive to the change of L_0 when L_0 is at low concentrations, while the sensitivity of $K_{d,det}$ to L_0 increases with the increase of L_0.

Note B1.6: Derivation of the Relationship Between Signal to Noise Ratio (S/N) and the Limit of Quantitation (LOQ)

The limit of quantitation (LOQ) of an instrument is defined as the analyte concentration (ligand concentration in our case) for which the signal-to-noise ratio (S/N) is equal to a certain value $X >> 1$ (*e.g.*, $X = 10$) which guarantees that the noise does not affect quantitation significantly, which can be expressed as:

$$\frac{S_{\mathrm{LOQ}}}{N} = X \Rightarrow S_{\mathrm{LOQ}} = XN \tag{B56}$$

where S_{LOQ} is the signal measured at LOQ. For one instrument and a specific ligand, since the measured signal S is proportional to the ligand concentration, we can assume:

$$S_{\mathrm{LOQ}} = k \cdot \mathrm{LOQ} \Rightarrow k = \frac{S_{\mathrm{LOQ}}}{\mathrm{LOQ}} = \frac{XN}{\mathrm{LOQ}} \tag{B57}$$

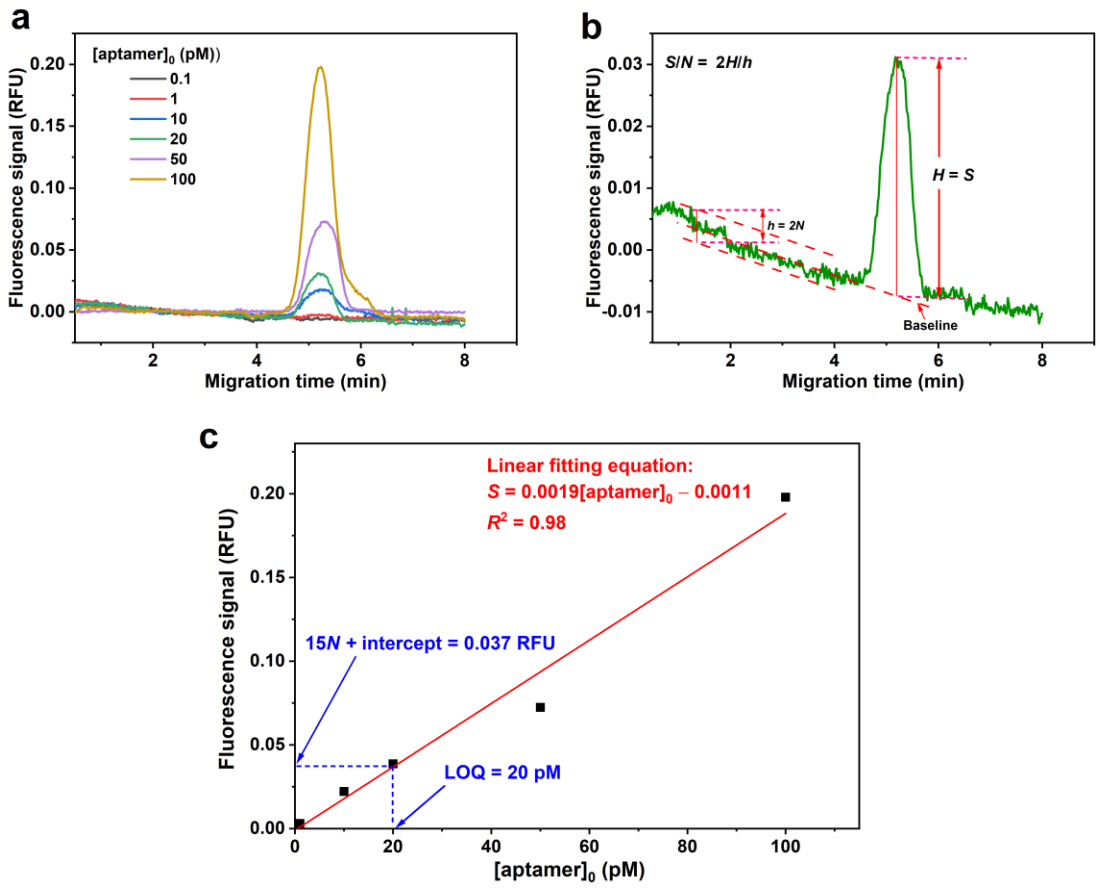
where factor k is the ratio of measured signal to ligand concentration. Now, if the noise N is independent of the signal S, the signal to noise ratio for any total ligand concentration L_0 can be calculated as:

$$\begin{aligned} &\frac{S}{N} = \frac{k \cdot L_0}{N} = \frac{XN}{\mathrm{LOQ}} \cdot \frac{L_0}{N} \\ &\Rightarrow \frac{S}{N} = X \frac{L_0}{\mathrm{LOQ}} \end{aligned} \tag{B58}$$

The minimum acceptable ligand concentration which satisfies **Eq (B58)** is $L_0 = \mathrm{LOQ}$.

Figure B1.2: Determination of Limit of Quantitation (LOQ)

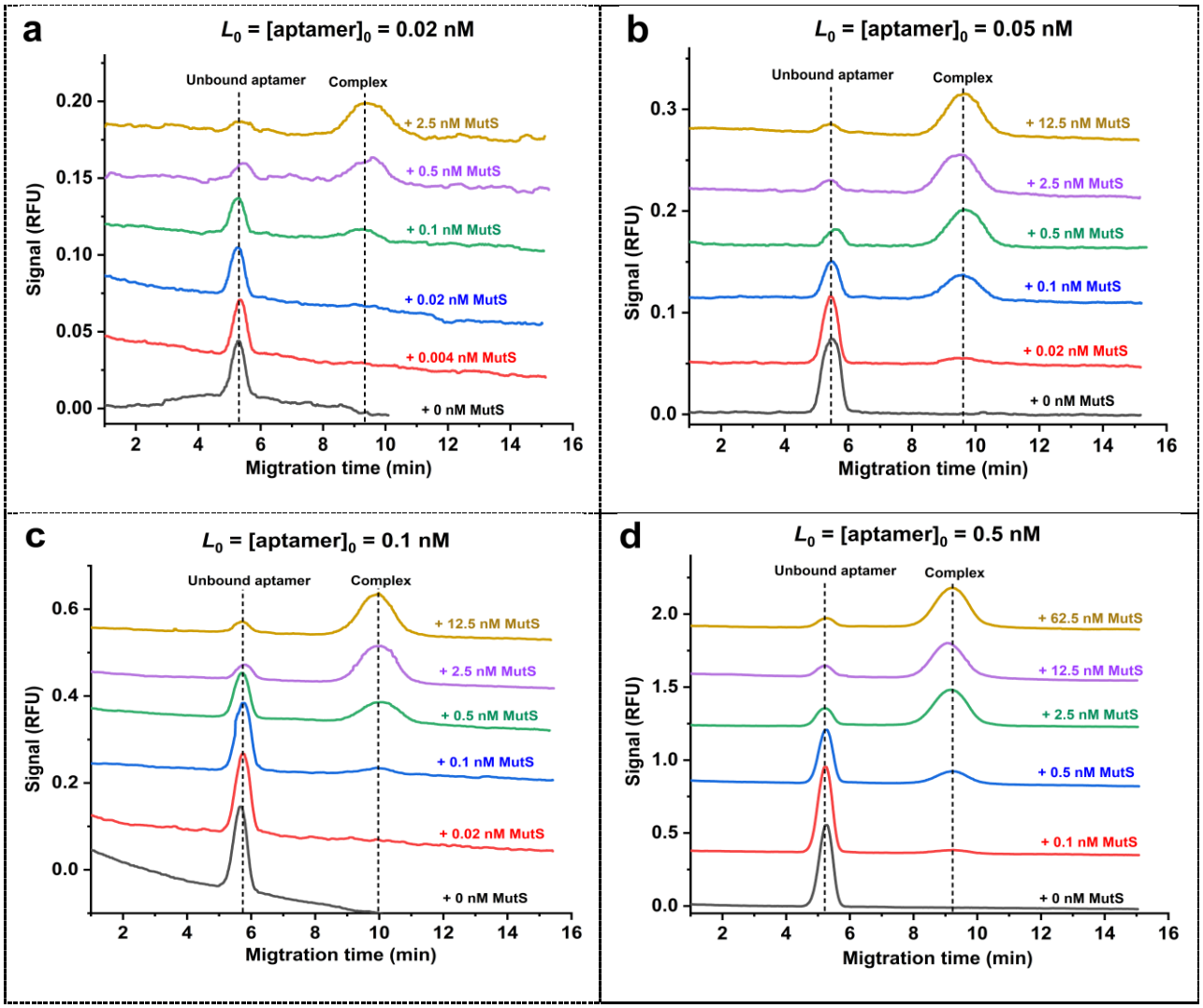
a) Electropherograms of the measurements for samples of MutS-binding aptamer with concentrations of 0.1, 1, 10, 20, 50, and 100 pM. The instrumental setup and experimental procedure are described in "Materials and Methods" section in the main text. **b)** An example ($[aptamer]_0$ = 20 pM) of determining signal and noise for an electropherogram: the intensity difference between the highest noise peak and the lowest noise peak was determined as the noise height *h*, which equals 2 times noise *N*; the midline of the noise peaks was determined as the baseline; the intensity difference between the highest signal of the signal peak and the baseline was determined as signal height *H*, which equal the signal *S*. **c)** Determining LOQ based on the linear plotting of "signal versus $[aptamer]_0$".

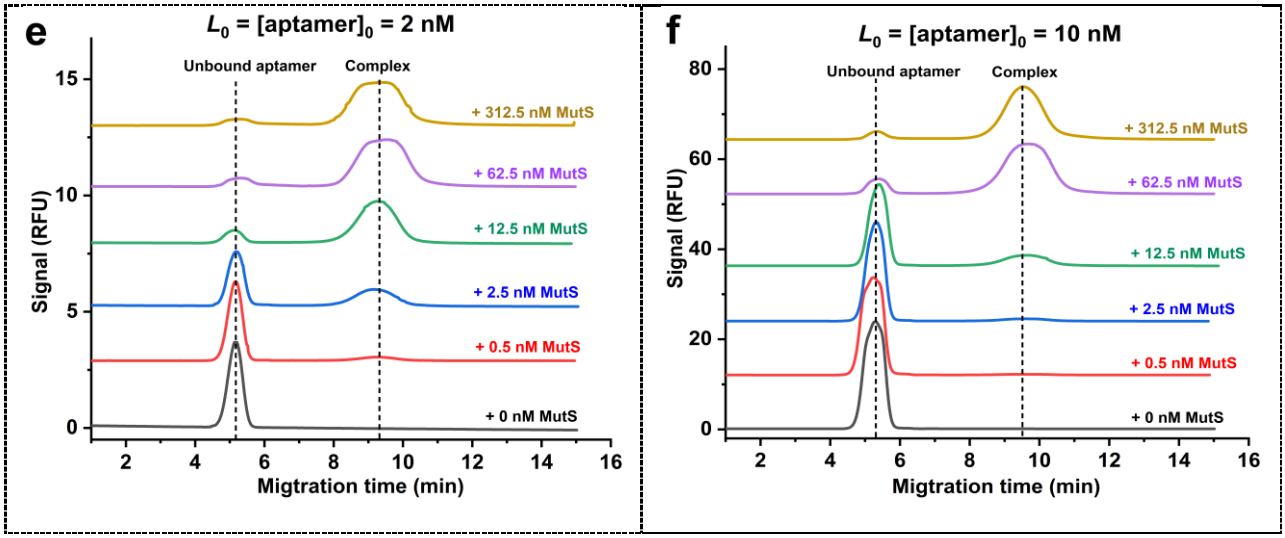


To determine the LOQ of our laser-induced fluorescence (LIF) detection system for the fluorescently labeled MutS-binding aptamer, we measured the aptamer samples with concentrations (L_0 or $[\text{aptamer}]_0$) of 0.1, 1, 10, 20, 50, and 100 pM. In the measurements, the instrumental setup and experimental procedure were identical to those used in the NECEEM experiments. The obtained electropherograms are shown in **Figure B1.2a**. Based on the electropherograms, the signals of aptamer at each concentration were determined to be 0.0020, 0.0031, 0.022, 0.039, 0.072, and 0.20 RFU, respectively. By analyzing all six electropherograms, the average noise was determined to be 0.0025 RFU. One example of determining signal (S), noise (N), and S/N is demonstrated in **Figure B1.2b**. Then, the results of "signal versus $[\text{aptamer}]_0$" were plotted (**Figure B1.2c**), and a linear fit was applied to the data points. Conventionally, the analyte concentration corresponding to a signal-to-noise ratio (S/N) of 10 (*i.e.*, $X = 10$ in **Eq (B56)**) is defined as LOQ. In this work, considering the signal of unbound aptamer decreases with the increase of protein concentration, we chose the $[\text{aptamer}]_0$ corresponding to $S/N = 15$ as LOQ. For different applications, one can choose different values for X in **Eq (B57)** to achieve a desired accuracy and precision of measured signals at LOQ. By inserting the signal equal to 15 times of average noise plus the intercept of the fitting line (for correcting the signal offset caused by background signal) into **Figure B1.2c**, LOQ for this work was determined to be 20 pM (*i.e.*, 0.02 nM).

Figure B1.3: Electropherograms of the NECEEM Experiments

Panels **a**–**f** show the electropherograms of the NECEEM experiments with L_0 ([aptamer]$_0$) of 0.02, 0.05, 0.1, 0.5, 2, and 10 nM, respectively. The instrumental setup and experimental procedure are explained in the "Materials and Methods" section (**Section 3.1.2**) in the main text.





Based on the electropherograms shown in **Figure B1.3**, the areas of the peaks for unbound aptamer and complex were measured with NECEEM Area Analysis Program (NAAP) that can be downloaded from https://www.yorku.ca/skrylov/resources.html. In each set of NECEEM experiments, since the corresponding highest MutS concentration was much greater than the expected K_d value range of $\leq$ 0.1nM, we presumed that the highest MutS concentration in each set of NECEEM experiments saturated the binding to the aptamer, and the remaining unbound aptamer was the impurity (*i.e.*, non-binders) in the aptamer sample. Therefore, from each unbound aptamer area, we subtracted the unbound aptamer area in corresponding "aptamer + highest MutS concentration" electropherogram to obtain the actual unbound aptamer area. Then, the determined areas were corrected by dividing them by their migration times. Based on the corrected areas, the actual fractions of unbound aptamer (R values) at different T_0 ([MutS]$_0$) were calculated, and the results are shown in **Table B1.1**. The binding isotherms for the NECEEM experiments are shown in **Figure 3.3a** in the main text.

Table B1.1. Results of R values at different T_0 ([MutS]$_0$) for the NECEEM experiments at different L_0 ([aptamer]$_0$).

T_0 (nM)	R (L_0 = 0.02 nM)	R (L_0 = 0.05 nM)	R (L_0 = 0.1 nM)	R (L_0 = 0.5 nM)	R (L_0 = 2 nM)	R (L_0 = 10 nM)
0	1.0	1.0	1.0	1.0	1.0	1.0
0.004	0.96	N/A	N/A	N/A	N/A	N/A
0.02	0.81	0.86	1.0	N/A	N/A	N/A
0.1	0.42	0.45	0.89	0.94	N/A	N/A
0.5	0.17	0.11	0.52	0.73	0.92	0.99
2.5	0.0	0.066	0.073	0.15	0.65	0.97
12.5	N/A	0.0	0.0	0.053	0.10	0.84
62.5	N/A	N/A	N/A	0.0	0.041	0.098
312.5	N/A	N/A	N/A	N/A	0.0	0.0

Appendix B2. Supporting Information for "Maximizing the Accuracy of Equilibrium Dissociation Constants (K_d) for Affinity Complexes: from Theory to Practical Recommendations"

Note B2.1: Definition of Signal Additivity and the Calculation of the Fraction of Unbound Ligand R with Additive Signals

In K_d determinations, we call the measured signal (S) "additive" when it can be directly decomposed into the sum of signals contributed by the target-unbound ligand (*i.e.*, free ligand L) and target-bound ligand (or complex C), *i.e.*, S_L and S_C, such as:

$$S = S_L + S_C \tag{B59}$$

If Q_L and Q_C are quantities of L and C in units of quantity, e.g., moles, then,

$$S = S_{L,unit} Q_L + S_{C,unit} Q_C \tag{B60}$$

where $S_{L,unit}$ and $S_{C,unit}$ are the signals from unit quantities of unbound ligand and complex (*i.e.*, target-bound ligand), respectively. If Q is the total quantity of L and C, such as:

$$Q = Q_L + Q_C \tag{B61}$$

then,

$$\begin{aligned} & S = (S_{L,unit} Q)\frac{Q_L}{Q} + (S_{C,unit} Q)\frac{Q_C}{Q} \\ & \Rightarrow S = S_{L,\text{per quantity } Q}\frac{Q_L}{Q} + S_{C,\text{per quantity } Q}\frac{Q_C}{Q} \\ & \Rightarrow S = S_{L,\text{per quantity } Q} f_L + S_{C,\text{per quantity } Q} f_C \end{aligned} \tag{B62}$$

where f_L and f_C are the fractions of L and C in the total quantity of them.

Let's assume the total quantity Q is a constant, S_L* is the signal from quantity Q of L, and S_C* is the signal from quantity Q of C, then **Eq (B62)** is converted to:

$$S = S_{\mathrm{L}}{}^{*} f_{\mathrm{L}} + S_{\mathrm{C}}{}^{*} f_{\mathrm{C}} \tag{B63}$$

If the unit of quantity is concentration, then **Eq (B63)** can be expressed as:

$$S = S_{\mathrm{L}}{}^{*} \frac{L}{L+C} + S_{\mathrm{C}}{}^{*} \frac{C}{L+C} \tag{B64}$$

where L and C are equilibrium concentrations of the ligand and complex, respectively. If the total concentration of ligand L_0 is a constant (*i.e.*, $L_0 = L + C = const$), then **Eq (B64)** can be rewritten as:

$$S = S_{\mathrm{L}}{}^{*} \frac{L}{L_0} + S_{\mathrm{C}}{}^{*} \frac{C}{L_0} \tag{B65}$$

According to the definition of the fraction of unbound ligand R (**Eq (2.1)**: $R = L/L_0$), **Eq (B65)** is rewritten as:

$$S = S_{\mathrm{L}}{}^{*} R + S_{\mathrm{C}}{}^{*} (1 - R) \tag{B66}$$

By solving R from **Eq (B66)**, we obtain:

$$R = \frac{S - S_{\mathrm{C}}{}^{*}}{S_{\mathrm{L}}{}^{*} - S_{\mathrm{C}}{}^{*}} \tag{B67}$$

which is **Eq (3.13)** in the main text. For using **Eq (B67)** to calculate R values (for constructing binding isotherms), **Eq (B59)** must be satisfied, which is the additivity of signals. Note that, **Eq (B63)** is applicable to surface-based methods, *e.g.*, SPR (Surface Plasmon Resonance) and BLI (Biolayer Interferometry), as well.

Note B2.2: Proof of Additivity of Signals That Linearly Depend on Concentrations

Assuming the detected signal is linearly related to components concentration, (*e.g.* fluorescence):

$$S_{\mathrm{L}} = \alpha L + N, \quad S_{\mathrm{C}} = \beta C + N$$
$$S_{\mathrm{L}}{}^{*} = \alpha L_0 + N, \quad S_{\mathrm{C}}{}^{*} = \beta C_0 + N \tag{B68}$$

where S_{L} is the signal from unbound ligand, S_{C} is the signal from complex (*i.e.*, target-bound ligand), L is the concentration of unbound ligand, C is the concentration of complex, α and β are the proportionality factors for ligand and complex, respectively, and N is the background that is (assumed to be) independent of components and their concentrations and thus only one N contributes to the cumulative signal S:

$$\begin{aligned} S &= S_{\mathrm{L}} + S_{\mathrm{C}} = \alpha L + \beta C + N \\ &= \alpha L_0 \frac{L}{L_0} + \beta L_0 \frac{C}{L_0} + N\frac{L_0}{L_0} = \alpha L_0 \frac{L}{L_0} + \beta L_0 \frac{C}{L_0} + N\frac{L+C}{L_0} \\ &= \alpha L_0 \frac{L}{L_0} + N\frac{L}{L_0} + \beta L_0 \frac{C}{L_0} + N\frac{C}{L_0} \\ &= \left(\alpha L_0 + N\right)\frac{L}{L_0} + \left(\beta L_0 + N\right)\frac{C}{L_0} \\ &= S_{\mathrm{L}}{}^{*}\frac{L}{L_0} + S_{\mathrm{C}}{}^{*}\frac{C}{L_0} \end{aligned} \tag{B69}$$

which is the superposition equation (**Eq (3.12)** in main text) required for an additive signal. Note that, if the background noise is dependent on the concentration of any reactant, *e.g.* donor/acceptor concentration in FRET (Fluorescence Resonance Energy Transfer), the background noise must be subtracted to calculate S, S_{L} and S_{C}. Otherwise, the measured signal would be non-additive.

Note B2.3: Sensitivity of Determined K_d ($K_{d,det}$) to Random Errors in Variables for R Approaching Either Zero or Unity

If K_{d} is determined with a single R value, the following formula is used:

$$K_{\mathrm{d}} = \frac{T_0 - L_0(1-R)}{(1/R-1)} \tag{B70}$$

By applying the regular error propagation rule to **Eq (B70)**, we have:

$$\begin{aligned}\delta K_{\mathrm{d,det}} &= \sqrt{\left(\frac{\partial K_{\mathrm{d}}}{\partial T_0}\right)^2 \delta {T_0}^2 + \left(\frac{\partial K_{\mathrm{d}}}{\partial L_0}\right)^2 \delta {L_0}^2 + \left(\frac{\partial K_{\mathrm{d}}}{\partial R}\right)^2 \delta R^2} \\ &= \frac{\sqrt{\delta {T_0}^2 + \delta {L_0}^2 (1-R)^2 + \delta R^2 \left(L_0(1/R-1) + \left(T_0 - L_0(1-R)\right)/R^2\right)^2}}{(1/R-1)} \\ &= \frac{\sqrt{\delta {T_0}^2 R^2 + \delta {L_0}^2 \left((1-R)R\right)^2 + \delta R^2 \left(L_0(1-R) + \left(T_0 - L_0(1-R)\right)/R\right)^2}}{1-R}\end{aligned} \tag{B71}$$

It has been proven that, with the single-point approach, $K_{\mathrm{d,det}}$ is the most accurate for $R \approx ½$. Now let's investigate the sensitivity of $K_{\mathrm{d,det}}$ to random errors of R, T_0, and L_0 when R value approaches two extrema: 0 and 1.

When R approaches 0,

$$\begin{aligned}\lim_{R\to 0} \delta K_{\mathrm{d,det}} &= \frac{\sqrt{\delta {T_0}^2 0^2 + \delta {L_0}^2 \left((1-0)\times 0\right)^2 + \delta R^2 \left(L_0(1-0) + \left(T_0 - L_0(1-0)\right)/0\right)^2}}{1-0} \\ &= \sqrt{\delta R^2 \left(L_0 - \infty\right)^2} \\ &= \infty\end{aligned} \tag{B72}$$

When R approaches 1,

$$\begin{aligned}\lim_{R\to 1} \delta K_{\mathrm{d,det}} &= \frac{\sqrt{\delta {T_0}^2 1^2 + \delta {L_0}^2 \left((1-1)\times 1\right)^2 + \delta R^2 \left(L_0(1-1) + \left(T_0 - L_0(1-1)\right)/1\right)^2}}{1-1} \\ &= \frac{\sqrt{\delta {T_0}^2 + \delta R^2 {T_0}^2}}{0} \\ &= \infty\end{aligned} \quad . \tag{B73}$$

Eqs (B72) and **(B73)** suggest that when R approaches 0 or 1, any random error of R (or T_0) can be largely magnified when propagated to the random error of the $K_{\mathrm{d,det}}$. If single-point K_{d}-determination experiments are repeated many times, the large random error of $K_{\mathrm{d,det}}$ (obtained for R being close to 0 or 1) will result in very poor precision of the determined K_{d} values. If the single-point K_{d}-determination experiment is only conducted once or repeated a few times, the

large random error of $K_{d,det}$ (obtained for R being close to 0 or 1) becomes the systematic error of $K_{d,det}$, resulting in very poor accuracy (and precision) of the determined K_d values.

Note B2.4: Non-additivity of Polarization

Fluorescence anisotropy (r) has been proven to be an additive signal, which satisfies:

$$r = r_L \frac{L}{L_0} + r_C \frac{C}{L_0} \tag{B74}$$

If polarization is additive, it must satisfy:

$$P = P_L \frac{L}{L_0} + P_C \frac{C}{L_0} \tag{B75}$$

where P_L and P_C are the polarizations of pure ligand and pure complex, respectively, with a concentration of L_0. However, by replacing r in **Eq (B74)** with **Eq (3.24)** in the main text, we obtain:

$$\frac{2P}{3-P} = \frac{2P_L}{3-P_L}\frac{L}{L_0} + \frac{2P_C}{3-P_C}\frac{C}{L_0} \tag{B76}$$

which cannot be converted to **Eq (B76)** with any transformations. Therefore, polarization is non-additive.

Note B2.5: The Dependence of Relative Systematic Error of $K_{d,det}$ on L_0/K_d for Using Non-Additive Polarization to Calculate R

To investigate the effect of using non-additive polarization to calculate R on the accuracy of $K_{d,det}$, we first used Data S1 (on the Figshare website https://doi.org/10.6084/m9.figshare.25464685) to produce the theoretical binding isotherms "*R*

versus T_0" for L_0/K_d from 0.01 to 1311 with a 2-fold step size. According to the definition of R (**Eq (2.1)**), **Eq (B74)** is rewritten as:

$$r = r_L R + r_C(1-R) \tag{B77}$$

By assuming r_L and r_C to be 0 and 0.4, respectively, anisotropy r corresponding to each R value was calculated. **Eq (3.24)** in the main text can be transformed to be an expression of polarization P with anisotropy r:

$$P = \frac{3r}{2+r} \tag{B78}$$

Using **Eq (B78)**, polarizations of pure ligand (P_L) and pure complex (P_C), and P corresponding to each r were computed. With P_L, P_C, and P calculated from each r, the R values (R') determined with non-additive P for different T_0 were calculated with:

$$R' = \frac{P - P_C}{P_L - P_C} \tag{B79}$$

and formed the binding isotherms "R' versus T_0" for $L_0/K_{d,true}$ from 0.01 to 1311. By fitting isotherms "R' versus T_0" with **Eq (2.2)** in the main text, $K_{d,det}$ obtained by using non-additive polarization to calculate R for different L_0/K_d were determined. The resulted dependence of accuracy of $K_{d,det}$ (*i.e.*,$|\Delta K_{d,det}/K_d|$) on L_0/K_d is summarized in **Figure B2.1.**

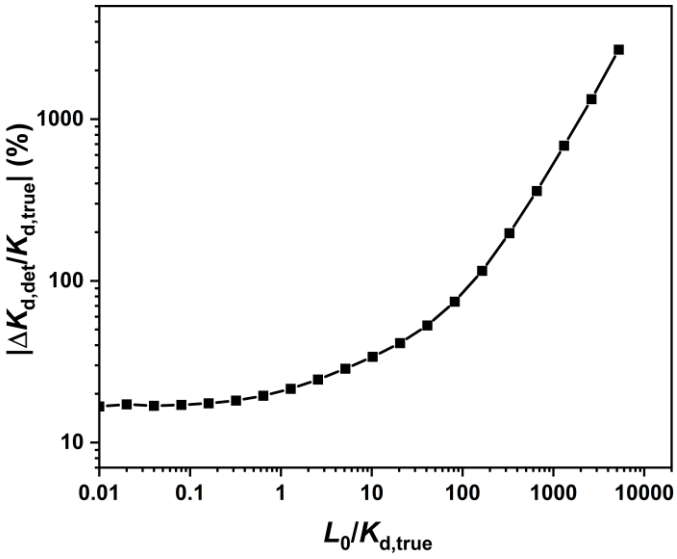


Figure B2.1. The dependence of relative systematic error of $K_{d,det}$ (*i.e.*,$|\Delta K_{d,det}/K_d|$) on L_0/K_d for using non-additive polarization to calculate R.

Note B2.6: Non-additivity of Diffusion Coefficients

B2.6.1: Determination of K_d Using Taylor Dispersion Analysis

Taylor dispersion analysis is a method that has been widely used to characterize the diffusion coefficient of analytes in a solute within a laminar flow pipe**.** Taylor described the concentration profile of the dispersion of an injected plug as given by**:**

$$C(x,t) = y_0 + A\exp\left(-\frac{1}{2}\frac{(t-t_R)^2}{\sigma^2}\right) \qquad \text{(B80)}$$

where $C(x,t)$ is the signal profile averaged over the cross-section of the flow pipe at time t and distance x; the fitting parameters y_0, A, t_R and σ^2 are the baseline, amplitude, average elution time and temporal variance respectively.

The value of σ^2 in **Eq (B80)** is related to the apparent diffusion coefficient D_{app} of the diffusing analyte through:

$$D_{app} = r^2 t_0 / (24\sigma^2) \tag{B81}$$

where r is the radius of the flow pipe.

Taylor dispersion analysis has been used to find K_d of an equilibrium mixture (EM) *via* the determination of the apparent diffusion coefficient D_{app}. To obtain $R = L/L_0$, from which $K_{d,det}$ can be obtained, an additivity assumption is used:

$$D_{app} = D_{app,L}\frac{L}{L_0} + D_{app,C}\frac{C}{L_0} = D_{app,L}R + D_{app,C}(1-R) \tag{B82}$$

where $D_{app,L}$ and $D_{app,C}$ are the apparent diffusion coefficients for pure ligand L and pure complex C, respectively, and D_{app} is the apparent diffusion coefficient of L (resulting from both unbound and protein-bound) in the EM.

B2.6.2: Equilibrium Mixture (EM) Ternary System

The dispersion of the EM involves 5 components: pure L, pure complex C, pure protein P and the solvent. Assuming that P and C are indistinguishable in their diffusion properties, this system can be reduced to a simpler ternary system (two components L and C in one solvent). This system is described by 4 diffusion coefficients (**Eq (B83)**): D_L and D_C which are the self-diffusion coefficients of L and C, respectively and $D_{C,L}$ and $D_{L,C}$ are the cross-diffusion coefficients describing the coupling between the diffusion flux of L and C.

$$\hat{D} = \begin{bmatrix} D_L & D_{L,C} \\ D_{C,L} & D_C \end{bmatrix} \tag{B83}$$

The apparent (*i.e.* observable) diffusion coefficients of L and C are given by the eigenvalues of the diffusion matrix (**Eq (B83)**) as**:**

$$D_{\mathrm{app,L}} = \frac{D_{\mathrm{L}} + D_{\mathrm{C}} + \sqrt{(D_{\mathrm{L}} - D_{\mathrm{C}})^2 + 4D_{\mathrm{L,C}}D_{\mathrm{C,L}}}}{2}$$
$$D_{\mathrm{app,C}} = \frac{D_{\mathrm{L}} + D_{\mathrm{C}} - \sqrt{(D_{\mathrm{L}} - D_{\mathrm{C}})^2 + 4D_{\mathrm{L,C}}D_{\mathrm{C,L}}}}{2} \quad \text{(B84)}$$

It is worth noting that the (apparent) self-diffusion coefficients $D_{\mathrm{app,L}}$ and $D_{\mathrm{app,C}}$ in **Eq (B82)** correspond to their binary diffusion coefficients D_{L} and D_{C} in the limit where the concentration of their partner component, *i.e.* C and L respectively, goes to zero. In general, cross-diffusion coefficients are functions of both the type and concentration of species and may have an impact on the resulting apparent diffusion coefficients. However, the cross-diffusion coefficients can be neglected in very dilute solutions, *i.e.* solutions where the molar fraction of the solvent is much greater than the molar fractions of L and C. The above considerations result in:

$$D_{\mathrm{app,L}} = D_{\mathrm{L}}$$
$$D_{\mathrm{app,C}} = D_{\mathrm{C}} \approx D_{\mathrm{P}} \quad \text{(B85)}$$

To describe the ternary system, an extension of the original Taylor dispersion expression for binary systems (**Eq (B80)**) was formulated. The adapted concentration profile $S(t)$ of such system, taking into consideration of **Eq (B85)** becomes:

$$S(t) = (t_{\mathrm{R}}/t)^{1/2}\left[\frac{B_1}{B_1+B_2}\exp\left(-\frac{12D_{\mathrm{app,L}}(t-t_{\mathrm{R}})^2}{r^2 t}\right) + \frac{B_2}{B_1+B_2}\exp\left(-\frac{12D_{\mathrm{app,C}}(t-t_{\mathrm{R}})^2}{r^2 t}\right)\right] \quad \text{(B86)}$$

where

$$B_1 = \left[\begin{array}{l}(D_{\mathrm{C}} - D_{\mathrm{C,L}})R \\ +(D_{\mathrm{L}} - D_{\mathrm{L,C}})(1-R) - D_{\mathrm{app,L}}\end{array}\right] D_{\mathrm{app,L}}^{1/2}$$

$$B_2 = -\left[\begin{array}{l}(D_{\mathrm{L}} - D_{\mathrm{L,C}})(1-R) \\ +(D_{\mathrm{C}} - D_{\mathrm{C,L}})R - D_{\mathrm{app,C}}\end{array}\right] D_{\mathrm{app,C}}^{1/2}$$

B2.6.3: Simulation Parameters

The simulation parameters used in **Eq (B86)** were: t_R = 150 s and r = 25 μm, which are typical scales used in Taylor dispersion analysis to satisfy conditions under which Taylor solution is valid. D_L and D_C were chosen to be 425 and 50 $\mu m^2/s$, respectively, which are in the range of typical values for diffusion coefficients of L and P

To simulate titration curves, R values obtained from the theoretical dependence of R on L_0, T_0 and K_d, were used in **Eq (B85)**. The expression for R is given by:

$$R = -\frac{K_d + T_0 - L_0}{2L_0} + \sqrt{\left(\frac{K_d + T_0 - L_0}{2L_0}\right)^2 + \frac{K_d}{L_0}} \qquad \text{(B87)}$$

B2.6.4: Binding Isotherms from Simulated Signals

The simulations were performed for K_d = 1 μM and L_0 = 0.01 and 100 μM. First, the signals $S(t)$ for different R values were simulated with Data S1 (https://doi.org/10.6084/m9.figshare.25464685). Then, the simulated signals were fitted with **Eq (B80)** to find σ^2. With the determined σ^2, the apparent diffusion coefficients D_{app} corresponding to different T_0 were calculated with **Eq (B81)**. Subsequently, the apparent R values (R_{app}) were calculated with **Eq (B82)** according to three types of signals: *i*) the signal of EM with pure L to find $D_{app,L}$ (*e.g.*, **Figure B2.2A**); *ii*) the signal of EM with pure C to find $D_{app,C}$ (*e.g.*, **Figure B2.2B**), and *iii*) the signal of EM with a mixture of L and C to find D_{app} (*e.g.*, **Figure B2.2C**). By fitting the binding isotherms "R_{app} versus T_0", the K_d values ($K_{d,det}$) determined with the assumption of additivity of D_{app} were found. The comparisons of the resulting binding isotherm "R_{app} versus T_0" (along with $K_{d,det}$) to the theoretical binding isotherms "R versus T_0" (along with input K_d) for L_0 = 0.01 and 100 μM are shown in **Figure 3.8** in the main text.

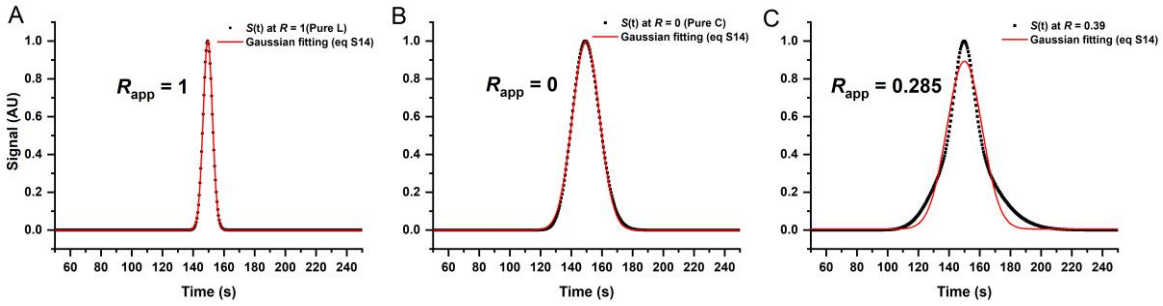


Figure B2.2. Examples of finding apparent diffusion coefficients along with apparent R values by fitting the simulated diffusion signals obtained with the analytical solution of Taylor dispersion. In these three examples, the input K_d = 1 μM and L_0 = 0.01 μM.

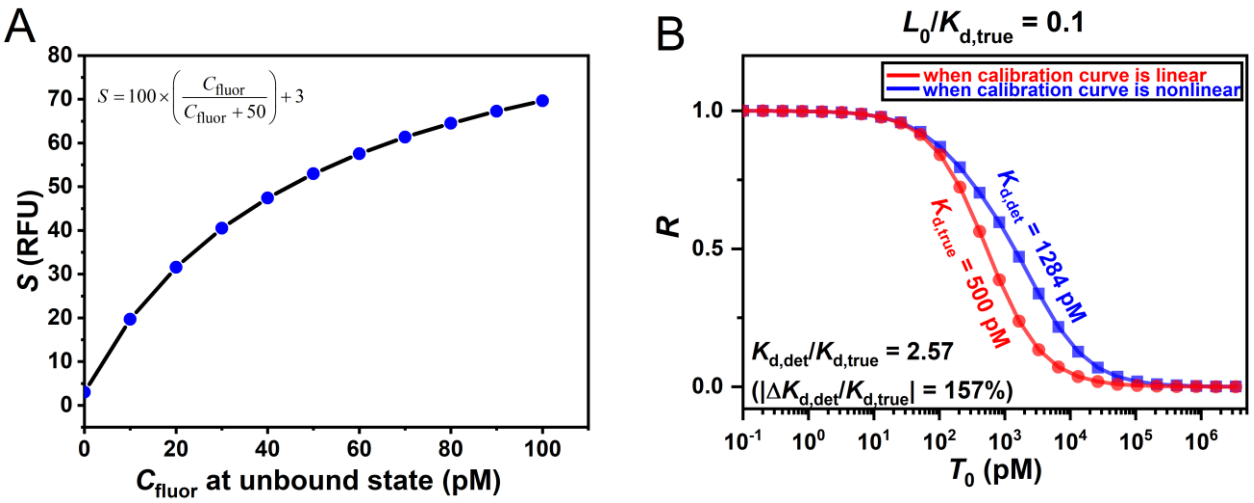
Note B2.7: Demonstration of the Effect of Mis-calibrated Instrument on the Accuracy of $K_{d,det}$

In this demonstration, we use fluorescence detection as an example. As we showed in Note B2.1, for fluorescence signal to be additive, it should be linearly related to the concentration of fluorophore, *i.e.*, the calibration curve of "fluorescence versus fluorophore concentration" should be linear. Otherwise, the detected signal can be non-additive. Here, we assume the dependence of fluorescence on fluorophore concentration is nonlinear, such as:

$$S = \lambda\left(\frac{C_{\text{fluor}}}{C_{\text{fluor}} + 50\ \text{pM}}\right) + 3 \tag{B88}$$

where S is the detected fluorescence signal with unit of RFU (Relative Fluorescence Unit), C_{fluor} is the concentration of fluorophore with unit of pM, λ is a magnification factor (with unit of RFU) that depends on the quantum yield of the fluorophore at different binding states. The constant "3" is the magnitude of background noise with unit of RFU.

Here, we set λ = 100 RFU for unbound ligand, and λ = 50 RFU for bound ligand caused by fluorescence quenching. The resulted (nonlinear) calibration curve **(A)** and the discrepancies of binding isotherms and determined K_d values (at L_0/K_d = 0.1) caused by the instrumental



nonlinear response to fluorophore concentrations **(B)** are demonstrated here. In the simulation, the input K_d was 500 pM.

Figure B2.3. Demonstration of the effect of mis-calibrated instrument on the accuracy of $K_{d,det}$: **(A)** An exemplified nonlinear calibration curve produced by **Eq (B88)** with setting λ = 100 RFU and 50 RFU for unbound and bound ligand, respectively; **(B)** the discrepancies of binding isotherms and determined K_d values (at $L_0/K_{d,true}$ = 0.1) caused by the exemplified instrumental nonlinear response to fluorophore concentrations. In the simulation, the input $K_{d,true}$ was 500 pM.

Note B2.8: Grating Factor in Fluorescence Anisotropy

Experimentally, the fluorescence anisotropy (r) of a fluorophore is determined with:

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}} \tag{B89}$$

where I_{VV} and I_{VH} are the detected emission intensities with vertical and horizontal polarizations, respectively, when the excitation light is vertically polarized. G is a grating factor used to correct the instrumental bias on vertically and horizontally polarized lights, which can be determined by:

$$G = \frac{I_{HV}}{I_{HH}} \tag{B90}$$

where I_{HV} and I_{HH} are the emission intensities with vertical and horizontal polarizations, respectively, when the excitation light is horizontally polarized. Any systematic error in G will lead to inaccurate r, and consequently translate into systematic errors in R and $K_{d,det}$. To study the effect of inaccurate G on the accuracy of $K_{d,det}$, we demonstrate the dependence of systematic error in $K_{d,det}$ on $L_0/K_{d,true}$ with adding an experimentally reasonable 10% relative systematic error in G (*i.e.*, $\Delta G/G = 0.1$) (**Figure B2.4**). **Figure B2.4** shows that the effect of the inaccuracy of G on the accuracy of $K_{d,det}$ is severe at high $L_0/K_{d,true}$, *e.g.*, at $L_0/K_{d,true} = 10^3$ and10% systematic error in G, the relative systematic error in $K_{d,det}$ is greater than 100%. Since the true value of K_d ($K_{d,true}$) is unknown *a priori* in a K_d-determination experiment, $L_0/K_{d,true}$ may be large even when using the lowest suitable L_0 for quantification, *i.e.*, L_0 = LOQ. Thus, determining an accurate G factor is crucial for determining accurate K_d with fluorescence anisotropy.

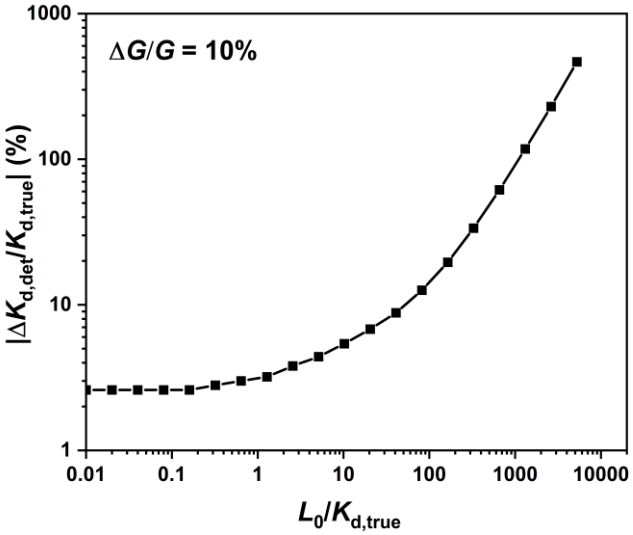


Figure B2.4. The dependence of relative systematic error of $K_{d,det}$ (*i.e.*,$|\Delta K_{d,det}/K_{d,true}|$) on $L_0/K_{d,true}$ with adding 10% relative systematic error in G factor (*i.e.*, $\Delta G/G =$ 10%) for anisotropy-based K_d determination.

To measure the grating factor G, as suggested by **Eq (B89)**, the excitation light must be horizontally polarized. Horizontally polarized excitation light ensures that the fluorophores' exited-state distribution is along the detection axis, thereby equalizing the intensities of emitted light with different polarization orientations on the plane of detection, which is perpendicular to the travel direction of the emitted light. In this scenario, any difference in the detected intensities of vertically and horizontally polarized emissions is caused by instrumental bias, which can be corrected by the determined G factor. Since G is dependent on wavelength and instrumental setup, its value should be re-determined when any optical component, *e.g.*, emission filter and excitation/emission polarizer, is changed in the instrument.

Note B2.9: Sufficiency of Incubation Time

For the lowest non-zero T_0, the pseudo-first order conditions of $T_0 << L_0$ and $L \approx L_0$ are satisfied. Thus, the exponential function of the dependence of the concentration of the formed complex ($C(t)$) on time (t) can be expressed as:

$$C(t) = C_{\text{eq}}\left(1 - e^{-t/t_{\text{eq}}}\right) \tag{B91}$$

with C_{eq} as the concentration of complex at equilibrium (*i.e.*, at $t = \infty$). By replacing T_0 with L_0, the classic characteristic equilibration time (t_{eq}) is converted to:

$$t_{\text{eq}} = \left(k_{\text{on}}L_0 + k_{\text{off}}\right)^{-1} \tag{B92}$$

To reach $C(t) \geq 95\%C_{\text{eq}}$, we have:

$$
\begin{aligned}
C_{\mathrm{eq}}\left(1-e^{-t_{\mathrm{inc}}/t_{\mathrm{eq}}}\right) &\geq 0.95C_{\mathrm{eq}} \\
1-e^{-t_{\mathrm{inc}}/t_{\mathrm{eq}}} \geq 0.95 &\Rightarrow e^{-t_{\mathrm{inc}}/t_{\mathrm{eq}}} \leq 0.05 \\
-\frac{t_{\mathrm{inc}}}{t_{\mathrm{eq}}} \leq \ln(0.05) &\Rightarrow t_{\mathrm{inc}} \geq -\ln(0.05)t_{\mathrm{eq}} \\
t_{\mathrm{inc}} &\geq 3t_{\mathrm{eq}}
\end{aligned}
\tag{B93}
$$

Therefore, the incubation time t_{inc} should be equal or longer than three times of characteristic equilibration time t_{eq} to reach ≥ 95% equilibration. For example, in a K_{d}-determination experiment, the K_{d} of the studied binding pair is 10^{-10} M with a k_{on} of 10^5 $\mathrm{M^{-1}s^{-1}}$, and the experimentally used (constant) total concentration of ligand $L_0 = 10^{-11}$ M which is the LOQ of the instrument for detection. With **Eq (1.2)** in the main text, k_{off} can be calculated as 10^{-5} $\mathrm{s^{-1}}$. Then, with the definition of t_{eq} in **Eq (B92)**, t_{eq} is calculated as 9.1×10^4 s, which is ~25 h. As a result, the sufficient incubation time is $t_{\mathrm{inc}} \geq 75$ h ($3t_{\mathrm{eq}}$), which is more than 3 days.

Note B2.10: Experimental and Mathematical Approaches to Estimate Sufficient Incubation Time

The sufficient incubation time can be determined experimentally. Some kinetic methods, *e.g.*, stopped-flow spectroscopy, can monitor a binding reaction in real-time, and intuitively show if the reaction reaches equilibrium. If a real-time reaction monitoring method capable of detecting the signal from the target (with low LOQ) is available in the lab, researchers can employ it to conduct a preliminary study for determining the incubation time with a mixture of ligand with concentration L_0 and target with concentration T_0 ($<< L_0$) as a sample. Here, L_0 and T_0 are the constant ligand concentration and lowest nonzero target concentration planned for the K_{d}-determination experiment. Since the target is the limiting component with a smaller concentration in this preliminary study, only the signal from the target should be measured, and

sufficient incubation time can be assumed to be the time for the reaction to reach 95% completion.

If a real-time reaction monitoring method is unavailable in the lab, researchers can measure the mixture of ligand with concentration L_0 and detectable target (with the lowest nonzero T_0) with a certain frequency and construct a curve for the dependence of the target's signal on time to determine the time after which the signal stops changing significantly, using this time as the incubation time. However, if there is no means of solely measuring the signal from the target or the lowest nonzero T_0 is smaller than the instrument's LOQ for the target, researchers should conduct multiple K_d-determination experiments with different incubation times. When $K_{d,det}$ stops decreasing with the increase of incubation time, it can be concluded that the incubation time is sufficient for all samples to reach equilibrium.

All experimental methods mentioned above for determining the sufficient incubation time require specialized equipment and expertise and are generally resource-intensive. Alternatively, a simple mathematical estimation with the help of **Eq (B92)** can serve the purpose of finding sufficient incubation time well in many cases. For example, when planning to study a molecular pair with an expected K_d value in the nanomolar range (*i.e.*, 1 – 1000 nM) using an instrument with LOQ = 0.1 nM for the ligand, the ligand concentration L_0 is chosen to be the LOQ (0.1 nM) to minimize the error of $K_{d,det}$ translated from the errors of variables. To assess the sufficient incubation time, k_{on} can be assumed to be near an "average" value of ~10^6 $M^{-1}s^{-1}$ since generally k_{on} ranges from 10^3 to 10^9 $M^{-1}s^{-1}$. Then, by assuming the true K_d value ($K_{d,true}$) to be the lowest possible value of 0.1 nM, k_{off} can be estimated (with **Eq (1.2)**, main text) to be around 10^{-4} s^{-1}. With all these considerations and **Eq (B92)**, t_{eq} is calculated to be 1.4 h, and accordingly, the minimum incubation time ($3t_{eq}$) is calculated to be 4.2 h. This mathematical estimation of

sufficient incubation time can save researchers time and experimental resources However, since true k_{on} can be much smaller than $10^6\ M^{-1}s^{-1}$, ideally at least one more K_d-determination experiment should be conducted with a longer incubation time to confirm that $K_{d,det}$ is stabilized.

Note B2.11: Common Sources of Systematic Errors in Concentrations

To emphasize the importance of eliminating the common sources of systematic errors in concentrations (*i.e.*, L_0 and T_0) used in K_d determination, here we simulated the effect of experimentally reasonable systematic errors in T_0 and L_0 on the accuracy of $K_{d,det}$ for a large range of $L_0/K_{d,true}$ (**Figure B2.5**). **Figure B2.5** suggests that, without minimizing the systematic errors in concentrations, the experimentally determined $K_{d,det}$ can potentially be orders of magnitude different from $K_{d,true}$ for high $L_0/K_{d,true}$. The most common sources of systematic

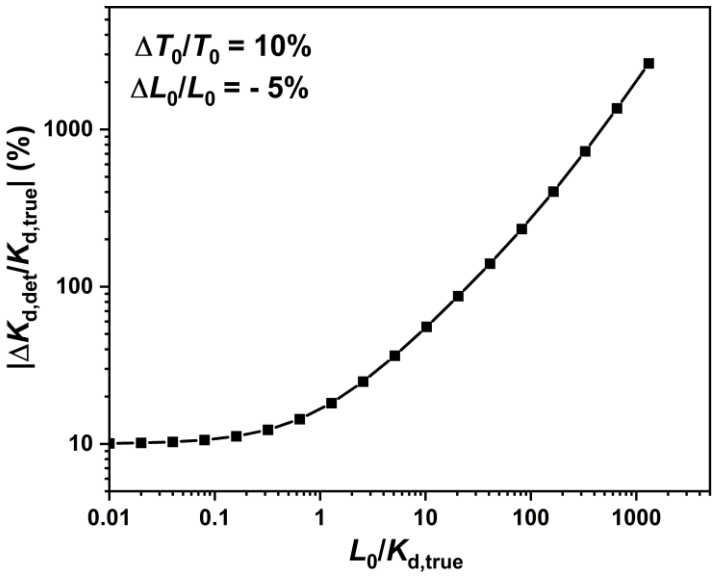


Figure B2.5. The dependence of relative systematic error of $K_{d,det}$ (*i.e.*,$|\Delta K_{d,det}/K_{d,true}|$) on $L_0/K_{d,true}$ with adding 10% relative systematic error in T_0 and –5% relative systematic error in L_0, *i.e.*, the nominal T_0 is 10% higher than true T_0 and the nominal L_0 is 5% lower than true L_0.

errors in concentrations include mis-calibration and improper operation of measuring equipment, product impurity, and solute adsorption.

B2.11.1: Mis-calibration and Improper Operation of Measuring Equipment

K_d determination using the nonlinear regression approach entails preparing a series of equilibrium mixtures with constant L_0 and varied T_0. Experimentally, all the equilibrium mixtures are prepared from the high-concentration stock solutions of ligand and target.Thus, any systematic errors in concentrations of stock solutions will translate to systematic errors in L_0 and T_0 for all equilibrium mixtures. In certain experiments, a stock solution is created by dissolving a specific weight of solid material into a volume of solvent (*e.g.*, buffer) to achieve a desired concentration. In such cases, the mis-calibration or improper operation of mass- and volume-measuring equipment can introduce systematic errors in the concentration of a stock solution.

B2.11.1.1: Calibration and Operation of Mass-measuring Equipment

Laboratory mass measuring equipment, including top-loading balances, portable balances, analytical balances, semi-micro balances, and microbalances, should be calibrated and operated meticulously. Before using a laboratory balance, it is essential to carefully read the user manual provided by the manufacturer to understand how to use and calibrate the device correctly. Regular calibration of any type of lab balance is crucial to minimize systematic errors in measured mass.

Traditionally, lab balances are manually calibrated by adjusting the balance readings using to standard weights. This manual calibration process is time-consuming and usually conducted once per month. However, many modern lab balances now come with built-in automatic self-calibration functionality, which offers convenience and time efficiency. Balances equipped with this function can be calibrated more often, even before every measurement. It is noteworthy that

the most reliable calibration for balances with self-calibration is a combination of manual and automatic calibrations.

In addition to proper calibration, users must adhere to certain rules to ensure high accuracy and precision when using a lab balance. These rules dictate:

i) level the balance perfectly,

ii) keep the device clean,

iii) place the sample in a suitable container for weighing to avoid damaging the weighing pan,

iv) tare the balance to subtract the weight of the sample container,

v) avoid vibration and heavy airflow around the balance during weight measurement,

vi) keep the balance away from any strong magnetic field.

vii) maintain the balance in an environment with a stable temperature (*i.e.*, room temperature).

Failure to comply with these rules can potentially result in significant systematic errors in measured mass.

B2.11.1.2: Calibration and Operation of Volume-measuring Equipment

A wide array of specialized volume-measuring equipment is usually available in well-equipped laboratories, including beakers, flasks, graduated cylinders, burettes, and pipettes. Since most studied binding partners are valuable materials (*e.g.*, proteins), well-established K_d-determination approaches typically require small quantities. Therefore, using volume pipettes alone is often sufficient for volume measurements in a K_d-determination experiment. To ensure high accuracy and precision in measured volumes, volume pipettes must be routinely calibrated and operated correctly. Pipettes should be calibrated every 3–6 months, and after thorough

cleaning. Detailed guidelines on correctly calibrating and operating pipettes are provided as follows.

The calibration of a pipette relies on the relationship between the volume and mass of distilled water aspirated/dispensed by the pipette. Therefore, calibrating volume pipettes necessitates the use of an analytical balance that has been properly calibrated and is operated correctly, as discussed in Note B2.5. Pipettes are calibrated using professional tools by adjusting the readings to correspond to the volumes of distilled water aspirated and dispensed, as weighed by a highly accurate analytical balance. This conversion between mass and volume is facilitated by the density of distilled water (volume = mass/density), which, assuming negligible effects of atmospheric pressure on water density, varies with environmental temperature.

Typically, a pipette is calibrated for both its lower- and upper-limit volumes. If the lower- and upper-limit volumes cannot be calibrated simultaneously, the pipette should not be used for volume measurement and should either be replaced or sent back to the vendor for maintenance.

When using a calibrated pipette, it is essential to adhere to the following rules to enhance the accuracy of pipetting:

i) pre-wet the pipette tip at least three times before aspirating the final volume,

ii) hold the pipette at a consistent angle not exceeding 20 degrees to the vertical,

iii) before aspirating, immerse the pipette tip to an appropriate depth (based on the volume to aspirate/dispense) below the liquid surface to ensure contact with the liquid throughout the aspiration process,

iv) touch off after each dispense,

v) pipette viscous liquids slowly, while pipetting volatile liquids quickly.

Failure to use calibrated pipettes or to follow the aforementioned rules can result in inaccuracies in measured liquid volume.

B2.11.2: Product Impurity

The purified material of binding partners (target and ligand) is either obtained from commercial vendors or produced internally in a laboratory. The quality and purity of the purified material vary from batch to batch. Since any impurity or degradation of binding partners will result in systematic errors in concentrations, preventing reagent degradation and determining their purity are crucial for reducing systematic errors in L_0 and T_0.

To avoid reagent degradation, it is crucial to store reagents under appropriate conditions and adhere to specific guidelines. For instance, for long-term storage of single-stranded DNA (ssDNA) with high concentration (> 100 ng/μL), the ssDNA suspension should be divided into small aliquots with small aliquots (*e.g.*, 2 μL) and stored at –20 ºC for no longer than 2 years. Repeated freeze-thaw cycles for DNA solutions should be avoided since they can lead to DNA degradation. Additionally, long-term exposure to light, including UV light and ambient lab light, should be avoided for fluorophore-modified ssDNA to avoid photobleaching. A working solution of ssDNA with a lower concentration can be stored in a refrigerator at +4 ºC for one year and should be placed on ice when used to prepare samples on the bench. Similarly, to prevent preserving proteins, solutions with high concentrations (> 1000 ng/μL) or lyophilized proteins should be divided into small aliquots and stored at –80 ºC for long-term storage. Repeated freeze-thaw cycles for lyophilized protein and protein solutions should also be avoided.

Due to the propensity of microorganisms to grow in protein solutions under nonfreezing conditions, working solutions of proteins with a lower concentration can only be stored at +4 ºC for days to weeks at the longest. Improper storage of reagents can significantly reduce the

concentrations of active ligand and target molecules compromising their native structures and unavoidably introducing large systematic errors in L_0 and T_0.

To minimize systematic errors in concentrations, researchers often employ spectrophotometry (Lambert-Beer law) to measure "true" concentrations. However, this method has two limitations: (*i*) spectrophotometry cannot distinguish impurities or contaminations from the product, and (*ii*) systematic errors in molar extinction coefficients will translate into systematic errors in concentrations. Thus, if the purities of ligand and target products are unknown *a priori*, they must be determined using analytical approaches. For example, high-performance liquid chromatography (HPLC) and CE can quantitatively determine the purity of ssDNA, and sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and liquid chromatography-mass spectrometry (LC-MS) can estimate the purity of proteins. Based on the quantitatively determined purities of products, the concentrations calculated from measured mass and volume or determined by spectrophotometry need to be corrected.

On the other hand, determining the accurate concentration of a pure product in solution using the Lambert-Beer law requires knowledge of the accurate molar extinction coefficient for the component in solution, which is influenced by various chemical conditions such as buffer type, ionic strength, and pH. Hence, when using a widely accepted molar extinction coefficient to calculate the concentration of a component, one should ensure that the conditions of the measured solution are similar (if not exactly the same) to conditions used to determine the molar extinction coefficient. Otherwise, the concentration determined with the Lambert-Beer law according to a known molar extinction coefficient should be considered inaccurate. In many cases, accurate molar extinction coefficients for studied components are unavailable; thus, the

systematic errors in L_0 and T_0 can only be reduced by correcting the concentrations based on the experimentally determined purities of the reagents.

B2.11.3: Solute Adsorption

Systematic errors in concentrations can also result from solute adsorption, *i.e.*, reagent adsorption, which occurs in most steps of a K_d-determination experiment. Reagent adsorption can significantly reduce reagent concentration at low concentrations, which is typically the case for L_0 = LOQ and T_0 in a low-concentration range. Solute adsorption to pipette tips and sample vials can occur during sample preparation, while solute adsorption to fluidic paths and channels may occur during sample analysis. Although reagent adsorption cannot be fully eliminated, it can be minimized by careful experimental design and optimization.

The most common and straightforward method to mitigate reagent adsorption and nonspecific binding to surfaces is by incorporating blocking reagents such as BSA protein, Tween® 20, and Triton X-100 into solutions. The addition of blocking agents to solutions has been proven to be an effective strategy for reducing nonspecific bindings.

Since the extent of reagent adsorption to a surface is significantly influenced by the surface's chemical properties, instrumental parts made of adsorption-resistant materials, such as titanium and polyether ether ketone (PEEK), are preferred for use along the fluidic path.

Furthermore, surface modification with biocompatible coatings is a widely adopted and effective approach for reducing reagent adsorption. Coatings such as Bovine Serum Albumin (BSA), Polyethylene glycol (PEG), and polyvinyl alcohol (PVA) have been shown to effectively reduce DNA and protein adsorption.

To minimize reagent adsorption to the inner walls of vials, some manufacturers have developed various types of low-binding tubes. These tubes can be used in conjunction with other

adsorption elimination methods to further mitigate reagent losses resulting from solute adsorption.

Table B2.1: Checklist for Minimizing the Systematic Errors of Variables in K_d Determination with Nonlinear Regression

This checklist is an extended version of the checklist proposed by Jarmoskaite et al. (2020). This checklist is also stored on Figshare (https://doi.org/10.6084/m9.figshare.25464685) as a PDF file.

Checklist for Improving the Accuracy of Determined K_d		
Ligand (limiting component): _____ **Target (excess component):** _____ **K_d-determination method:** _____ **CONDITIONS: Temperature:** _____ **Buffer:** _____ **Other:** _____		
ABBREVIATIONS AND SYMBOLS: ***R***, fraction of unbound ligand; **ΔR**, systematic error of R; **$\Delta K_{d,det}$**, systematic error of determined K_d; **$K_{d,det}$**, experimentally determined K_d; ***L***, equilibrium concentration of ligand; ***L*₀**, total concentration of ligand; ***T***, equilibrium concentration of target; ***T*₀**, total concentration of target; ***C***, equilibrium concentration of target–ligand complex; ***S***, measured cumulative signal; ***Q***, total quantity of ligand and complex; ***Q*L**, the quantity of ligand at equilibrium; ***Q*C**, the quantity of complex at equilibrium; **$S_{L,\text{per quantity } Q}$**, signal of ligand with quantity Q; **$S_{C,\text{per quantity } Q}$**, signal of complex with quantity Q; **LOQ**, limit of quantitation; **EM**, equilibrium mixture.		
☐ **Minimize ΔR** (and minimize $\Delta K_{d,det}$ caused by too high L_0)	• Has the additivity of the signal used to calculate R been proven? Additivity of signal: $S = S_{L,\text{per quantity } Q}\frac{Q_L}{Q} + S_{C,\text{per quantity } Q}\frac{Q_C}{Q}$	☐ Yes. Continue. ☐ No/Not sure. Prove the additivity or change to a different method.
	• Has the instrument used to detect signals been calibrated?	☐ Yes. Continue. ☐ No/Not sure. Calibrate the instrument.
	• Is the LOQ of the instrument for the ligand known?	☐ Yes. Use L_0 = LOQ. ☐ No. Determine LOQ.
	• Vary incubation time to confirm equilibration. Time range: _____ Number of time points: _____ ➢ Is $K_{d,det}$ independent to the tested incubation time?	☐ Yes. Incubation time is sufficient. ☐ No. Increase incubation time.
	• Does the binding isotherm "R vs T_0" reach saturation? ➢ *Check approach*: 1) Linear fit the data points of "S vs T_0" obtained from the EMs with the three highest T_0 to get a fitting equation of "$S = a + bT_0$" with a slope of "$b \pm \delta b$", *i.e.*, uncertainty range of slope is $[b - \delta b, b + \delta b]$. 2) If $0 \in [b - \delta b, b + \delta b]$, the binding isotherm reaches	☐ Yes. Continue. ☐ No. Increase the highest T_0 (if solubility allows).

	saturation. Note: The highest T_0 should be at least twofold greater than the lowest T_0 in the three EMs.	
☐ **Minimize ΔL_0 and ΔT_0**	• Have the measuring instruments (*e.g.*, balance and pipette) been calibrated?	☐ Yes. Continue. ☐ No/Not sure. Calibrate the measuring instruments.
	• Have the purities of ligand and target been measured with analytical approaches (*e.g.*, HPLC and SDS-PAGE)?	☐ Yes. Correct L_0 and T_0 based on the determined purities. ☐ No. Determine the purities of ligand and target.
	• Are the ligand and/or target prone to adsorption to surfaces? ➢ *Common measures to reduce reagent adsorptions*: - Adding blocking agents (*e.g.*, BAS and Tween® 20) into solutions. - Surface modifications with biocompatible coatings - Using instrumental parts and lab supplies that are made of low-binding materials.	☐ Yes/Not sure. Apply the measures of reducing reagent adsorptions. ☐ No.
Comments		

Table B2.2: Determination of the Random Error in the Concentration of Fluorescein (Ligand) Solutions

To demonstrate the determination of the random error in the concentration of fluorescein solutions, we prepared five 800 µL fluorescein stock solutions with nominal (desired) concentration of 300 µM. Each solution was prepared separately from scratch. For each sample, the absorbance at light wavelength of 488 nm (A_{488}) with path length l = 1 mm was measured with a NanoDrop 1000 (Thermo Scientific) spectrophotometer. The measured absorbance for

each sample and the calculated relative standard deviation (RSD) are summarized here. The absorbance of each sample was measured in triplicate.

Sample #	A_{488}	**Average**	**Overall Average**	**Standard Deviation (SD)**	**RSD ($\delta L_0/L_0$)**
1	1.721	1.72	1.69	0.025	0.015
	1.708				
	1.723				
2	1.694	1.69			
	1.675				
	1.685				
3	1.648	1.67			
	1.667				
	1.684				
4	1.658	1.66			
	1.654				
	1.666				
5	1.710	1.71			
	1.702				
	1.714				

Appendix C. Supplementary Information for Chapter 4

Appendix C1. Supporting Information for "A Practical Approach to Quantitatively Assessing Equilibrium-Constant Accuracy from a Single Binding Isotherm"

Note C1.1: Generalization of ACI-Concept

To generalize the ACI-Concept to scenarios involving the determination of physicochemical parameters with multiple independent variables held constant, we assume that the parameter p is a known function of measurable independent variables x_0, x_1, x_2, …, x_n (which can be expressed as a vector $\boldsymbol{x}$) and a measurable dependent variable, the signal S:

$$p = F(\boldsymbol{x}, S) \tag{C1}$$

Accordingly, **Eq (4.2)** in the main text becomes:

$$S = f(x_0, \boldsymbol{x_c}, p) \text{ with } \boldsymbol{x_c} = x_1, x_2, ..., x_n = const \tag{C2}$$

using p as a fitting parameter and fixing the rest of the independent variables, which can be represented as a vector $\boldsymbol{x_c}$. The vector of systematic errors in variables is defined as $\boldsymbol{\Delta} = \{\boldsymbol{\Delta x}, \Delta S\}$.

The sensitivity of the determined value p_{det} to changes in vector $\boldsymbol{\Delta}$ (regression stability) can be quantified by the ratio p_{det}'/p_{det}, where p_{det} and p_{det}' represent the values of parameter p determined at $\boldsymbol{x_c}/p$ for two different vectors of systematic errors in variables, such as $\boldsymbol{\Delta}$ and $\boldsymbol{\Delta'}$, respectively. Similar to the theory developed for the case with $\boldsymbol{x_c}$ as a one-dimensional scalar x_c, here we postulate that if the theoretical relation between p_{det}'/p_{det} and $\boldsymbol{x_c}/p$ can be solved analytically or numerically, the accurate value of the studied parameter can be computed by evaluating the regression stability and investigating the vector $\boldsymbol{x_c}$. The generalization for **Eqs (4.3)** and **(4.4)** in the main text only requires replacing scalar x_c with vector $\boldsymbol{x_c}$:

$$\begin{aligned} &\Delta p = G\ \ \boldsymbol{x_c}, \boldsymbol{\Delta} \\ &\Rightarrow \Delta p = p_{\text{det}} - p = G(\boldsymbol{x_c}, \boldsymbol{\Delta}) \\ &\Rightarrow p_{\text{det}} = G(\boldsymbol{x_c}, \boldsymbol{\Delta}) + p \\ &\Rightarrow p_{\text{det}} = g(\boldsymbol{x_c}, \boldsymbol{\Delta}) \end{aligned} \tag{C3}$$

Then, the expression linking a unitless accuracy ratio p_{det}/p to unitless ratio $\boldsymbol{x_c}/p$ and unitless relative errors of variables can be written as:

$$\frac{p_{\text{det}}}{p} = H\left(\frac{\boldsymbol{x_c}}{p}, \boldsymbol{\Delta}^*\right), \text{ where } \boldsymbol{\Delta}^* = \left\{\frac{\Delta x_0}{x_0}, \frac{\Delta x_1}{x_1}, \frac{\Delta x_2}{x_2}, ..., \frac{\Delta x_n}{x_n}, \frac{\Delta S}{S}\right\} \tag{C4}$$

To study the sensitivity of p_{det}/p to the errors of variables, we need to have at least two points, corresponding to errors $\boldsymbol{\Delta}^*$ and $\boldsymbol{\Delta}^{*\prime}$, where $\boldsymbol{\Delta}^{*\prime} - \boldsymbol{\Delta}^*$ should be large to ensure that the values of p_{det} obtained with these two errors are distinguishable. **Eq (C4)** can be written for $\boldsymbol{\Delta}^{*\prime}$ as:

$$\frac{p_{\text{det}}'}{p} = H\left(\frac{\boldsymbol{x_c}}{p}, \boldsymbol{\Delta}^{*\prime}\right) \tag{C5}$$

To investigate the sensitivity of the determined parameter to the errors of variables at different $\boldsymbol{x_c}/p$, we divide **Eq (C5)** over **Eq (C4)** to obtain:

$$\frac{p_{\text{det}}'}{p_{\text{det}}} = H\left(\frac{\boldsymbol{x_c}}{p}, \boldsymbol{\Delta}^{*\prime}\right) \Big/ H\left(\frac{\boldsymbol{x_c}}{p}, \boldsymbol{\Delta}^*\right) = h\left(\frac{\boldsymbol{x_c}}{p}, \boldsymbol{\Delta}^*, \boldsymbol{\Delta}^{*\prime}\right) \tag{C6}$$

The theoretical relation shown in **Eq (C6)** can be numerically tabulated by fitting simulated datasets (involving $\boldsymbol{\Delta}^*$) with **Eq (C2)** using $\boldsymbol{x_c}/p$ and $\boldsymbol{x_c}/p + (\boldsymbol{\Delta}^{*\prime} - \boldsymbol{\Delta}^*)$ as the variables kept constant. To determine the $\boldsymbol{x_c}/p$ ratio for fixed $\boldsymbol{\Delta}^*$ and $\boldsymbol{\Delta}^{*\prime}$, one can consider the inverse relationship between $p_{\text{det}}'/p_{\text{det}}$ and $\boldsymbol{x_c}/p$, as follows:

$$\frac{\boldsymbol{x_c}}{p} = h^{-1}\left(\frac{p_{\text{det}}'}{p_{\text{det}}}\right), \text{ at } \boldsymbol{\Delta}^* \text{ and } \boldsymbol{\Delta}^{*\prime} \tag{C7}$$

In practice, p_{det} and p_{det}' can be determined by using nominal values of $\boldsymbol{x_c}$ and $\boldsymbol{x_c} + (\boldsymbol{\Delta}^{*\prime} - \boldsymbol{\Delta}^*)$, respectively, in a regression procedure with a correct regression model (fitting **Eq (C2)** into

experimental data that includes experimental $\boldsymbol{\Delta^*}$). Note, in real experiments, the value of $\boldsymbol{\Delta^{*\prime} - \Delta^*}$ is controllable in the data fitting process.

We presume that the value of $p_{\text{det}}'/p_{\text{det}}$ can be obtained experimentally. Then, by plugging it into the numerically established **Eq (C7)** (involving the confidence interval of $\boldsymbol{\Delta^*}$), the confidence interval of $\boldsymbol{x}_{\mathbf{c}}/p$ can be determined. Finally, since the nominal $\boldsymbol{x}_{\mathbf{c}}$ is known, we can, in principle, determine the confidence interval of p, *i.e.*, ACI of p.

Eq (C6) is used to compute $p_{\text{det}}'/p_{\text{det}}$ as functions of multi-dimensional $\boldsymbol{x}_{\mathbf{c}}/p$ for the boundary systematic error vectors, $\boldsymbol{\Delta^*_{\min}}$ and $\boldsymbol{\Delta^*_{\max}}$ (which cause the smallest and largest accuracy ratio p_{det}/p in **Eq (C4)**, respectively), with a constant value of $\boldsymbol{\Delta^{*\prime} - \Delta^*}$ to build a multi-dimensional ACI graph. A known value of $p_{\text{det}}'/p_{\text{det}}$ is placed on the graph as a horizontal surface for multi-dimensional $\boldsymbol{x}_{\mathbf{c}}/p$, and the intersections of this surface with the two border surfaces define the boundary vectors of $(\boldsymbol{x}_{\mathbf{c}}/p)_{\min} = \{(x_1/p)_{\min}, (x_2/p)_{\min}, \ldots, (x_n/p)_{\min}\}$ and $(\boldsymbol{x}_{\mathbf{c}}/p)_{\max} = \{(x_1/p)_{\max}, (x_2/p)_{\max}, \ldots, (x_n/p)_{\max}\}$. Finally, the nominal vector of $\boldsymbol{x}_{\mathbf{c}} = \{x_1, x_2, \ldots, x_n\}$ is used to calculate the confidence interval in which p lies: between $p_{\min} = \min\{x_1/(x_1/p)_{\max}, x_2/(x_2/p)_{\max}, \ldots, x_n/(x_n/p)_{\max}\}$ and $p_{\max} = \max\{x_1/(x_1/p)_{\min}, x_2/(x_2/p)_{\min}, \ldots, x_n/(x_n/p)_{\min}\}$. The presented graphical strategy can be presumably applied to any parameter computed with the regression approach.

Note C1.2: Derivation of the Dependence of Accuracy Ratio of Determined K_d on the Relative Systematic Errors of Variables

Eq (3.1) in the main text shows that:

$$K_{\text{d}} = \frac{T_0 - L_0(1-R)}{(1/R-1)} \tag{C8}$$

When the error sources of T_0, L_0, and R are strongly correlated, the maximum systematic error of K_d can be calculated with the error propagation rule:

$$|\Delta K_d| = \left|\left(\frac{\partial K_d}{\partial T_0}\right)\Delta T_0\right| + \left|\left(\frac{\partial K_d}{\partial L_0}\right)\Delta L_0\right| + \left|\left(\frac{\partial K_d}{\partial R}\right)\Delta R\right| \quad \text{(C9)}$$

However, for calculating the systematic error of K_d propagated from specific systematic errors of variables, which can have different signs and results in ΔK_d of different signs, we should use the formula (without using absolute values):

$$\Delta K_d = \left(\frac{\partial K_d}{\partial T_0}\right)\Delta T_0 + \left(\frac{\partial K_d}{\partial L_0}\right)\Delta L_0 + \left(\frac{\partial K_d}{\partial R}\right)\Delta R \quad \text{(C10)}$$

By applying the error propagation rule **Eq (C10)** to **Eq (C8)**, we obtain:

$$\Delta K_d = \frac{R}{1-R}\Delta T_0 - R\Delta L_0 + \left(-L_0 + \frac{T_0}{(1-R)^2}\right)\Delta R \quad \text{(C11)}$$

By solving **Eq (C8)** for T_0, we obtain:

$$T_0 = (1/R - 1)K_d + (1-R)L_0 \quad \text{(C12)}$$

To obtain an analytical solution of error propagation for **Eq (C8)**, it's necessary to focus on a single data point by fixing R. Here, we only focus on the scenario of the determined R value being equal to 0.5 with a systematic error ΔR, which is close to that of the least erroneous case. Then, the actual R is:

$$R = 0.5 - \Delta R \quad \text{(C13)}$$

and accordingly, **Eq (C12)** is converted to:

$$T_0 = \frac{(0.5+\Delta R)K_d}{0.5-\Delta R} + (0.5+\Delta R)L_0 \quad \text{(C14)}$$

By dividing both sides of **Eq (C11)** by K_d, we have:

$$\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}=\frac{R}{1-R}\frac{\Delta T_0}{K_{\mathrm{d}}}-R\frac{\Delta L_0}{K_{\mathrm{d}}}+\left(-\frac{L_0}{K_{\mathrm{d}}}+\frac{1}{(1-R)^2}\frac{T_0}{K_{\mathrm{d}}}\right)\Delta R \tag{C15}$$

Eq (C15) can be rewritten as:

$$\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}=\frac{R}{1-R}\left(\frac{\Delta T_0}{T_0}\right)\frac{T_0}{K_{\mathrm{d}}}-R\frac{\Delta L_0}{L_0}\frac{L_0}{K_{\mathrm{d}}}+\left(-\frac{L_0}{K_{\mathrm{d}}}+\frac{1}{(1-R)^2}\frac{T_0}{K_{\mathrm{d}}}\right)\Delta R \tag{C16}$$

By inserting **Eq (C12)** and **Eq (C13)** into **Eq (C16)**, we get:

$$\begin{aligned}\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}=&\frac{0.5-\Delta R}{0.5+\Delta R}\left(\frac{\Delta T_0}{T_0}\right)\left[\frac{0.5+\Delta R}{0.5-\Delta R}+\frac{L_0}{K_{\mathrm{d}}}(0.5+\Delta R)\right]-(0.5-\Delta R)\frac{\Delta L_0}{L_0}\frac{L_0}{K_{\mathrm{d}}}\\&+\left[-\frac{L_0}{K_{\mathrm{d}}}+\frac{1}{(0.5+R)^2}\left(\frac{0.5+\Delta R}{0.5-\Delta R}+\frac{L_0}{K_{\mathrm{d}}}(0.5+\Delta R)\right)\right]\Delta R\end{aligned} \tag{C17}$$

To express ΔR through the relative systematic error of R (*i.e.*, $\Delta R/R$), we have:

$$\begin{aligned}\Delta R=R\frac{\Delta R}{R}&=(0.5-\Delta R)\frac{\Delta R}{R}\\\Rightarrow\Delta R&=0.5\frac{\Delta R}{R}-\Delta R\frac{\Delta R}{R}\\\Rightarrow\left(1+\frac{\Delta R}{R}\right)\Delta R&=0.5\frac{\Delta R}{R}\\\Rightarrow\Delta R&=\frac{0.5\dfrac{\Delta R}{R}}{1+\dfrac{\Delta R}{R}}\end{aligned} \tag{C18}$$

Finally, by inserting **Eq (C18)** into **Eq (C17)**, we obtain:

$$\begin{aligned}\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}=&\frac{0.5\Delta T_0/T_0}{0.5+\Delta R/R}\left[\frac{0.5+\Delta R/R}{0.5}+\frac{0.5+\Delta R/R}{1+\Delta R/R}\frac{L_0}{K_{\mathrm{d}}}\right]-\frac{0.5\Delta L_0/L_0}{1+\Delta R/R}\frac{L_0}{K_{\mathrm{d}}}\\&+\left[-\frac{L_0}{K_{\mathrm{d}}}+\left(\frac{1+\Delta R/R}{0.5+\Delta R/R}\right)^2\left(\frac{0.5+\Delta R/R}{0.5}+\frac{0.5+\Delta R/R}{1+\Delta R/R}\frac{L_0}{K_{\mathrm{d}}}\right)\right]\frac{0.5\Delta R/R}{1+\Delta R/R}\\\Rightarrow\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}=&\left[\frac{0.5(\Delta T_0/T_0-\Delta L_0/L_0-\Delta R/R)}{1+\Delta R/R}+\frac{0.5\Delta R/R}{0.5+\Delta R/R}\right]\frac{L_0}{K_{\mathrm{d}}}\\&+\left[\Delta T_0/T_0+\frac{\Delta R/R(1+\Delta R/R)}{0.5+\Delta R/R}\right]\end{aligned} \tag{C19}$$

By introducing the definition of systematic error of K_d (*i.e.*, $\Delta K_d = K_{d,det} - K_d$) into **Eq (C19)**, we obtain:

$$\begin{aligned}\frac{K_{d,det}-K_d}{K_d} &= \frac{K_{d,det}}{K_d} - 1 \\ &= \left[\frac{0.5\left(\Delta T_0/T_0 - \Delta L_0/L_0 - \Delta R/R\right)}{1+\Delta R/R} + \frac{0.5\,\Delta R/R}{0.5+\Delta R/R}\right]\frac{L_0}{K_d} \\ &+ \left[\Delta T_0/T_0 + \frac{\Delta R/R\left(1+\Delta R/R\right)}{0.5+\Delta R/R}\right] \\ \Rightarrow \frac{K_{d,det}}{K_d} &= \left[\frac{0.5\left(\Delta T_0/T_0 - \Delta L_0/L_0 - \Delta R/R\right)}{1+\Delta R/R} + \frac{0.5\,\Delta R/R}{0.5+\Delta R/R}\right]\frac{L_0}{K_d} \\ &+ \left[\Delta T_0/T_0 + \frac{\Delta R/R\left(1+\Delta R/R\right)}{0.5+\Delta R/R} + 1\right]\end{aligned} \tag{C20}$$

which is **Eq (4.13)** in the main text.

If the guidelines established in our previous work for minimizing systematic errors of variables in K_d determinations are strictly followed, we can reasonably assume that the confidence intervals of ΔT_0 and ΔL_0 are defined by the random errors in the concentrations of stock solutions. In contrast, at this stage, we have not identified a quantitative approach for assessing the confidence interval of ΔR. Therefore, due to the absence of input data for ΔR, we will only consider the systematic errors of T_0 and L_0, ignoring ΔR, to determine the narrower confidence intervals for K_d at the specified confidence levels.. In this case, **Eq (C20)** can be simplified to the following (by excluding $\Delta R/R$):

$$\frac{K_{d,det}}{K_d} = 0.5\left(\frac{\Delta T_0}{T_0} - \frac{\Delta L_0}{L_0}\right)\frac{L_0}{K_d} + \frac{\Delta T_0}{T_0} + 1 \tag{C21}$$

which is **Eq (4.14)** in the main text.

Note C1.3: Determination of the Choice of Ligand Concentrations Used for Fitting a Binding Isotherm

For one single binding isotherm, as mentioned in the main text, to obtain two distinguishable determined parameters $K_{d,det}$ and $K_{d,det}'$, the ligand concentrations $L_{0,fit}$ and $L_{0,fit}'$ used for curve-fitting should differ by a large degree. Here, let's assume that $L_{0,fit}' = \alpha L_{0,fit}$, where α is the ratio of $L_{0,fit}'$ to $L_{0,fit}$. Apparently, α must be much greater or much smaller than unity to obtain $L_{0,fit}'$ to be very different from $L_{0,fit}$. Our goal was to find an α value that can at least result in $K_{d,det}'$ different from $K_{d,det}$ (*i.e.*, $|K_{d,det}' - K_{d,det}|/K_{d,det}$) by 5% at $L_0/K_d = 0.1$, and $K_{d,det}' \geq 0$ at $L_0/K_d = 100$ ($K_{d,det}' < 0$ often indicates a failed fitting) since 0.1–100 is the range in which L_0/K_d lies in for most K_d-determination experiments for stable complexes.[5]

To conduct this investigation, using Data S1 (https://doi.org/10.6084/m9.figshare.25777164.v3), we generated the binding isotherms with an input value of $K_d = 1$ for $L_0/K_d = 0.1$ and 100. Then, we fitted the two binding isotherms with different α values for $L_{0,fit}' = \alpha L_{0,fit}$ to obtain $K_{d,det}'$ (note that, by using $L_{0,fit} = L_0$, we obtain $K_{d,det}$ which is always equal to the input $K_d = 1$). According to the results summarized in **Table C1.1**, we cannot choose $\alpha >> 1$ because this choice will cause $K_{d,det}' < 0$ for $L_0/K_d = 100$. In this investigation, we found that α was required to be equal or smaller than 0.01 to satisfy our requirements for $|K_{d,det}' - K_{d,det}|/K_{d,det} \geq 5\%$ at $L_0/K_d = 0.1$ and $K_{d,det}' > 0$ at $L_0/K_d = 100$. By testing different choices of α values that were equal to or smaller than 0.01 in the ACI-Kd program, we found that $\alpha = 0.005$ was the smallest value we could choose without dramatically increasing computing time. Note that, the more α is different from unity, the more iteration steps (and computing time) are required for the program to find $K_{d,det}'$ with the ligand concentration of $L_{0,fit}'$.

Table C1.1. Investigation of the difference between $K_{d,det}'$ and $K_{d,det}$ for different ratios α of $L_{0,fit}'$ to $L_{0,fit}$.

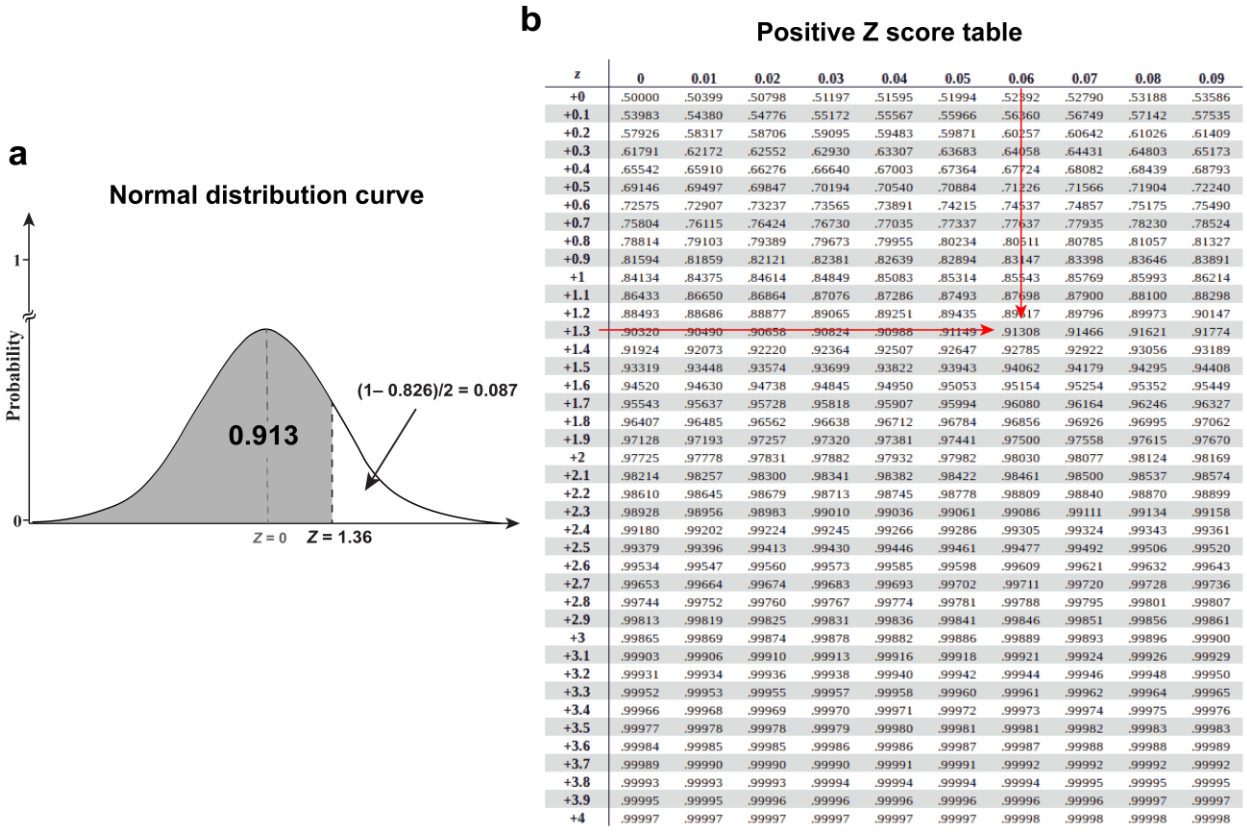
$\alpha = L_{0,fit}'/L_{0,fit}$	$L_0/K_d = 0.1$		$L_0/K_d = 100$	
	$K_{d,det}'$	$\lvert K_{d,det}' - K_{d,det}\rvert/K_{d,det}$	$K_{d,det}'$	$\lvert K_{d,det}' - K_{d,det}\rvert/K_{d,det}$
0.001	1.05	5%	40.1	3910%
0.005	1.05	5%	39.8	3880%
0.01	1.05	5%	39.5	3850%
0.1	1.04	4%	34.4	3340%
10	0.613	39%	–468	46900%
100	–2.41	341%	–9378	937900%
1000	–88.8	898%	–98340	9834100%

Note C1.4: Determination of the Required Confidence Intervals of the Relative Systematic Errors of Concentrations for Finding ACI of K_d with Different Confidence Levels

Conventionally, the K_d value determined with nonlinear regression (*i.e.*, $K_{d,det}$) is presented as "the mean value ± standard deviation (SD)" with a confidence level of 68.3%, which defines the PCI of K_d. To harmonize the confidence levels of PCI and ACI, we also demonstrate ACI with the 68.3% confidence level unless otherwise stated, and present both PCI and ACI as value ranges since the probability distribution of K_d in ACI can be non-Gaussian. As we explained in a previous work, L_0 and T_0 are subject to unavoidable systematic errors that stem from the normally-distributed random errors in concentrations of stock solutions. Thus, the lower limit of ACI of K_d is defined by the confidence intervals of L_0 and T_0. In this work, we assume that the

probability distribution for L_0 and T_0 are both normal (Gaussian) centred at their nominal values, and the confidence levels for L_0 and T_0 are identical. Therefore, to achieve the 68.3% confidence level for K_d, the confidence levels for both L_0 and T_0 should be equal to $(68.3\%)^{1/2} = 82.6\%$. To find the number of standard deviations required to cover the 82.6% probability of the concentration of ligand or target (that are centred at their nominal values), we consulted with the positive standard normal table (positive *Z*-score table).

To determine the number of standard deviations (*Z*-score) required to cover an area (probability) of 0.826 (symmetric to $Z = 0.00$), we first calculated the corresponding area (probability) to the left of the *Z* value by subtracting half of the undesired area (probability) from unity: $1 – (1 – 0.826)/2 = 0.913$ (**Figure C1.1a**). According to the standard normal table (**Figure C1.1b**), we found that probability P ($Z < 1.36$) = 0.913, which suggested that a confidence interval ranging from –1.36 × SD to 1.36 × SD is with the 82.6% confidence level (probability) because of the symmetry of the normal distribution. Therefore, we chose $\Delta L_0/L_0 \in [–1.36\delta L_0/L_0, 1.36\delta L_0/L_0]$ and $\Delta T_0/T_0 \in [–1.36\delta T_0/T_0, 1.36\delta T_0/T_0]$ as the confidence intervals of relative systematic errors of L_0 and T_0, respectively, to find ACI of K_d with the 82.6% confidence level. Using the same method, we found $\Delta L_0/L_0 \in [–2.27\delta L_0/L_0, 2.27\delta L_0/L_0]$ and $\Delta T_0/T_0 \in [–2.27\delta T_0/T_0, 2.27\delta T_0/T_0]$ as the confidence intervals for the relative systematic errors of L_0 and T_0 for the 95.5% confidence level of K_d, and $\Delta L_0/L_0 \in [–3.18\delta L_0/L_0, 3.18\delta L_0/L_0]$ and $\Delta T_0/T_0 \in [–3.18\delta T_0/T_0, 3.18\delta T_0/T_0]$ as the confidence intervals for the relative systematic errors of L_0 and T_0 for the 99.7% confidence level of K_d.



z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
+0.1	.53983	.54380	.54776	.55172	.55567	.55966	.56360	.56749	.57142	.57535
+0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
+0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
+0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
+0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
+0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
+0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
+0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
+0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
+1	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
+1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
+1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
+1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91308	.91466	.91621	.91774
+1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
+1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
+1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
+1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
+1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
+1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
+2	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
+2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
+2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
+2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
+2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
+2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
+2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
+2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
+2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
+2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
+3	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
+3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
+3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
+3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
+3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
+3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
+3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
+3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
+3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
+3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997
+4	.99997	.99997	.99997	.99997	.99997	.99997	.99998	.99998	.99998	.99998

Figure C1.1. Determination of the number of standard deviations required to cover 82.6% probability in a normal distribution: **(a)** calculate the area to the left of the Z value required to exclude half of the undesired area with the help of a normal-distribution curve, and **(b)** find the Z value corresponding to the area determined in panel **(a)** using a positive Z score table.

Note C1.5: Determination of the Conditions for the Two Boundary Cases

As discussed in Note C1.4, to achieve a confidence level of 68.3% for ACI, we should choose a confidence level of 82.63% for both $\Delta L_0/L_0$ and $\Delta T_0/T_0$, which leads to $\Delta L_0/L_0 \in [-1.36\delta L_0/L_0, 1.36\delta L_0/L_0]$ and $\Delta T_0/T_0 \in [-1.36\delta T_0/T_0, 1.36\delta T_0/T_0]$ to be the confidence intervals of relative systematic errors of L_0 and T_0. Now, we need to determine the combinations of $\Delta L_0/L_0$ and $\Delta T_0/T_0$ corresponding to $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$. In the main text, we define $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$ as the

combinations of relative systematic errors of variables that cause the smallest and largest accuracy ratio p_{det}/p, respectively. By applying this general definition to the specific case of K_d determination, $\mathbf{\Delta^*_{min}}$ and $\mathbf{\Delta^*_{max}}$ are the combinations of $\Delta L_0/L_0$ and $\Delta T_0/T_0$ (and $\Delta R/R$) that cause the smallest and largest accuracy ratio $K_{d,det}/K_d$, respectively. Based on **Eq (C21)**, we can infer that when $\Delta L_0/L_0$ is the largest value in its confidence interval (*i.e.*, $1.36\delta L_0/L_0$) and $\Delta T_0/T_0$ is the smallest value in its confidence interval (*i.e.*, $-1.36\delta T_0/T_0$), the accuracy ratio $K_{d,det}/K_d$ is minimal. Hence, we can conclude that the vector $\mathbf{\Delta^*_{min}} = \{\Delta L_0/L_0 = 1.36\delta L_0/L_0,\ \Delta T_0/T_0 = -1.36\delta T_0/T_0\}$. With a similar analysis, we can conclude that the vector $\mathbf{\Delta^*_{max}} = \{\Delta L_0/L_0 = -1.36\delta L_0/L_0,\ \Delta T_0/T_0 = 1.36\delta T_0/T_0\}$.

Note C1.6: Computational Workflow and Algorithm for the ACI-Kd Webapp

To create a user-friendly webapp for calculating the ACI of K_d, we followed a structured development process. First, we designed a detailed computational workflow (**Figure C1.2**), based on the step-by-step ACI-Kd workflow presented in **Figure 4.5** of the main text. Next, we developed a Python program implementing the computational algorithm shown in **Figure C1.3.** Finally, we deployed the Python code on a web server to build the ACI-Kd webapp, which is accessible at https://aci.sci.yorku.ca.

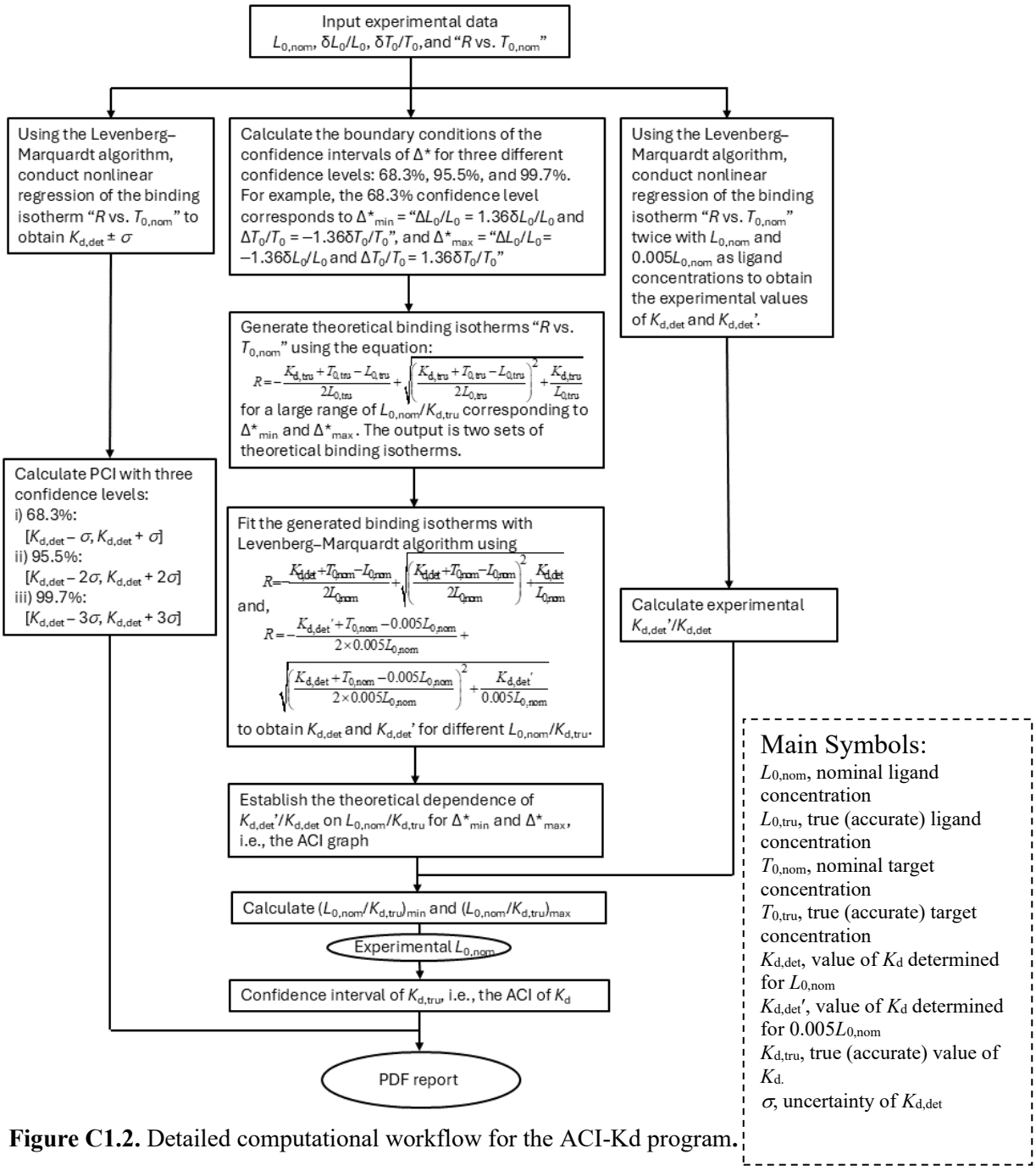


Figure C1.2. Detailed computational workflow for the ACI-Kd program.

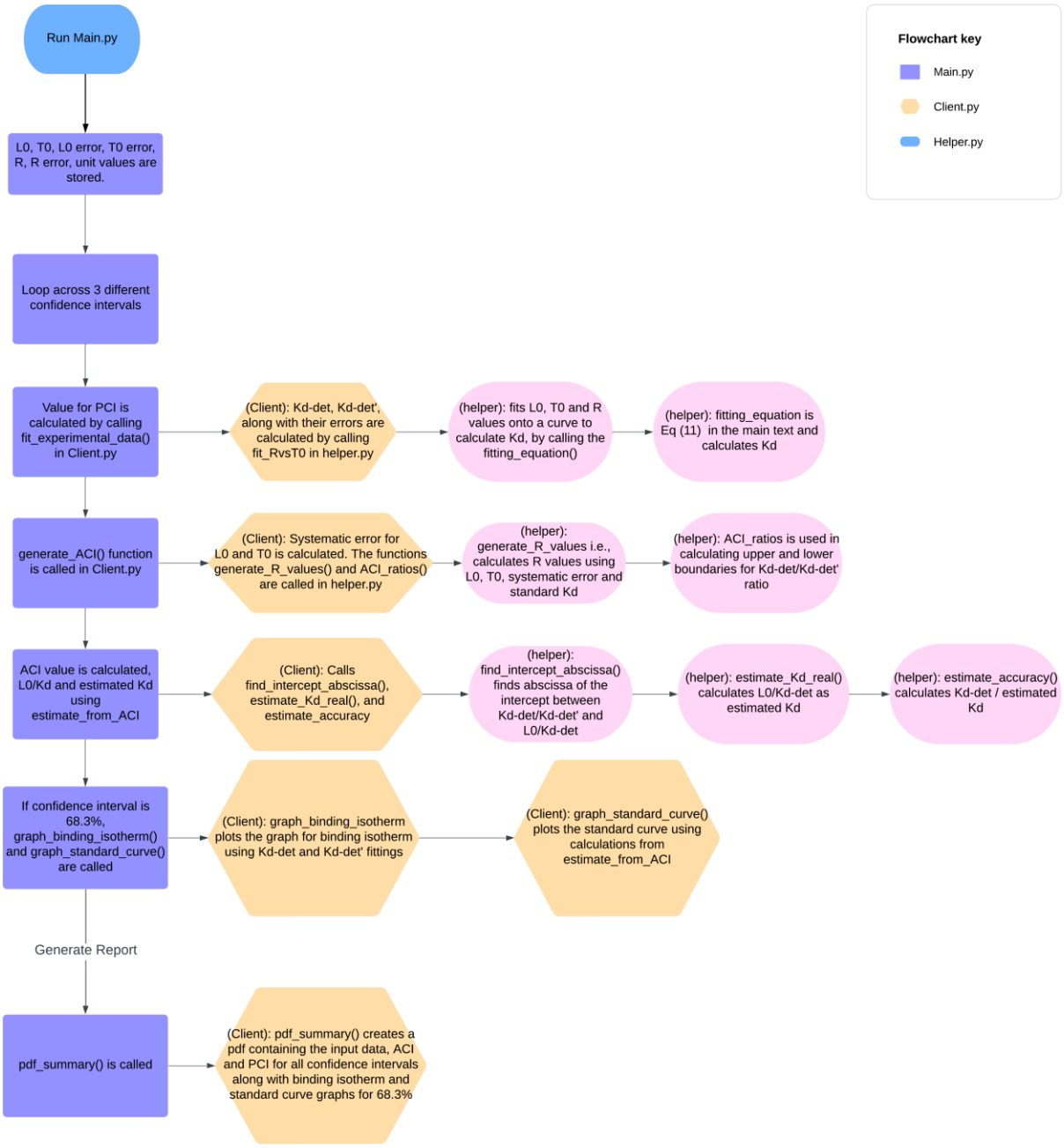


Figure C1.3. Detailed computational algorithm for the ACI-Kd Python program.

Note C1.7: Virtual ACTIS Experiments

All virtual experiments were simulated with COMSOL Multiphysics software version 5.6a. "Transport of diluted species and Laminar Flow" modules were used, and the simulation files can be found in the Supporting files (**COMSOL_AAEDC.zip**, https://doi.org/10.6084/m9.figshare.25777164.v3). In all virtual experiments, we considered a ligand L with a typical diffusion coefficient of μ_L = 100 μm^2/s and a target T with a diffusion coefficient of μ_T = 20 μm^2/s. The input rate constants were $k_{on,inp}$ = 10^5 M^{-1}s^{-1} and $k_{off,inp}$ = 5 × 10^{-5} s^{-1}, which led the true equilibrium dissociation constant $K_d = k_{off,inp} / k_{on,inp}$ = 0.5 nM. In three virtual experiments, the input (true) ligand concentrations were $L_{0,tru}$ = 2.2, 8.8 and 22 nM to achieve $\Delta L_0/L_0$ = 0.10 for nominal ligand concentration $L_{0,nom}$ = 2.0, 8.0 and 20 nM, respectively. The input (true) target concentration $T_{0,tru}$ was varied from 0.009 to 900 nM (10 different non-zero concentrations plus the zero concentration) to achieve $\Delta T_0/T_0$ = –0.10 for the nominal target concentration $T_{0,nom}$ that was varied from 0, 0.01 to 1000 nM. The other input parameters and the geometry used in the virtual experiments were identical to the settings used in one of our previous works.

To determine whether the ACI-Kd workflow (including the webapp) can correctly compute the ACI of K_d (for the virtual experiments) when given correct values of $\delta L_0/L_0$ and $\delta T_0/T_0$, we set these input values to be identical: $\delta L_0/L_0 = \delta T_0/T_0 = |\pm 0.1|/1.36 \approx 0.074$. Theoretically, if the ACI-Kd is valid for the virtual experiments, the true K_d value of 0.5 nM should be near the lower boundary of the computed ACI of K_d with a 68.3% confidence level (similar to the results shown in **Figure 4.7b1** in the main text). The results summarized in **Figure C1.4** show that the K_d = 0.5 nM are near the lower boundary of the determined ACI for all three nominal ligand concentrations ($L_{0,nom}$ = 2.0, 8.0 and 20 nM), which is consistent with the theoretical

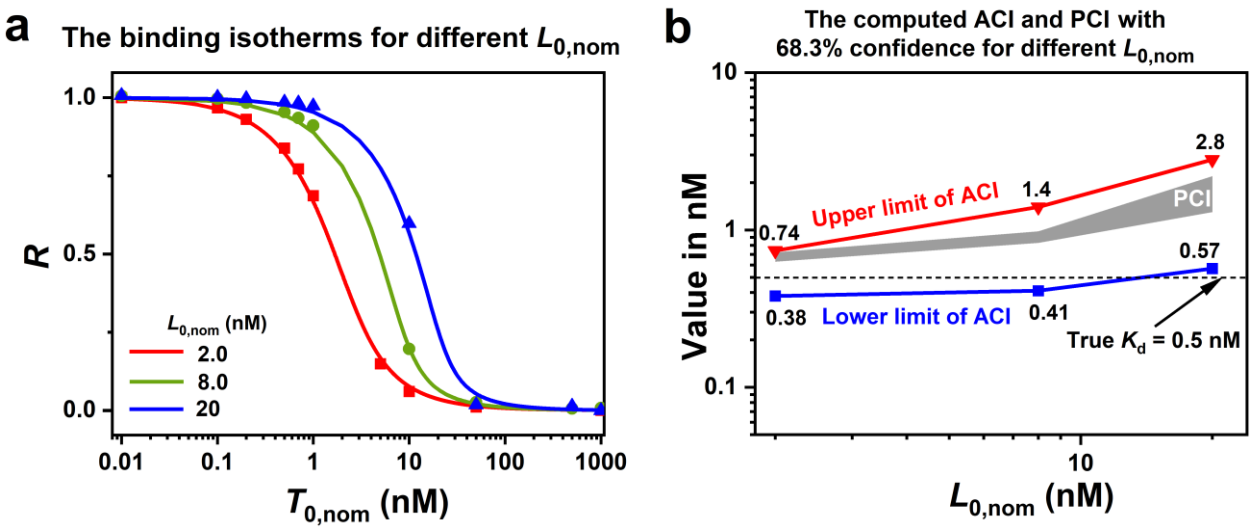


Figure C1.4. Application of ACI-Kd to results of virtual ACTIS experiments with a virtual binding pair with true K_d = 0.5 nM: **(a)** the dependence of binding isotherms to nominal ligand concentrations ($L_{0,nom}$ = 2.0, 8.0, and 20 nM); **(b)** the dependence of computed ACI and PCI on $L_{0,nom}$ with 68.3% confidence.

expectations. Note that, unlike binding isotherms generated directly from the theoretical dependence of R on concentrations (**Eq (2.2)** of the main text), the data from the virtual experiments performed in COMSOL include small, unknown errors in the signals, leading to slight systematic errors in the R values. As a result, the theoretical K_d (0.5 nM) is close to, but not exactly on, the boundary of the computed ACI (**Figure C1.7b**). This highlights that the current ACI-Kd workflow calculates a narrower limit for ACI of K_d, without accounting for systematic errors in R, due to the absence of a strategy for estimating the confidence interval of ΔR. The results in **Figure C1.4** also confirm that relying on PCI only can largely overestimate/underestimate K_d.

Note C1.8: Materials and Solutions Used in Real ACTIS Experiments for BSA–Fluorescein

All reagents were obtained from Sigma-Aldrich (Oakville, Ontario, Canada) unless otherwise stated. Fused-silica capillaries were purchased from Molex Polymicro (Phoenix, AZ, USA). The molecular pair used in the experiments was bovine serum albumin (BSA, catalog A2514) and fluorescein sodium salt (catalog F2456). A 30 mM ammonium acetate buffer pH 7.5 was utilized to prepare all solutions and used as the sample propagation buffer; accordingly, we simply refer to it as the buffer. BSA−fluorescein equilibrium mixtures (EMs) were prepared with the strategy of "new sample-preparation workflow" that was introduced in our previous work. In four ACTIS experiments, the fluorescein concentrations (L_0) were fixed to be 0.07, 7.0, 70 and 210 μM, respectively (after correction based on the purity of fluorescein). For the ACTIS experiment with L_0 = 0.07 μM, the total BSA concentrations (T_0) in the EMs varied from 0 (lowest non-zero T_0 = 1.0 μM) to 1000 μM with a total of 10 EMs with different T_0 values. For the other three ACTIS experiments with L_0 = 7.0, 70 and 210 μM, the total BSA concentrations (T_0) in the EMs varied from 0 (lowest non-zero T_0 = 0.5 μM) to 1024 μM with a total of 11 EMs in each ACTIS experiment. Prior to (incompletely) separating the EMs (and measuring the fluorescence signals) with an ACTIS instrument, all the mixtures were incubated for 2 h to reach equilibrium. All sample handling and measurements were carried out at a room temperature of 24 ± 1 °C.

Note C1.9: Details of Real ACTIS Experiments

Instrumental Setup

The ACTIS instrumental setup mimicked the previously described ACTIS instrument with the following two modifications: (*i*) all the capillaries were uncoated fused-silica capillaries

instead of PVA-coated capillaries and (*ii*) the inner diameter of the separation capillary was 200 μm.

ACTIS Experimental Procedure

The ACTIS experimental procedure was similar to the previously described one. The following changes were made. For Step 1, the valve was in position I, and the injection loop was filled with the sample using a syringe pump to aspirate the sample at a flow rate of Q_s = 50 μL/min for 25 s. Since there was a delay for the desired volume of the sample to flow through the fluidic path, the valve was set to stay in position I for 30 s. For Step 4, the separation capillary was rinsed with the buffer at a flow rate of $Q_{rinse} \approx$ 500 μL/min for 1 min allowing for approximately 20 volumes of the separation capillary to pass through the fluidic system. For the "between-sample" rinsing step, a 1.5 mL sample vial containing the buffer was placed at the end of the sample aspiration capillary, and the syringe pump was set to continuously withdraw at a flow rate of 75 μL/min, and the main pump kept injecting the buffer at a flow rate of 500 μL/min. During the "between-sample" rinsing step, the valve was switching between positions I and II 60 times. For each switch, the valve spent 2 s in each position I and II. During this rinsing, the buffer volume of approximately 80 volumes of the separation capillary passed through the fluidic system. This rinsing step took approximately 4 min. A total of 3 ACTIS runs were performed for each concentration of the protein. An ACTIS experiment with 13 protein concentrations along with all the rinsing steps (excluding the sample preparation procedure) took approximately 3 h. All experiments were conducted at a room temperature of 24 ± 1 ºC.

Determination of the Random Errors in The Concentrations of Fluorescein (Ligand) Stock Solutions

To determine the random errors in the concentration of fluorescein stock solutions, we first prepared five 50 mL fluorescein stock solutions with a nominal (expected) concentration of 1000 μM. According to the product specification of the used fluorescein sodium salt, the purity of the product was 70%. Thus, the nominal concentrations of the fluorescein stock solutions were corrected to 700 μM. Each solution was prepared separately from scratch. Then, to avoid saturation of the spectrophotometer detector by the excessively high fluorescein concentration, the fluorescein stock solutions were diluted 100 times to obtain five fluorescein solutions with a nominal concentration of 7.0 μM. For each diluted solution, the absorbance at a light wavelength of 488 nm (A_{488}) with the optical path length of l = 1 mm was measured with a NanoDrop 1000 (Thermo Scientific) spectrophotometer. The measured absorbance for each sample and the calculated relative standard deviation (RSD = $\delta L_0/L_0$) are summarized in **Table C1.2**. The absorbance of each sample was measured in triplicate.

The molar extinction coefficient (ε) of an analyte is related to the solvent used to dissolve it. Since ε of fluorescein sodium salt in the buffer of 30 mM ammonium acetate (pH 7.5) has never been accurately determined, like in most cases of K_d determinations for biomolecular complexes, the real ligand concentrations here cannot be accurately calculated with the Beer-Lambert law ($A = \varepsilon lc$, where A is absorbance and c is concentration). However, the RSD of the concentrations can be reliably determined using spectroscopic signals.

Table C1.2. Determination of the random error in the concentrations of the fluorescein stock solutions.

Sample #	A_{488}	$A_{488,ave}$	**Overall** $A_{488,ave}$	**Standard Deviation (SD)**	**RSD** ($\delta L_0/L_0$)

	0.060				
	0.062				
1	0.061	0.061			
	0.060				
	0.061				
2	0.061	0.061			
	0.058				
	0.057				
3	0.059	0.058			
	0.057				
	0.060				
4	0.058	0.058			
	0.059				
	0.061				
5	0.059	0.060	0.060	0.0014	0.023

Determination of the Random Errors in the Concentrations of BSA (Target) Stock Solutions

For determining the random errors in the concentrations of BSA stock solutions, we first prepared five 2 mL BSA stock solutions each with a nominal BSA concentration of 2048 μM (each solution was prepared separately from scratch). Using SDS-PAGE, the purity of the used BSA protein was measured to be > 98% without quantitatively measurable impurities (**Figure C1.5**), thus we assumed the BSA protein to be 100% pure and the systematic errors in BSA concentrations were solely from sample preparations. Then, the BSA stock solutions were diluted to a nominal concentration of 32 μM (five samples). We chose the concentration of 32 μM to (*i*) avoid the saturation of the spectrophotometric signal by an excessively high BSA

concentration and (*ii*) to better estimate the "average" confidence interval of $\Delta T_0/T_0$ for all the EMs since the number of consecutive two-fold dilutions we applied to achieve the concentration of 32 μM (from 2048 μM) was six, which was the average number of dilution times required to prepare the samples (EMs) used in the ACTIS experiment. Note that, the error in the concentration of a protein solution prepared with serial dilutions was related to the dilution times, and we applied 12 consecutive two-fold dilutions to obtain 0.5 μM (the lowest non-zero concentration used in the ACTIS experiment) BSA from a 2048 μM BSA solution.

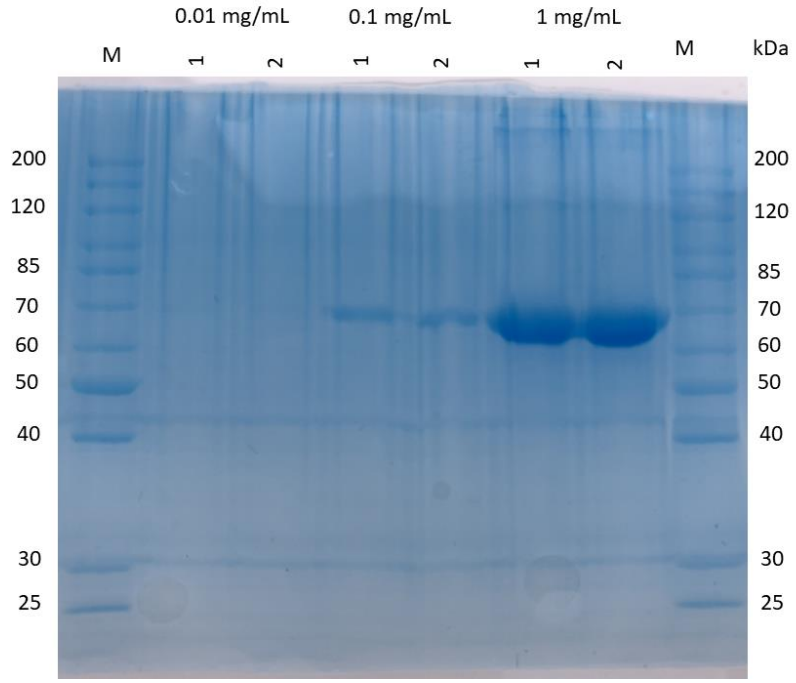


Figure C1.5. Determination of the purity of BSA protein with SDS-PAGE. For determining the purity of BSA (catalog A2514), proteins were separated by SDS-PAGE (sodium dodecyl-sulfate polyacrylamide gel electrophoresis) and visualized by Coomassie Brilliant Blue (Diagonal, Münster, Germany) staining. The BSA stock solution was diluted and applied to the gel in a concentration of 0.01 mg/mL, 0.1 mg/mL and 1 mg/mL. No other protein bands, except the expected protein band at approximately 66 kDa were visible in the gel after staining.

For each sample, the absorbance at a light wavelength of 280 nm (A_{280}) with an optical path length of l = 10 mm was measured with a NanoDrop 1000 (Thermo Scientific) spectrophotometer. The measured absorbance for each sample and the calculated relative standard deviation (RSD = $\delta T_0/T_0$) are summarized in **Table C1.3**. As we mentioned, the type of solvent can affect the extinction coefficient (ε) of an analyte, so we cannot simply calculate the real protein concentrations with the Beer-Lambert law.

Table C1.3. Determination of the random error in the concentration of BSA stock solutions.

Sample #	A_{280}	$A_{280,ave}$	**Overall** $A_{280,ave}$	**Standard Deviation (SD)**	**RSD** ($\delta T_0/T_0$)
1	1.081	1.05	1.14	0.094	0.082
	1.085				
	0.971				
2	1.191	1.21			
	1.239				
	1.213				
3	1.156	1.19			
	1.210				
	1.197				
4	1.083	1.02			
	0.928				
	1.055				
5	1.241	1.21			
	1.171				
	1.209				

Note C1.10: Plasmid Construction and Mutagenesis

The plasmid for expression and purification of human HSP90α-mNeonGreen (HSP90α UniProtKB: P07900) was constructed in a previous work. The mNeonGreen (FPbase ID: ZRKRV) gene sequence from *Branchiostoma lanceolatum* was obtained from Addgene plasmid #58179 (Addgene HQ, NY, Watertown, USA). The DNA sequence of human CDC37 (UniProtKB: Q16543) was obtained from Addgene plasmid #163838. The gene sequence for CDC37 was cloned into the expression vector pET-11d (Art. No. 69439, Novagen, Merck, Darmstadt, Germany) with an N-terminal His_6-tag followed by a sequence coding for an FXa interface, an XhoI restriction site, the sequence of the protein of interest and a KpnI restriction site via In-Fusion cloning (Clontech Laboratories, Takara, Saint- Germain-en-Laye, France). The gene sequence of mScarlet-I (FPbase ID: 6VVTK) followed by the sequence of a flexible linker (GGGGS) was fused to the 5' of the CDC37 gene sequence (corresponding to the protein sequence N-terminus) by inserting it between the His_6-Tag and the XhoI restriction site of pET-11d. mScarlet-I gene sequence originated from Addgene plasmid #98839.

Note C1.11: Protein Purification

Protein expression and purification were similar to a previous work. For expression of His_6-HSP90α-mNeonGreen and mScarlet-I-CDC37, *E. coli* Rosetta2(DE3) cells (71403 Novagen, Merck Darmstadt, Germany) containing the corresponding plasmids were cultivated in TB-medium (15 g tryptone, 30 g yeast extract, 5 mL glycerol per litre) at 37 °C, 200 rpm to an optical density of 0.6-0.8. Protein expression was induced with 0.5 mM IPTG for HSP90α-mNeonGreen and 1 mM IPTG for mScarlet-I-CDC37 and cells henceforth incubated for another 24–72 h at 18 °C, 180 rpm. For expression of His_6-mNeonGreen, corresponding plasmids were

transformed into *E. coli* BL21(DE3) cells and cultivated at 37 °C in LB-medium to an optical density of 0.5. Protein expression was induced by the addition of 1 mM IPTG. Cells were further cultivated for 4 h at 30 °C, 200 rpm. Cells were harvested by centrifugation. The cell pellet was resuspended in HEPES buffer (40 mM HEPES/KOH pH 7.5, 20 mM KCl). Lysozyme (0.2 g/L, final concentration), DNAseI (0.05 g/L), benzamidine (1 mM) and PMSF (1 mM) were added to the cell suspension and the cells were incubated for 20 min on ice. Cells were lyzed by ultrasonification and centrifuged at 100.000 × *g* for 30 min at 4 °C. The supernatant was applied to Ni-NTA columns that had previously been equilibrated with HEPES buffer containing 10 mM imidazole. The column was washed thoroughly with HEPES buffer containing 20 mM imidazole to remove unspecific protein binding. Elution of the desired protein was performed with HEPES buffer containing 500 mM imidazole. As a second purification step, this elution fraction was further purified by gel filtration with a HiLoad 26/600 Superdex 200 pg column and an Äkta Start (Cytiva Marlborough, MA, USA) using the HEPES buffer as the mobile phase. After gel filtration, elution fractions containing the desired protein size and respective fluorescent properties were united and concentrated using an Amicon® Ultra-15 filter unit (MWCO: 10 kDa) (Merck, Darmstadt, Germany). The qualities of protein purifications were assessed via SDS-PAGE analysis (see **Figure C1.6**). The purities of the protein solutions were analyzed with ImageJ. The spectroscopic signals (*i.e.*, absorbance) of protein solutions were determined with a nanophotometer (Pearl, Implen GmbH; Munich, Germany) at 280 nm. For long-term storage, the protein solution was aliquoted, flash-frozen in liquid nitrogen and stored at –80 °C. mNeonGreen was purified solely via His_6-tag purification.

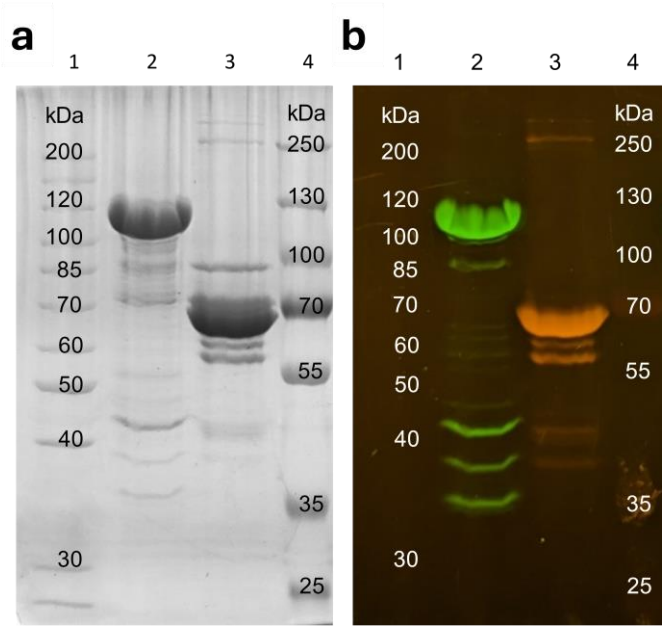


Figure C1.6. Determination of the purities of HSP90α-mNeonGreen and mScarlet-I-CDC37 with SDS-PAGE. Shown is the 10% polyacrylamide SDS-gel after SDS-PAGE and staining with Coomassie Brilliant Blue **(a)** and prior to staining under UV-light exposure **(b)** to visualize the fluorescence of the fusion proteins. The purity of the protein solutions was assessed via densitometric analysis after Coomassie staining with ImageJ. Lane 1: protein ladder PAGE Ruler Unstained. Lane 2: HSP90α-mNeonGreen (MW: 113 kDa), purity: 74.5%. Lane 3: mScarlet-I–CDC37 (MW: 72 kDa), purity: 79.5%. Lane 4: protein ladder PAGE Ruler Prestained Plus.

Note C1.12: Details of FRET Experiments

Instrumental Setup

FRET measurements were carried out with two different microplate readers, *i.e.*, a monochromator-based instrument Infinite M200Pro (Tecan, Männedorf, Switzerland) and a filter-based instrument Infinite F200Pro (Tecan, Männedorf, Switzerland).

With the instrument Tecan Infinite M200Pro, the fluorescence was measured in three different channels, as specified in **Table C1.4**.

Table C1.4. Specifications of fluorescence measurements with the Tecan Infinite M200Pro.

Channel	**Excitation wavelength (nm)**	**Emission wavelength (nm)**
mNeonGreen	488	525
mScarlet-I	561	610
FRET	488	610

Bandpass: excitation wavelength: 9 nm, emission wavelength: 20 nm

With the instrument Tecan Infinite F200Pro, the fluorescence was measured in three different channels, as specified in **Table C1.5**.

Table C1.5. Specifications of fluorescence measurements with the Tecan Infinite F200Pro.

Channel	**Excitation wavelength (nm)**	**Emission wavelength (nm)**
mNeonGreen	485 ± 20	535 ± 20
mScarlet-I	535 ± 20	612 ± 10

Channel	Excitation wavelength (nm)	Emission wavelength (nm)
FRET	485 ± 20	612 ± 10

Further specifications that were independent of the instrument used are specified in the following **Table C1.6**.

Table C1.6. Further specifications of the fluorescence measurements.

Parameter	Setting
Number of flashes	25
Gain	Optimal
Fluorescence measurement	Top
Z-Position	20,000 µm
Integration time	20 µs

FRET Experimental Procedure

FRET experiments were carried out in 384-well microtiter plates (Art. No. 781906, Greiner Bio-One GmbH, Frickenhausen, Germany) containing a final volume of 20 µL per well.

FRET emission (Em_{FRET}) was calculated from the fluorescence signal of the three channels (see instrumental setup) according to Song *et al.* 2011:

$$Em_{FRET} = (FL_{DA}) - (x \cdot FL_{DD}) - (y \cdot FL_{AA}) \quad \text{(C22)}$$

where FL_{DA} is the fluorescence signal of the FRET channel; FL_{DD} and FL_{AA} are fluorescence signals from mNeonGreen and mScarlet-I channels, respectively. The correction factors x and y were calculated as follows:

$$x = \frac{FL_{\mathrm{DA}}}{FL_{\mathrm{DD}}} \tag{C23}$$

for samples containing only the donor fluorophore or fluorescent fusion proteins with mNeonGreen, and

$$x = \frac{FL_{\mathrm{DA}}}{FL_{\mathrm{AA}}} \tag{C24}$$

for samples containing only the acceptor fluorophore or fluorescent fusion proteins with mScarlet-I in the same concentration as samples containing both donor and acceptor.

Determination of the LOQ of Two Different Microplate Readers for the Ligand of HSP90α-Mneongreen

The limit of quantitation (LOQ) is defined as the analyte concentration that can produce a signal that is X ($X >> 1$) times the standard deviation (SD) of background noise, with the choice of X varying with applications.[1] Unlike most analytical approaches that produce relatively stable background signals with noises in well-defined ranges, the "background" signals and noises of FRET experiments are dependent on the concentration of the excess component (*e.g.*, acceptor/target) due to unspecific effects, such as molecular crowding at high acceptor concentrations, *i.e.*, stochastic FRET. In a previous study, it was shown that the background signal and its noise increase with the concentration of the acceptor/target. Therefore, to determine the LOQ for the two microplate readers, *i.e.*, Tecan Infinite M200Pro and Tecan Infinite F200Pro, we measured the signals caused by binding events (*i.e.*, "FRET signal from the mixture of HSP90α-mNeonGreen and mScarlet-I-CDC37" – "FRET signal from the mixture of mNeonGreen and mScarlet-I-CDC37 (negative control background)") for different donor concentrations ($[\text{HSP90α-mNeonGreen}]_0$) at the highest acceptor concentration ([mScarlet-I-

CDC37]$_0$ = 3,500 nM after purity correction) planned to be used in the K_d-determination experiments. The background noises here were the standard deviations (SD) of the FRET signals for mNeonGreen–mScarlet-I-CDC37 in the negative control experiments. Then, we plotted the dependence of the FRET signal caused by binding events on the concentration of donor ([HSP90α-mNeonGreen]$_0$) and set X = 15 to find the LOQ of the two instruments (**Figure C1.7**). Our analysis showed that the LOQ of Tecan Infinite M200Pro and Tecan Infinite F200Pro (for HSP90α-mNeonGreen) were 630 nM (**Figure C1.7a**) and 140 nM (**Figure C1.7b**), respectively.

Determination of the K_d of Mneongreen–Mscarlet-I-CDC37 with Two Different Microplate Readers at Two Different Ligand Concentrations

For determining the equilibrium dissociation constant (K_d) of HSP90α-mNeonGreen–mScarlet-I-CDC37 with two different instruments (Tecan Infinite M200Pro and Tecan Infinite

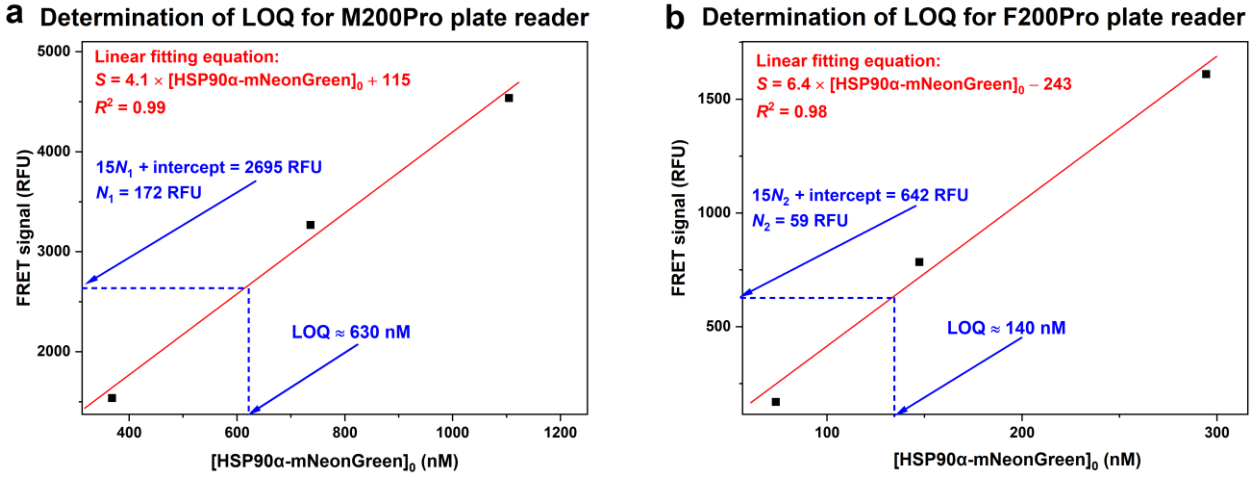


Figure C1.7. Determination of the limit of quantitation (LOQ) of **(a)** Tecan Infinite M200Pro microplate reader (Männedorf, Switzerland), and **(b)** Tecan Infinite F200Pro microplate reader (Männedorf, Switzerland). The determinations were based on the linear plotting of "FRET signal caused by binding versus [HSP90α-mNeonGreen]$_0$". The background noises N_1 and N_2 were determined to be 172 RFU and 59 RFU, respectively.

F200Pro) and two different donor concentrations, 10 µL HSP90α-mNeonGreen or mNeonGreen in the negative control (with concentrations of 1,480 and 300 nM, respectively, in the two experiments) in HEPES buffer (40 mM HEPES/KOH pH 7.5, 20 mM KCl) with 0.2% Triton X-100 were mixed with varied concentration of 10 µL of mScarlet-I-CDC37 (0–7 µM, lowest non-zero concentration 20 nM) in HEPES buffer. Resulting ultimately in constant HSP90α-mNeonGreen (or mNeonGreen) concentrations of 740 and 150 nM and mScarlet-I-CDC37 concentrations between 0 and 3,500 nM (lowest non-zero concentration 10 nM) in HEPES buffer containing 0.1% Triton X-100. A total of 15 different mScarlet-I-CDC37/HSP90α-mNeonGreen or mScarlet-I-CDC37/mNeonGreen equilibrium mixtures (EM) (plus a solution with no mScarlet-I-CDC37) with a stepwise two-fold dilution were prepared. The samples were incubated at 37 °C, 300 rpm in a microplate thermoshaker (Grant-Bio PHMP, Grant Instruments, Cambridgeshire, UK) for 120 min. Measurements for $[\text{HSP90α-mNeonGreen}]_0$ = 740 nM were conducted in triplicate with the monochromator-based platereader Tecan Infinite M200Pro. Measurements for $[\text{HSP90α-mNeonGreen}]_0$ = 150 nM were conducted in triplicate with the filter-based instrument Tecan Infinite F200Pro. The FRET signals (Em_{FRET}) of the negative control mScarlet-I-CDC37/mNeon-Green were subtracted from the Em_{FRET} of the mScarlet-I-CDC37/HSP90α-mNeonGreen interaction, to account for non-binding related FRET effects (also termed stochastic FRET). The resulting Em_{FRET} was normalized to fraction unbound (R). Fraction unbound was calculated based on the Em_{FRET} signal for the highest acceptor concentration (3,500 nM), assuming that at this concentration all of the co-chaperone was bound to HSP90α. The obtained binding isotherms are shown in **Figure 4.9a**, **4.9b** in the main text.

Determination of the Random Errors in the Concentrations of HSP90α-Mneongreen (Ligand) Stock Solutions

To determine the random errors in the stock solutions of HSP90α-mNeonGreen, we measured the absorbance at a light wavelength of 280 nm (A_{280}) with path length l = 1 mm for the HSP90α-mNeonGreen stock solutions from 10 different aliquots. The concentrations of HSP90α-mNeonGreen stock solutions were ~20 μM, which would not cause saturation of the spectrophotometer detector. The absorbance was measured with a nanophotometer (Pearl, Implen GmbH; Munich, Germany). The measured absorbance for each sample and the calculated relative standard deviation (RSD = $\delta L_0/L_0$) are summarized in **Table C1.7**.

Table C1.7. Determination of the random error in the concentration of HSP90α-mNeonGreen stock solutions.

Sample #	A_{280}	**Overall A_{280}**	**Standard Deviation (SD)**	**RSD ($\delta L_0/L_0$)**
1	0.172			
2	0.181			
3	0.180			
4	0.175			
5	0.178			
6	0.175			
7	0.175			
8	0.174			
9	0.170			
10	0.168	0.174	0.004	0.024

Determination of the Random Errors in the Concentrations of Mscarlet-I-CDC37 (Target) Stock Solutions

To determine the random errors in the concentrations of mScarlet-I-CDC37 stock solutions, we measured the absorbance at a light wavelength of 280 nm (A_{280}) with an optical path length of l = 1 mm for the mScarlet-I-CDC37 stock solutions from 10 different aliquots. The absorbance was measured with a nanophotometer (Pearl, Implen GmbH; Munich, Germany). The measured absorbance for each sample and the calculated relative standard deviation (RSD = $\delta T_0/T_0$) are summarized in **Table C1.8**.

Table C1.8. Determination of the random error for the mScarlet-I-CDC37 concentration.

Sample #	A_{280}	**Overall A_{280}**	**Standard Deviation (SD)**	**RSD ($\delta T_0/T_0$)**
1	0.950			
2	0.990			
3	1.040			
4	1.022			
5	1.073			
6	0.951			
7	1.038			
8	0.971			
9	1.090			
10	1.096	1.022	0.055	0.054

Note C1.13: Computational Workflow for Bootstrapping Program

To compare the confidence intervals produced by ACI-Kd with those obtained through bootstrapping, we developed a Python script (uploaded to https://doi.org/10.6084/m9.figshare.25777164.v3) following the computational workflow shown in **Figure C1.8.** Comparing **Figure C1.8** with **Figure C1.2** highlights that ACI-Kd is a distinct approach from Monte Carlo-based methods, such as bootstrapping.

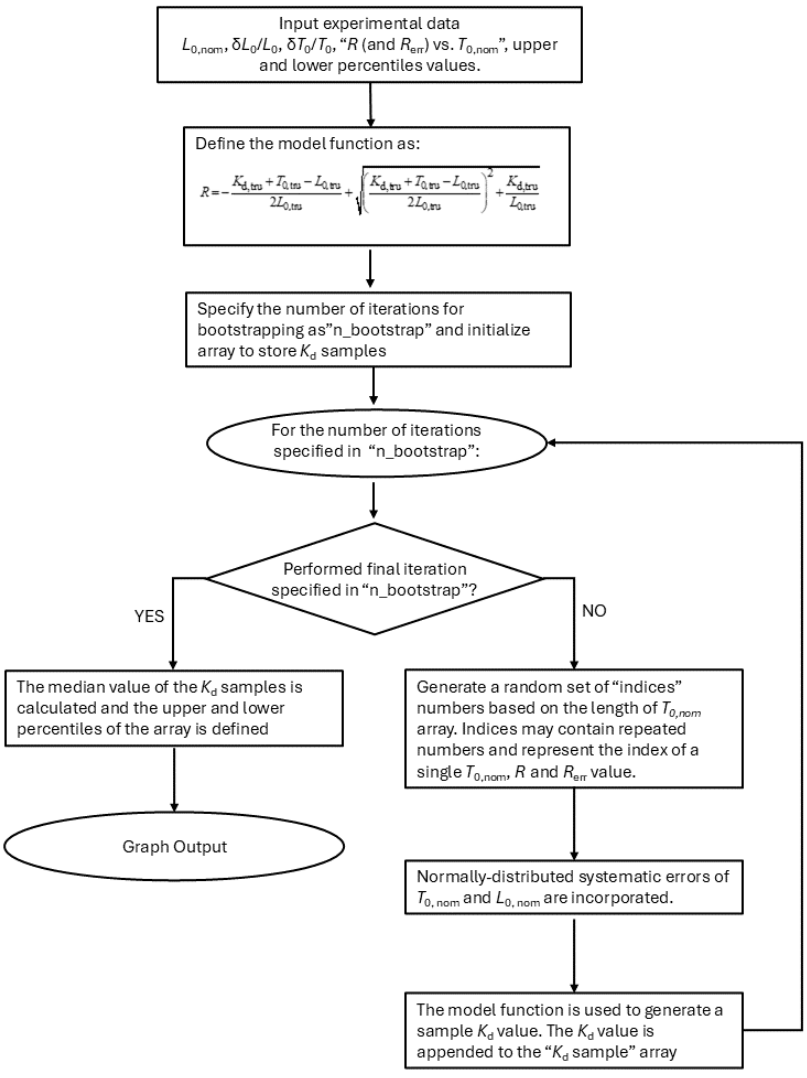


Figure C1.8. Detailed computational workflow for the bootstrapping program.

Appendix C2. Supporting Information for "Introducing Quantitative Assessment of Michaelis Constant (K_m) Accuracy"

Note C2.1: Dependencies of the Constants on the Relative Systematic Errors of Variables

The detailed dependencies of constants a and b as well as α, λ, and β on $|\Delta T_0/T_0|$, $|\Delta L_0/L_0|$, and $|\Delta R/R|$ were presented in our previous work on K_d determination. In this study, due to the mathematical equivalence between K_d and K_m determinations, we can extend these results by replacing $|\Delta T_0/T_0|$ and $|\Delta L_0/L_0|$ to $|\Delta S_0/S_0|$ and $|\Delta E_0/E_0|$, respectively. This substitution allows us to derive the dependencies of a, b, α, λ, and β on $|\Delta S_0/S_0|$, $|\Delta E_0/E_0|$, and $|\Delta R/R|$.

By defining the relative systematic errors in S_0, E_0, and R as $|\Delta S_0/S_0| = s$, $|\Delta E_0/E_0| = e$, and $|\Delta R/R| = r$, he following dependencies can be obtained:

$$
\begin{aligned}
&a = s + \frac{r(1+r)}{0.5+r},\ b = \frac{0.5(s+e-r)}{1+r} + \frac{0.5r}{0.5+r} \ \text{when } 0 \le \Delta R < 0.5 \\
&a = s + \frac{r(1-r)}{0.5-r},\ b = \frac{0.5(s+e-r)}{1-r} + \frac{0.5r}{0.5-r} \ \text{when } -0.5 < \Delta R < 0 \\
&\alpha = \sqrt{s^2 + \left(\frac{r(1+r)}{0.5}\right)^2},\ \beta = \frac{0.5}{1+r}\sqrt{s^2+e^2+r^2},\ \lambda = \frac{s^2}{1+r} + 2r^2 \ \text{when } 0 \le \Delta R < 0.5 \\
&\alpha = \sqrt{s^2 + \left(\frac{r(1-r)}{0.5}\right)^2},\ \beta = \frac{0.5}{1-r}\sqrt{s^2+e^2+r^2},\ \lambda = \frac{s^2}{1-r} + 2r^2 \ \text{when } -0.5 < \Delta R < 0
\end{aligned}
\tag{C25}
$$

Note C2.2: Determination of the Relative Systematic Errors in Concentrations that Define the ACI Of K_m with Different Confidence Levels

Eq (4.16) in the main text shows that:

$$K_m = \frac{S_0 - E_0(1-R)}{(1/R-1)} \tag{C26}$$

When the error sources of S_0, E_0, and R are strongly correlated, the maximum systematic error of K_m can be calculated with the error propagation rule:

$$|\Delta K_m| = \left|\left(\frac{\partial K_m}{\partial S_0}\right)\Delta S_0\right| + \left|\left(\frac{\partial K_m}{\partial E_0}\right)\Delta E_0\right| + \left|\left(\frac{\partial K_m}{\partial R}\right)\Delta R\right| \quad \text{(C27)}$$

However, when calculating the systematic error in K_m propagated from the specific systematic errors of variables — which may have different signs and result in ΔK_m with varying signs — it is important to use the formula without applying absolute value:

$$\Delta K_m = \left(\frac{\partial K_m}{\partial S_0}\right)\Delta S_0 + \left(\frac{\partial K_m}{\partial E_0}\right)\Delta E_0 + \left(\frac{\partial K_m}{\partial R}\right)\Delta R \quad \text{(C28)}$$

Note that, for uncorrelated error sources of S_0, E_0, and R, we can also use **Eq (C28)** to find the relative systematic errors in variables that define the ACI of K_m at various confidence levels.

Applying the error propagation rule **Eq (C28)** to **Eq (C26)** yields the following result:

$$\Delta K_m = \frac{R}{1-R}\Delta S_0 - R\Delta E_0 + \left(-E_0 + \frac{S_0}{(1-R)^2}\right)\Delta R \quad \text{(C29)}$$

By solving **Eq (C26)** for S_0, we obtain:

$$\boldsymbol{S_0 = (1/R - 1)K_m + (1 - R)E_0} \quad \text{(C30)}$$

To derive an analytical solution for error propagation in **Eq (C26)**, it is essential to focus on a single data point by fixing R. In this case, we focus on the scenario where the determined value of R is 0.5 with a systematic error ΔR, which approximates the least erroneous case. The actual value of R is then given by:

$$R = 0.5 - \Delta R \quad \text{(C31)}$$

and accordingly, **Eq (C30)** is converted to:

$$S_0 = \frac{(0.5+\Delta R)K_{\mathrm{m}}}{0.5-\Delta R} + (0.5+\Delta R)E_0 \tag{C32}$$

Dividing both sides of **Eq (C29)** by K_{m} yields:

$$\frac{\Delta K_{\mathrm{m}}}{K_{\mathrm{m}}} = \frac{R}{1-R}\frac{\Delta S_0}{K_{\mathrm{m}}} - R\frac{\Delta E_0}{K_{\mathrm{m}}} + \left(-\frac{E_0}{K_{\mathrm{m}}} + \frac{1}{(1-R)^2}\frac{S_0}{K_{\mathrm{m}}}\right)\Delta R \tag{C33}$$

which can be rewritten as:

$$\frac{\Delta K_{\mathrm{m}}}{K_{\mathrm{m}}} = \frac{R}{1-R}\left(\frac{\Delta S_0}{S_0}\right)\frac{S_0}{K_{\mathrm{m}}} - R\frac{\Delta E_0}{E_0}\frac{E_0}{K_{\mathrm{m}}} + \left(-\frac{E_0}{K_{\mathrm{m}}} + \frac{1}{(1-R)^2}\frac{S_0}{K_{\mathrm{m}}}\right)\Delta R \tag{C34}$$

By inserting **Eqs (C31)** and **(C32)** into **Eq (C34)**, we obtain:

$$\begin{aligned}\frac{\Delta K_{\mathrm{m}}}{K_{\mathrm{m}}} = {} & \frac{0.5-\Delta R}{0.5+\Delta R}\left(\frac{\Delta S_0}{S_0}\right)\left[\frac{0.5+\Delta R}{0.5-\Delta R} + \frac{E_0}{K_{\mathrm{m}}}(0.5+\Delta R)\right] - (0.5-\Delta R)\frac{\Delta E_0}{E_0}\frac{E_0}{K_{\mathrm{m}}} \\ & + \left[-\frac{E_0}{K_{\mathrm{m}}} + \frac{1}{(0.5+R)^2}\left(\frac{0.5+\Delta R}{0.5-\Delta R} + \frac{E_0}{K_{\mathrm{m}}}(0.5+\Delta R)\right)\right]\Delta R\end{aligned} \tag{C35}$$

To express ΔR through the relative systematic error of R (*i.e.*, $\Delta R/R$), we have:

$$\begin{aligned}&\Delta R = R\frac{\Delta R}{R} = (0.5-\Delta R)\frac{\Delta R}{R} \Rightarrow \Delta R = 0.5\frac{\Delta R}{R} - \Delta R\frac{\Delta R}{R} \Rightarrow \left(1+\frac{\Delta R}{R}\right)\Delta R = 0.5\frac{\Delta R}{R} \\ &\Rightarrow \Delta R = \frac{0.5\dfrac{\Delta R}{R}}{1+\dfrac{\Delta R}{R}}\end{aligned} \tag{C36}$$

Finally, substituting **Eq (C36)** into **Eq (C35)** gives:

$$\begin{aligned}\frac{\Delta K_{\mathrm{m}}}{K_{\mathrm{m}}} = {} & \frac{0.5\Delta S_0/S_0}{0.5+\Delta R/R}\left[\frac{0.5+\Delta R/R}{0.5} + \frac{0.5+\Delta R/R}{1+\Delta R/R}\frac{E_0}{K_{\mathrm{M}}}\right] - \frac{0.5\Delta E_0/E_0}{1+\Delta R/R}\frac{E_0}{K_{\mathrm{M}}} \\ & + \left[-\frac{E_0}{K_{\mathrm{M}}} + \left(\frac{1+\Delta R/R}{0.5+\Delta R/R}\right)^2\left(\frac{0.5+\Delta R/R}{0.5} + \frac{0.5+\Delta R/R}{1+\Delta R/R}\frac{E_0}{K_{\mathrm{M}}}\right)\right]\frac{0.5\Delta R/R}{1+\Delta R/R} \\ \Rightarrow \frac{\Delta K_{\mathrm{m}}}{K_{\mathrm{m}}} = {} & \left[\frac{0.5(\Delta S_0/S_0 - \Delta E_0/E_0 - \Delta R/R)}{1+\Delta R/R} + \frac{0.5\Delta R/R}{0.5+\Delta R/R}\right]\frac{E_0}{K_{\mathrm{M}}} \\ & + \left[\Delta S_0/S_0 + \frac{\Delta R/R(1+\Delta R/R)}{0.5+\Delta R/R}\right]\end{aligned} \tag{C37}$$

If we define $K_{m,det}$ and K_m as the experimentally determined and true (unknown) K_m values, respectively, by incorporating the definition of the systematic error of K_m (*i.e.*, $\Delta K_m = K_{m,det} - K_m$) into **Eq (C37)**, we obtain:

$$\begin{aligned}\frac{K_{m,det}-K_m}{K_m} &= \frac{K_{m,det}}{K_m} - 1 \\ &= \left[\frac{0.5\left(\Delta S_0/S_0 - \Delta E_0/E_0 - \Delta R/R\right)}{1+\Delta R/R} + \frac{0.5\,\Delta R/R}{0.5+\Delta R/R}\right]\frac{E_0}{K_m} \\ &\quad + \left[\Delta S_0/S_0 + \frac{\Delta R/R\left(1+\Delta R/R\right)}{0.5+\Delta R/R}\right] \\ \Rightarrow \frac{K_{m,det}}{K_m} &= \left[\frac{0.5\left(\Delta S_0/S_0 - \Delta E_0/E_0 - \Delta R/R\right)}{1+\Delta R/R} + \frac{0.5\,\Delta R/R}{0.5+\Delta R/R}\right]\frac{E_0}{K_m} \\ &\quad + \left[\Delta S_0/S_0 + \frac{\Delta R/R\left(1+\Delta R/R\right)}{0.5+\Delta R/R} + 1\right]\end{aligned} \tag{C38}$$

If the guidelines from our previous work for minimizing systematic concentration errors are strictly followed, we can reasonably assume that the confidence intervals for ΔS_0 and ΔE_0 are governed by normally-distributed random errors in the stock solution concentrations, denoted as δS_0 and δE_0. In contrast, at this point, we have not established a quantitative method for assessing the confidence interval of ΔR. Therefore, in the absence of input data for ΔR, we will focus solely on the systematic errors of S_0 and E_0, disregarding ΔR, to derive narrower confidence intervals for K_m at the specified confidence levels. Under these conditions, **Eq (C38)** simplifies as follows (with $\Delta R/R$ excluded):

$$\frac{K_{m,det}}{K_m} = 0.5\left(\frac{\Delta S_0}{S_0} - \frac{\Delta E_0}{E_0}\right)\frac{E_0}{K_m} + \frac{\Delta S_0}{S_0} + 1 \tag{C39}$$

If we assume that the confidence levels for ΔE_0 and ΔS_0 are always identical, then to achieve a 68.3% confidence level for K_m, the confidence levels for both ΔE_0 and ΔS_0 should be equal to $(68.3\%)^{1/2}$, which is approximately 82.6%. To determine the number of standard deviations

needed to cover 82.6% of the probability distribution for the concentrations of enzyme or substrate (centred around their nominal values), we consulted the positive standard normal (Z-score) table. As a result, we selected the intervals $\Delta E_0/E_0 \in [-1.36\delta E_0/E_0, 1.36\delta E_0/E_0]$ and $\Delta S_0/S_0 \in [-1.36\delta S_0/S_0, 1.36\delta S_0/S_0]$ as the confidence intervals for the relative systematic errors of E_0 and S_0, respectively, to achieve an 82.6% confidence level for K_m.

Using the same approach, we determined that for a 95.5% confidence level of K_m, the confidence intervals for the relative systematic errors are $\Delta E_0/E_0 \in [-2.27\delta E_0/E_0, 2.27\delta E_0/E_0]$ and $\Delta S_0/S_0 \in [-2.27\delta S_0/S_0, 2.27\delta S_0/S_0]$. For a 99.7% confidence level, the intervals expand to $\Delta E_0/E_0 \in [-3.18\delta E_0/E_0, 3.18\delta E_0/E_0]$ and $\Delta S_0/S_0 \in [-3.18\delta S_0/S_0, 3.18\delta S_0/S_0]$.

Next, we need to determine the combinations of $\Delta E_0/E_0$, $\Delta S_0/S_0$ (and $\Delta R/R$) that correspond to the relative systematic errors in the variables, yielding the smallest and largest accuracy ratios, $K_{m,det}/K_m$. From **Eq (C39)**, we can infer that the accuracy ratio $K_{m,det}/K_m$ is minimized when $\Delta E_0/E_0$ is at the upper bound of its confidence interval (*i.e.*, $1.36\delta E_0/E_0$) and $\Delta S_0/S_0$ is at the lower bound (*i.e.*, $-1.36\delta S_0/S_0$). Therefore, the vectors $\{\Delta E_0/E_0 = 1.36\delta E_0/E_0, \Delta S_0/S_0 = -1.36\delta S_0/S_0\}$ and $\{\Delta E_0/E_0 = -1.36\delta E_0/E_0, \Delta S_0/S_0 = 1.36\delta S_0/S_0\}$ will define the ACI for K_m at a 68.3% confidence level through the error propagation of nonlinear regression.

A similar approach can be applied to identify the values of $\Delta E_0/E_0$ and $\Delta S_0/S_0$ that define the ACI for K_m at higher confidence levels, such as 95.5% and 99.7%.

Note C2.3: Virtual K_m-Determination Experiments and the Analysis of the Synthetic Data

In one set of K_m-determination experiments, the enzyme at a constant concentration E_0 is mixed with its substrate at varying concentrations of S_0 to allow enzymatic interactions and

product formation. The reaction progress curve, showing product concentration (P) versus reaction time, is recorded to determine the initial reaction velocity (v_0). Then, the Michaelis-Menten (MM) plot "v_0 vs S_0" is analyzed to determine $K_{m,det}$.

To model the K_m-determination process, we developed a Python program named "Progress_Curve_Simulator". The script for this program is available in the supplementary files uploaded to Figshare website (https://doi.org/10.6084/m9.figshare.26961733). This program simulates the relationship between product concentration and reaction time under specified enzyme and substrate concentrations. It utilizes user-defined kinetic rate constants k_{on}, k_{off}, and k_{cat}, respectively, in this study. The dependence of product concentration on reaction time is determined by numerically solving the following system of ordinary differential equations that describe the reaction mechanisms:

$$\begin{aligned} \frac{dS}{dt} &= -k_{on}ES + k_{off}C \\ \frac{dC}{dt} &= k_{on}ES - (k_{off} + k_{cat})C \\ \frac{dP}{dt} &= k_{cat}C \end{aligned} \tag{C40}$$

In the simulations, we set k_{on} = 100 s^{-1}, k_{off} = 1 s^{-1}, and k_{cat} = 99 s^{-1}, resulting in $K_m = (k_{cat} + k_{off})/k_{on} = 1$ (unitless). 1% random noise was added to the product concentrations in the generated data. For each set of progress-curve simulations, the nominal total substrate concentration (S_0) ranged from 1000 to 0.01 (both unitless), decreasing in 3-fold increments to generate eleven progress curves of "P vs time". Ten sets of simulations were performed with nominal total enzyme concentrations (E_0) fixed at 0.005, 0.01, 0.05, 0.1, and 0.5 (unitless). Each nominal E_0 value corresponded to two specific combinations of systematic errors in E_0 and S_0: one where $\Delta E_0/E_0 = 0.2$ and $\Delta S_0/S_0 = -0.2$, and another where $\Delta E_0/E_0 = -0.2$ and $\Delta S_0/S_0 = 0.2$.

For instance, when the nominal E_0 was set to 0.01, two sets of progress curves were simulated: in one case, the inputted true E_0 was set as 0.012 and the inputted true S_0 varied from 800 to 0.008, reflecting relative systematic errors of $\Delta E_0/E_0 = 0.2$ and $\Delta S_0/S_0 = -0.2$. In the other case, the inputted true E_0 was set as 0.008 and the inputted true S_0 varied from 1200 to 0.012, reflecting relative systematic errors of $\Delta E_0/E_0 = -0.2$ and $\Delta S_0/S_0 = 0.2$.

In the following, we demonstrate the process of simulating progress curves and performing downstream data analysis for a specific case: nominal $E_0 = 0.01$, nominal S_0 ranging from 1000 to 0.01, with $\Delta E_0/E_0 = 0.2$ and $\Delta S_0/S_0 = -0.2$. The procedures for numerical simulations and data analysis in the other nine cases followed a similar approach. As described in the previous paragraph, introducing a relative systematic error of $\Delta E_0/E_0 = 0.2$ to the nominal E_0 and $\Delta S_0/S_0 = -0.2$ to the nominal S_0 resulted in true $E_0 = 0.012$ and true S_0 ranging from 800 to 0.008 (11 different values). These values were used as inputs in the "Progress_Curve_Simulator". Representative simulated progress curves are shown in **Figure C2.1A**. The linear regions of these 11 progress curves, corresponding to the 0.06–0.14 s time range, were analyzed using linear fitting to determine the initial reaction velocity (v_0) for each nominal S_0. Representative linear analyses are presented in **Figure C2.1B**, and the resulting MM plot, "v_0 vs nominal S_0", for nominal $E_0 = 0.01$ is shown in **Figure C2.1C**. The ACI-Km program automatically and internally converts the dataset of "v_0 vs nominal S_0" into the binding isotherm "R vs nominal S_0" (**Figure C2.1D**) to calculate the PCI and ACI of K_m at various confidence levels.

In our previous work (Wang et al., 2024), we demonstrated that the true input value of K_m should lie near one boundary of the computed ACI for K_m at a 68.3% confidence level when $\delta E_0/E_0 = |\Delta E_0/E_0|/1.36$ and $\delta S_0/S_0 = |\Delta S_0/S_0|/1.36$ are used as inputs in the ACI-Km program. To verify the ACI-Km program and its associated web application (https://aci.sci.yorku.ca), we

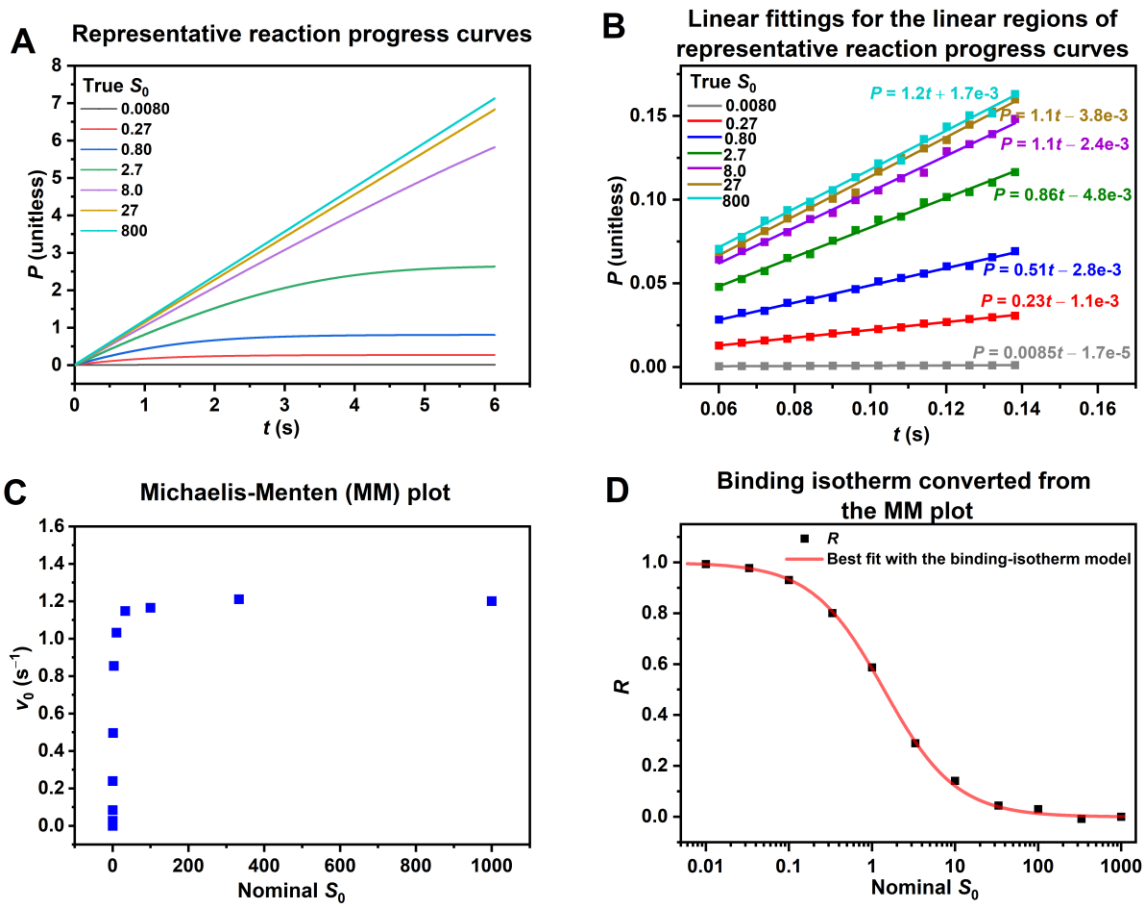


Figure C2.1. Virtual K_m-determination experiments and the downstream data analysis. **(A)** Representative simulated enzyme–substrate reaction progress curves generated using the Python program "Progress_Curve_Simulator". **(B)** Linear fits of the linear regions from the reaction progress curves shown in panel **(A)**. **(C)** MM plot ("v_0 vs nominal S_0") derived from the analysis of the simulated reaction progress curves. **(D)** Binding isotherm ("R vs nominal S_0") was converted from the MM plot in panel **(C)** with the best fitting line using the binding-isotherm model (**Eq (4.18)** in the main text). The R values were calculated using **Eq (4.15)** in the main text by setting the v_0 obtained at the highest nominal S_0 as $v_{0,max}$. In this demonstration, the nominal E_0 was 0.01, while the true E_0 was 0.012, incorporating a relative systematic error of $\Delta E_0/E_0 = 0.2$.

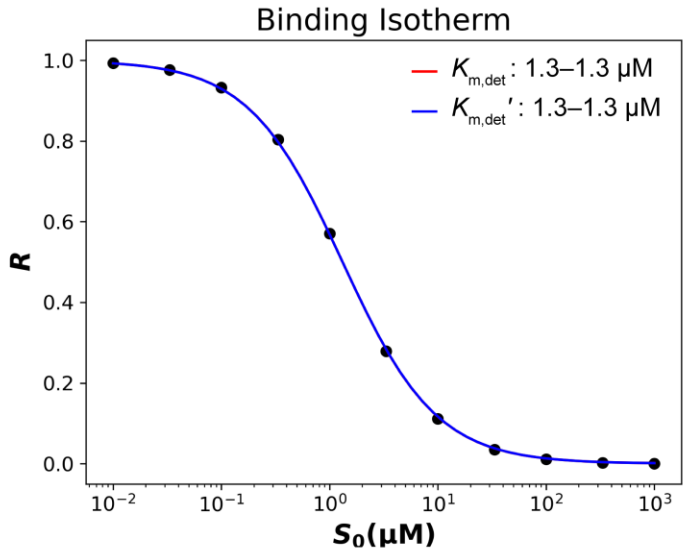
input the following parameters into the web app: nominal $E_0 = 0.01$, $\delta E_0/E_0 = 0.2/1.36 = 0.15$, $\delta S_0/S_0 = 0.2/1.36 = 0.15$, and the dataset of "$v_0$ vs nominal S_0" (generated from simulated reaction progress curves). Since the unit was set as "required" input for the web app, we input "μM" as an arbitrary unit to proceed with the calculations. The web app then calculated the PCI and ACI of K_m, producing a PDF report (**Figure C2.2**). As shown in the report, the true $K_m = 1$ (μM) is located at the lower boundary of the ACI of K_m at the 68.3% confidence level, consistent with the theoretical prediction. Verification was subsequently extended to nine additional simulated datasets with varied nominal E_0 values and different combinations of relative systematic errors in E_0 and S_0. A summary of the results from these validation processes, conducted using synthetic data, is presented in **Figure 4.13** of the main text.

ACI-Km Assessment

21/01/2025 11:31

Results: Confidence intervals of precision and accuracy of K_m with different confidence levels

	Confidence levels		
	68.3%	**95.5%**	**99.7%**
Precision confidence interval of K_m (PCI)	1.3–1.3 µM	1.3–1.3 µM	1.3–1.3 µM
Accuracy confidence interval of K_m (ACI)	1–1.6 µM	0.85–1.7 µM	0.67–1.9 µM



Input Data of MM plot

(E_0 = 0.01 µM)

S_0 (µM)	v_0 values
0.0	0.0
0.01	0.0085
0.033333	0.0279
0.1	0.0802
0.33333	0.2332
1.0	0.5089
3.33333	0.8559
10.0	1.0543
33.3333	1.1449
100.0	1.1733
333.3333	1.1836
1000.0	1.1865

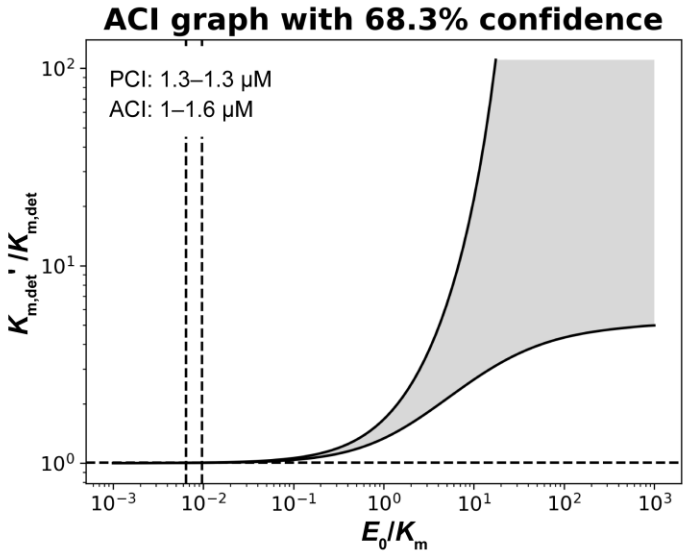


Figure C2.2. Example of a PDF report generated by the ACI-Km webapp (https://aci.sci.yorku.ca).

Note C2.4: Calculation of the ACI of K_m at Low E_0/K_m Ratio

Since the sensitivity of $K_{m,det}$ to E_0 values used in data fitting is insufficient for performing the proposed regression-stability test when $E_0/K_m < 0.001$, an alternative approach is needed to compute the ACI of K_m for very small E_0/K_m values. Based on **Eq (C25)** and **Eqs (4.20)** and **(4.22)** in the main text, we infer that for very small E_0/K_m values, the contributions of systematic errors in E_0 and E_0/K_m to the systematic error of $K_{m,det}$ are negligible.

If we assume R values are only subject to random errors with no systematic error ($r = 0$), then the systematic error in $K_{m,det}$ is solely determined by the systematic error in S_0. This implies that for extremely small E_0/K_m values, the ACI of K_m can be defined by errors in S_0 alone. Without contributions from E_0/K_m or systematic error in E_0, we propose that the ACI of K_m can be estimated as $[K_{m,det}(1 - N\delta S_0/S_0), K_{m,det}(1 + N\delta S_0/S_0)]$, where N takes values of 1.36, 2.27, and 3.18 for confidence levels of 68.3%, 95.5%, and 99.7%, respectively [2].

To validate this proposition, we investigated the ACI of K_m for E_0/K_m values ranging from 0.005 to 0.01 in increments of 0.001. Using the method described in Note S3, we simulated reaction progress curves for these E_0/K_m values while introducing two specific combinations of systematic errors: (*i*) $\Delta E_0/E_0 = 0.2$, $\Delta S_0/S_0 = -0.2$ and (*ii*) $\Delta E_0/E_0 = -0.2$, $\Delta S_0/S_0 = 0.2$. From these simulated reaction progress curves, we derived the corresponding MM plots ("v_0 vs S_0") and analyzed them using the ACI-Km webapp.

First, we confirmed that the computed ACI of K_m using $\delta E_0/E_0 = \delta S_0/S_0 = 0.15$ was identical to that obtained when only $\delta S_0/S_0 = 0.15$ ($\delta E_0/E_0 = 0$). This confirmed that the contribution of $\delta E_0/E_0$ to the ACI of K_m at very low E_0/K_m is negligible. Next, we compared the computed ACI of K_m with the proposed formula, $[K_{m,det}(1 - N\delta S_0/S_0), K_{m,det}(1 + N\delta S_0/S_0)]$ for the confidence levels mentioned above (**Figure C2.3**). The results in **Figure C2.3** demonstrate strong

consistency with our hypothesis, indicating that the ACI of K_m for low E_0/K_m (< 0.01) can be reliably computed using this alternative approach.

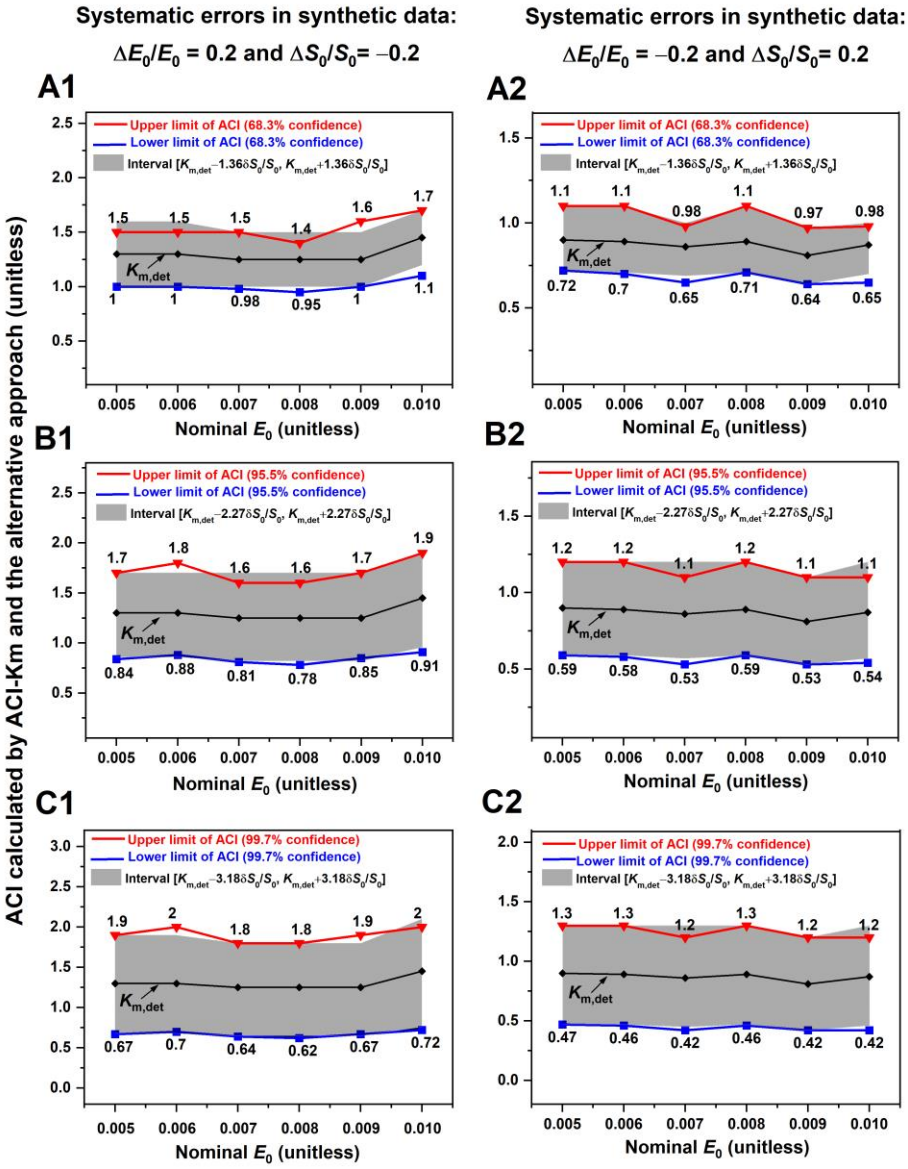


Figure C2.3. Comparison of ACI calculated using ACI-Km and the proposed alternative approach. The ratio of nominal E_0 to the true K_m values was varied from 0.005 to 0.01 in increments of 0.001. Two different combinations of systematic errors were introduced into the synthetic data: **(A1**, **B1**, **C1)** $\Delta E_0/E_0 = 0.2$ and $\Delta S_0/S_0 = -0.2$, and **(A2**, **B2**, **C2)** $\Delta E_0/E_0 = -0.2$ and $\Delta S_0/S_0 = 0.2$. Results were analyzed for three confidence levels: 68.3% **(A1**, **A2)**, 95.5% **(B1**, **B2)**, and 99.7% **(C1**, **C2)**.

Note C2.5: Data Analysis for the K_m-Determination Experiment

Here, we demonstrate the data analysis process for the K_m-determination experiment using a nominal enzyme concentration of $[LDH]_0 = 0.87$ nM. In this experiment, the enzyme concentration was fixed at 0.87 nM, while the pyruvate concentration was varied from the lowest

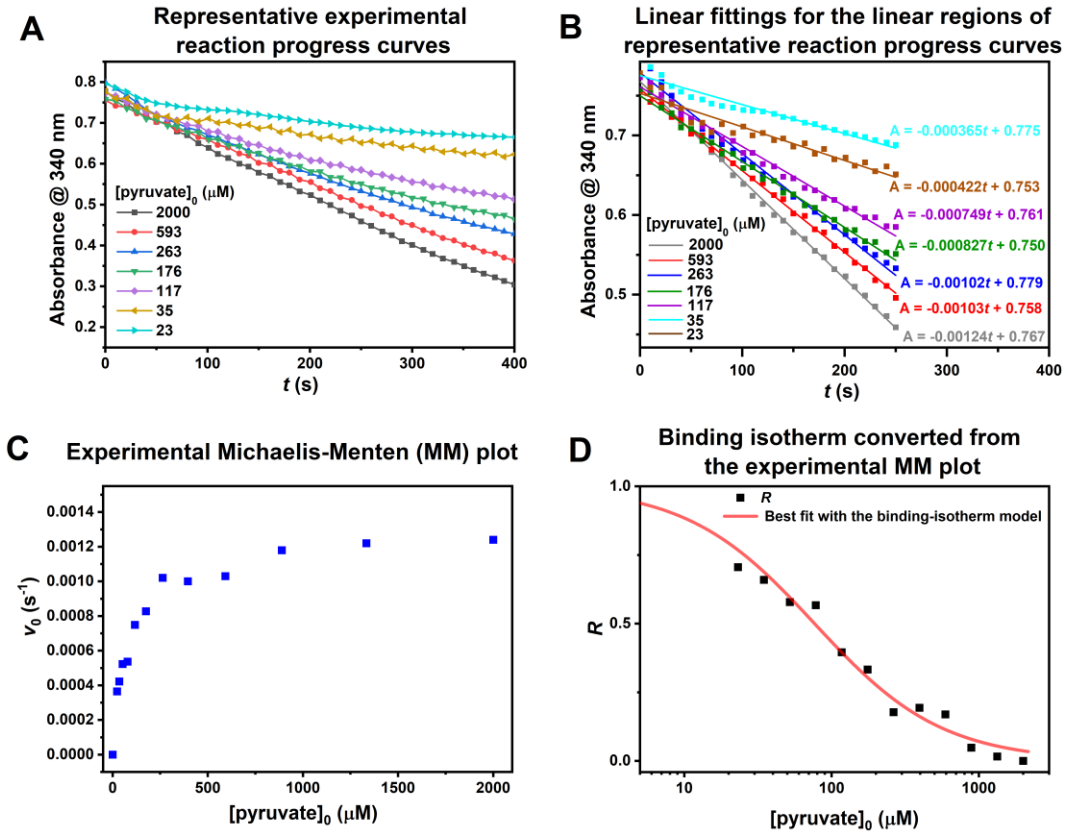


Figure C2.4. Illustration of the data-analysis process for a K_m-determination experiment in the LDH–pyruvate system. **(A)** Representative experimental reaction progress curves. **(B)** Linear fits of the linear regions (for time period 0–250 s) from the reaction progress curves shown in panel **(A)**. **(C)** Data points of the MM plot ("v_0 vs $[pyruvate]_0$") derived from the analysis of the reaction progress curves. **(D)** Binding isotherm "R vs $[pyruvate]_0$" generated from the MM plot in panel **(C)**, along with the best-fit line obtained using the binding isotherm model (**Eq (4.18)** in the main text). The R values were calculated using **Eq (4.15)** in the main text by setting the v_0 determined at the highest $[pyruvate]_0$ as $v_{0,max}$.

nonzero concentration of 23 μM to 2000 μM in 1.5-fold increments, resulting in 12 experimental reaction progress curves. Representative reaction progress curves are shown in **Figure C2.4A.** The linear regions of these 12 experimental progress curves, corresponding to the time range of 0–250 s, were analyzed using linear fitting to determine the initial reaction velocity (v_0) for each nominal [pyruvate]$_0$. Representative linear fits are presented in **Figure C2.4B**, and the resulting data points of "v_0 vs [pyruvate]$_0$" are shown in **Figure C2.4C**, including one data point with $v_0 = 0$ s^{-1} at [pyruvate]$_0 = 0$ nM.

The ACI-Km program automatically transforms the dataset of "v_0 vs [pyruvate]$_0$" into the binding isotherm "R vs [pyruvate]$_0$". **Figure C2.4D** illustrates the binding isotherm derived from the experimental "v_0 vs [pyruvate]$_0$" data points, along with the fitting line generated using the binding isotherm model (**Eq (4.18)** in the main text). In this study, reaction progress curve measurements were performed in triplicate under identical conditions. Based on the linear fitting results of the triplicate progress curves, the average v_0 (along with its associated uncertainty) for each [pyruvate]$_0$ was calculated to construct the experimental MM plot, "average v_0 vs [pyruvate]$_0$" (**Figure 4.15a** in the main text). This MM plot was then automatically converted to a corresponding binding isotherm "average R vs [pyruvate]$_0$" (**Figure 4.15b** in the main text) by the ACI-Km program (https://aci.sci.yorku.ca) to compute the PCI and ACI of K_m for various confidence levels. The raw experimental data is provided in the supplementary file "Experimental_data.xlsx" (https://doi.org/10.6084/m9.figshare.26961733).

Appendix D. Supplementary Information for Chapter 5

Appendix D1. Supporting Information for "Analytical Solution for Error Propagation in ITC: Revealing Multi-Fold Errors in K_d Values Under Standard Conditions"

Note D1.1: Calculation of Fraction of Unbound Ligand (R) Based on the Heats Measured in ITC Experiments

For the classic non-titration equilibrium K_d-determination methods, we have shown that the fraction of unbound ligand R can be determined through:

$$R = \frac{S - S_C^*}{S_L^* - S_C^*} \tag{D1}$$

Here S is the sum signal from the equilibrium mixture of the target and ligand. S_L^* and S_C^* are the signals from pure ligand and pure complex which are obtained for zero and saturating concentrations of the target, respectively. The saturating concentration satisfies the conditions of $T_0 >> L_0, K_d$.

In an ITC titration experiment, while the total concentration of ligand decreases with each injection of the target, the total amount of ligand (N_{total}, in moles) remains constant throughout the experiment. Therefore, for a 1:1 binding process:

$$N_{total} = N_L + N_C = const \tag{D2}$$

where N_L represents the amount of free ligand (in moles), and N_C denotes the amount of formed complexes (in moles). Consequently, the fraction of unbound ligand, R, is given by:

$$R = \frac{N_L}{N_{total}} \tag{D3}$$

In an ITC measurement, the total heat absorbed or released after the i-th injection (of the target)

is directly proportional to the amount of complex (N_C) formed, as expressed by:

$$Q_i = Q_{C,unit} N_C \tag{D4}$$

Here, $Q_{C,unit}$ represents the heat absorbed or released per unit amount (1 mole) of complex formed, where $Q_{C,unit} = \Delta H°$ (typically expressed in kcal/mol). Since no heat change occurs for free ligand in the absence of a binding event, we can rewrite **Eq (D4)** as:

$$Q_i = 0 \cdot N_L + Q_{C,unit} N_C = Q_{L,unit} N_L + Q_{C,unit} N_C \tag{D5}$$

where $Q_{L,unit}$ is the heat absorbed or released by unit amount (1 mole) of free ligand, which is equal to zero. Consequently, **Eq (D5)** can be reformulated as:

$$\begin{aligned} Q_i &= \left(Q_{L,unit} N_{total}\right)\frac{N_L}{N_{total}} + \left(Q_{C,unit} N_{total}\right)\frac{N_C}{N_{total}} \\ &= \left(Q_{L,unit} N_{total}\right)\frac{N_L}{N_{total}} + \left(Q_{C,unit} N_{total}\right)\frac{N_{total} - N_L}{N_{total}} \\ &= \left(Q_{L,unit} N_{total}\right)\frac{N_L}{N_{total}} + \left(Q_{C,unit} N_{total}\right)\left(1 - \frac{N_L}{N_{total}}\right) \\ &= \left(Q_{L,unit} N_{total}\right)R + \left(Q_{C,unit} N_{total}\right)(1 - R) \end{aligned} \tag{D6}$$

If the ITC measurements reach saturation — where no unbound ligand remains — at the highest target concentration used in the experiment ($T_0 >> L_0$, K_d), the total heat change corresponding to the formation of N_{total} of complex becomes measurable. Defining the measured total heat change at saturation as Q_C* (= $Q_{C,unit} N_{total}$), and the measured total heat change for free ligand as Q_L* (= $Q_{L,unit} N_{total} = 0$), we have:

$$Q_i = Q_L * R + Q_C * (1 - R) \tag{D7}$$

which suggests the total heat change measured in ITC is an additive function. As a result, R values can be calculated using:

$$R = \frac{Q_i - Q_C *}{Q_L * - Q_C *} = \frac{Q_i - Q_C *}{0 - Q_C *} = \frac{Q_C * - Q_i}{Q_C *} \tag{D8}$$

Note D1.2: Derivation of the Dependence of the Accuracy of ITC-Derived K_d ($|\Delta K_d/K_d|$) on c-Value for Strongly Correlated Error Sources

The equilibrium dissociation constant (K_d) and the mass balance equations for the target and ligand are defined as follows:

$$K_d = \frac{TL}{C} \tag{D9}$$

$$T_0 = T + C \Rightarrow T = T_0 - C \tag{D10}$$

$$L_0 = L + C \Rightarrow C = L_0 - L \tag{D11}$$

In ITC experiments, T, L, and C represent the equilibrium concentrations of the target, ligand, and complex, respectively, in the ITC sample cell after the i-th injection. T_0 and L_0 denote the total concentrations of the target and ligand, respectively, in the sample cell (after the i-th injection). Here, we refer to the injected reactant (with varied concentrations) in the ITC experiment as the target, and the reactant in the cell (that remains approximately constant concentration) as the ligand for consistency with the terminology used in our previous works

In the analysis of ITC data, the fitting equation can be generalized as follows:

$$\begin{aligned} Q_i &= nV_0\Delta H^\circ C \\ Q_i &= V_0\Delta H^\circ C \text{ for } n = 1 \end{aligned} \tag{D12}$$

where Q_i is the total heat released or absorbed after the i-th injection, n represents the stoichiometry number, V_0 is the volume of the sample cell, and ΔH° is the molar enthalpy change (heat released or absorbed per mole of formation of complex). In this study, we only focus on a 1:1 binding model with $n = 1$.

Recall that the definition of the fraction of bound ligand F ($0 \le F \le 1$) is:

$$F = \frac{C}{L_0}$$
$$\Rightarrow C = FL_0 \tag{D13}$$

By substituting **Eq (D13)** into **Eq (D12)**, we obtain:

$$Q_i = V_0 \Delta H^\circ F L_0 \tag{D14}$$

To simplify this study, we assume that V_0 is accurate (without systematic error), and ΔH° is directly measurable and with a systematic error of $\Delta(\Delta H^\circ)$. Solving F from **Eq (D14)**, we have:

$$F = \frac{Q_i}{V_0 \Delta H^\circ L_0} \tag{D15}$$

If the error sources of ΔH°, L_0, and Q_i are strongly correlated, then the error propagation for **Eq (D15)** can be calculated through:

$$\begin{aligned} |\Delta F| &= \left|\frac{\partial F}{\partial Q_i}\Delta Q_i\right| + \left|\frac{\partial F}{\partial(\Delta H^\circ)}\Delta(\Delta H^\circ)\right| + \left|\frac{\partial F}{\partial L_0}\Delta L_0\right| \\ \Rightarrow |\Delta F| &= \frac{1}{V_0}\left(\left|\frac{1}{\Delta H^\circ L_0}\Delta Q_i\right| + \left|\frac{Q_i}{\Delta H^{\circ 2} L_0}\Delta(\Delta H^\circ)\right| + \left|\frac{Q_i}{\Delta H^\circ {L_0}^2}\Delta L_0\right|\right) \end{aligned} \tag{D16}$$

To calculate the relative systematic error of F, we divide both sides of **Eq (D16)** by **Eq (D15)**:

$$\left|\frac{\Delta F}{F}\right| = \left|\frac{\Delta Q_i}{Q_i}\right| + \left|\frac{\Delta(\Delta H^\circ)}{\Delta H^\circ}\right| + \left|\frac{\Delta L_0}{L_0}\right| \tag{D17}$$

If we define $|\Delta F/F| = f$, $|\Delta Q_i/Q_i| = q$, $|\Delta(\Delta H^\circ)/\Delta H^\circ| = h$, and $|\Delta L_0/L_0| = l$, then, **Eq (D17)** can be rewritten as:

$$f = q + h + l \tag{D18}$$

In our previous works, we denote the fraction of unbound ligand as R. The relation between R and F is:

$$R = 1 - F \tag{D19}$$

By applying the error propagation rule (for correlated error sources) to **Eq (D19)**, we have:

$$|\Delta R| = |\Delta F| \tag{D20}$$

By dividing both sides in **Eq (D20)** by R, we obtain:

$$\left|\frac{\Delta R}{R}\right| = \left|\frac{\Delta F}{R}\right| = \left|\frac{\Delta F}{1-F}\right| \tag{D21}$$

To streamline the analysis, we focus on the scenario where the determined fraction of unbound ligand (R) is 0.5. This corresponds to the determined fraction of bound ligand (F) also being 0.5, which represents a case with minimal error.

Then, the actual F is:

$$F = 0.5 - \Delta F,\ -0.5 < \Delta F < 0.5 \tag{D22}$$

According to $|\Delta F/F| = f$ and $F = 0.5 - \Delta F$ (**Eq (D22)**):

When $-0.5 < \Delta F \leq 0$ (*i.e.*, $0 \leq \Delta R < 0.5$),

$$\begin{aligned}
&\Delta F = -Ff = (\Delta F - 0.5)f \Rightarrow \Delta F = f\Delta F - 0.5f \Rightarrow (f-1)\Delta F = 0.5f \\
&\Rightarrow \Delta F = \frac{0.5f}{f-1} \\
&\Rightarrow 1 - F = 1 - (0.5 - \Delta F) = 0.5 + \Delta F = 0.5 + \frac{0.5f}{f-1} = \frac{f-0.5}{f-1} \\
&\Rightarrow \left|\frac{\Delta F}{1-F}\right| = -\frac{\Delta F}{1-F} = -\frac{0.5f}{f-1}\cdot\frac{f-1}{f-0.5} = \frac{0.5f}{0.5-f}
\end{aligned} \tag{D23}$$

When $0 < \Delta F < 0.5$ (*i.e.*, $-0.5 < \Delta R < 0$),

$$\begin{aligned}
&\Delta F = Ff = (0.5 - \Delta F)f \Rightarrow \Delta F = 0.5f - f\Delta F \Rightarrow (1+f)\Delta F = 0.5f \\
&\Rightarrow \Delta F = \frac{0.5f}{1+f} \\
&\Rightarrow 1 - F = 1 - (0.5 - \Delta F) = 0.5 + \Delta F = 0.5 + \frac{0.5f}{1+f} = \frac{0.5+f}{1+f} \\
&\Rightarrow \left|\frac{\Delta F}{1-F}\right| = \frac{\Delta F}{1-F} = \frac{0.5f}{1+f}\cdot\frac{1+f}{0.5+f} = \frac{0.5f}{0.5+f}
\end{aligned} \tag{D24}$$

As a result, if we define $|\Delta R/R| = r$, when $0 \leq \Delta R < 0.5$, **Eq (D21)** is rewritten as:

$$r = \frac{0.5f}{0.5 - f} = \frac{0.5(q + h + l)}{0.5 - (q + h + l)} \tag{D25}$$

When $-0.5 < \Delta R < 0$, **Eq (D21)** is rewritten as:

$$r = \frac{0.5f}{0.5 + f} = \frac{0.5(q + h + l)}{0.5 + (q + h + l)} \tag{D26}$$

In the ITC data analysis process, the variables include the independent variables of total ligand concentration L_0 and the molar ratio between target and ligand $\Gamma = T_0/L_0$, and the dependent variable of the measured heat Q_i. The systematic error Γ is propagated from the systematic errors of T_0 and L_0:

$$\Gamma = \frac{T_0}{L_0} \Rightarrow \left|\frac{\Delta\Gamma}{\Gamma}\right| = \left|\frac{\Delta T_0}{T_0}\right| + \left|\frac{\Delta L_0}{L_0}\right| \tag{D27}$$

The absolute value of the relative systematic error of L_0 has been defined as $|\Delta L_0/L_0| = l$. If we define $|\Delta\Gamma/\Gamma| = \gamma$ and $|\Delta T_0/T_0| = t$, then:

$$\gamma = t + l \Rightarrow t = \gamma - l \tag{D28}$$

Our previous work shows that, for strongly correlated error sources:

$$\begin{aligned} &\left|\frac{\Delta K_\mathrm{d}}{K_\mathrm{d}}\right| = a + b\frac{L_0}{K_\mathrm{d}}, \\ &a = t + \frac{r(1+r)}{0.5+r},\ b = \frac{0.5(t+l-r)}{1+r} + \frac{0.5r}{0.5+r} \ \text{ when } 0 \le \Delta R < 0.5 \\ &a = t + \frac{r(1-r)}{0.5-r},\ b = \frac{0.5(t+l-r)}{1-r} + \frac{0.5r}{0.5-r} \ \text{ when } -0.5 < \Delta R < 0 \end{aligned} \tag{D29}$$

Since the equation for K_d determination used in ITC experiments is fundamentally the same as the one used in our previous study (as explained in the main text), we only need to express t and r (from **Eq (D29)**) in terms of q, h, l, and γ to derive the relationship between the relative systematic error in K_d ($|\Delta K_\mathrm{d}/K_\mathrm{d}|$) and L_0/K_d (which is termed as c-value in ITC terminology) for correlated error sources in ITC. The resulting equation is:

$$\left|\frac{\Delta K_{\mathrm{d}}}{K_{\mathrm{d}}}\right| = A + Bc,$$

For $f = q + h + l$:

When $0 \le \Delta R < 0.5$,

$$A = (\gamma - l) + \frac{f(1-f)}{0.5-f},\ B = \frac{(\gamma + f)(0.5-f)}{1-f} \tag{D30}$$

When $-0.5 < \Delta R < 0$,

$$A = (\gamma - l) + \frac{f(1+f)}{0.5+f},\ B = \frac{(\gamma + f)(0.5+f)}{1+f}$$

Note D1.3: Derivation of the Dependence of the Accuracy of ITC-Derived K_d ($\lvert\Delta K_d/K_d\rvert$) on c-Value for Weakly Correlated Error Sources

If the systematic errors of Q_i, ΔH, and L_0 are weakly correlated, the error propagation for **Eq (D15)** can be calculated through:

$$|\Delta F| = \sqrt{\left(\frac{\partial F}{\partial Q_i}\right)^2 \Delta {Q_i}^2 + \left(\frac{\partial F}{\partial (\Delta H^\circ)}\right)^2 \Delta(\Delta H^\circ)^2 + \left(\frac{\partial F}{\partial L_0}\right)^2 \Delta {L_0}^2} \tag{D31}$$

To calculate the relative systematic error of F, we divide both sides of **Eq (D31)** by **Eq (D15)**, and rearrange the equation to be:

$$\left|\frac{\Delta F}{F}\right| = \sqrt{\left(\frac{\Delta Q_i}{Q_i}\right)^2 + \left(\frac{\Delta(\Delta H^\circ)}{\Delta H^\circ}\right)^2 + \left(\frac{\Delta L_0}{L_0}\right)^2} \tag{D32}$$
$$f = \sqrt{q^2 + h^2 + l^2}$$

Since **Eqs (D15)**–**(D24)** are still valid, we have:

when $0 \le \Delta R < 0.5$, **Eq (D21)** is rewritten as:

$$r = \frac{0.5f}{0.5 - f} = \frac{0.5\sqrt{q^2 + h^2 + l^2}}{0.5 - \sqrt{q^2 + h^2 + l^2}} \tag{D33}$$

when $-0.5 < \Delta R < 0$, **Eq (D21)** is rewritten as:

$$r = \frac{0.5f}{0.5 + f} = \frac{0.5\sqrt{q^2 + h^2 + l^2}}{0.5 + \sqrt{q^2 + h^2 + l^2}} \tag{D34}$$

In cases where the error sources are weakly correlated, the systematic error in Γ can be propagated from the systematic errors in T_0 and L_0 with the following formula:

$$\left|\Delta\Gamma\right| = \sqrt{\left(\frac{\partial\Gamma}{\partial T_0}\right)^2 {\Delta T_0}^2 + \left(\frac{\partial\Gamma}{\partial L_0}\right)^2 {\Delta L_0}^2} \tag{D35}$$

To calculate the absolute value of the relative systematic error of Γ, we divide both sides of **Eq (D35)** by Γ and rearrange the equation to be:

$$\left|\frac{\Delta\Gamma}{\Gamma}\right| = \sqrt{\left(\frac{\Delta T_0}{T_0}\right)^2 + \left(\frac{\Delta L_0}{L_0}\right)^2} \text{ or } \gamma = \sqrt{t^2 + l^2} \tag{D36}$$

which induces that:

$$t = \sqrt{\gamma^2 - l^2} \tag{D37}$$

Our previous work demonstrates that for weakly correlated error sources:

$$\left|\frac{\Delta K_\mathrm{d}}{K_\mathrm{d}}\right| = \sqrt{\alpha^2 + \lambda\frac{L_0}{K_\mathrm{d}} + \beta^2\left(\frac{L_0}{K_\mathrm{d}}\right)^2},$$

When $0 \le \Delta R < 0.5$,

$$\alpha = \sqrt{t^2 + \left(\frac{r + r^2}{0.5 + r}\right)^2},\ \beta = \frac{0.5}{1 + r}\sqrt{t^2 + l^2 + \left(\frac{0.5r}{0.5 + r}\right)^2},\ \lambda = \frac{t^2}{1 + r} + 2\left(\frac{0.5r}{0.5 + r}\right)^2 \tag{D38}$$

When $-0.5 < \Delta R < 0$,

$$\alpha = \sqrt{t^2 + \left(\frac{r - r^2}{r - 0.5}\right)^2},\ \beta = \frac{0.5}{1 - r}\sqrt{t^2 + l^2 + \left(\frac{0.5r}{r - 0.5}\right)^2},\ \lambda = \frac{t^2}{1 - r} + 2\left(\frac{0.5r}{r - 0.5}\right)^2$$

We now express t and r (from **Eq (D38)**) in terms of q, h, l, and γ to derive the relationship between the absolute value of the relative systematic error in K_d ($|\Delta K_d/K_d|$) and c-value (L_0/K_d) for weakly correlated error sources in ITC. The resulting equation is:

$$\left|\frac{\Delta K_d}{K_d}\right| = \sqrt{\Theta^2 + \Lambda c + \Omega^2 c^2},$$

For $f = \sqrt{q^2 + h^2 + l^2}$:

When $0 \le \Delta R < 0.5$,

$$\Theta = \sqrt{\left(\gamma^2 - l^2\right) + \left[\frac{f(1-f)}{0.5-f}\right]^2},\ \Omega = \frac{0.5-f}{1-f}\sqrt{\gamma^2 + f^2},\ \Lambda = \frac{0.5-f}{0.5(1-f)}\left(\gamma^2 - l^2\right) + 2f^2 \qquad \text{(D39)}$$

When $-0.5 < \Delta R < 0$,

$$\Theta = \sqrt{\left(\gamma^2 - l^2\right) + \left[\frac{f(1+f)}{0.5+f}\right]^2},\ \Omega = \frac{0.5+f}{1+f}\sqrt{\gamma^2 + f^2},\ \Lambda = \frac{0.5+f}{0.5(1+f)}\left(\gamma^2 - l^2\right) + 2f^2$$

Figure D1.1: General Trends in Dependence of $|\Delta K_d/K_d|$ on c-Value for Constant Relative Systematic Errors in Variables (For $-0.5 < \Delta R < 0$)

Reasonable values were used for A, B, Θ, Λ, and Ω, calculated from typical values of $|\Delta Q_i/Q_i|$, $|\Delta(\Delta H^\circ)/\Delta H^\circ|$, $|\Delta L_0/L_0|$, and $|\Delta\Gamma/\Gamma|$ as explained in the section "Dependence of Relative Systematic Errors in K_d on c-Value in a Wide Range of c-Values" in the main text. These values are: $|\Delta Q_i/Q_i| = |\Delta(\Delta H^\circ)/\Delta H^\circ| = 0.01$, $|\Delta L_0/L_0| = 0.05$, and $|\Delta\Gamma/\Gamma| = 0.1$. The dependence of $|\Delta K_d/K_d|$ on c is shown for panel **(a)** using **Eq (D30)** and for panel **(b)** using **Eq (D38)**. Data are presented on a log-log scale to enhance the clarity of details for low c-values. Notably, the general trend of $|\Delta K_d/K_d|$ on the c-value (for $-0.5 < \Delta R < 0$) obtained here highly agrees with the trend obtained for $0 \leq \Delta R < 0.5$ (**Figure 5.1** in the main text).

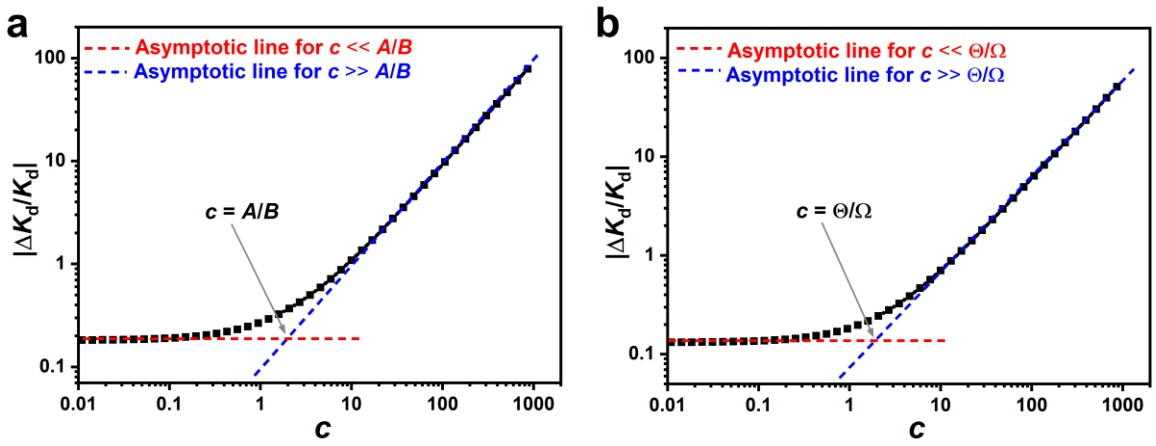
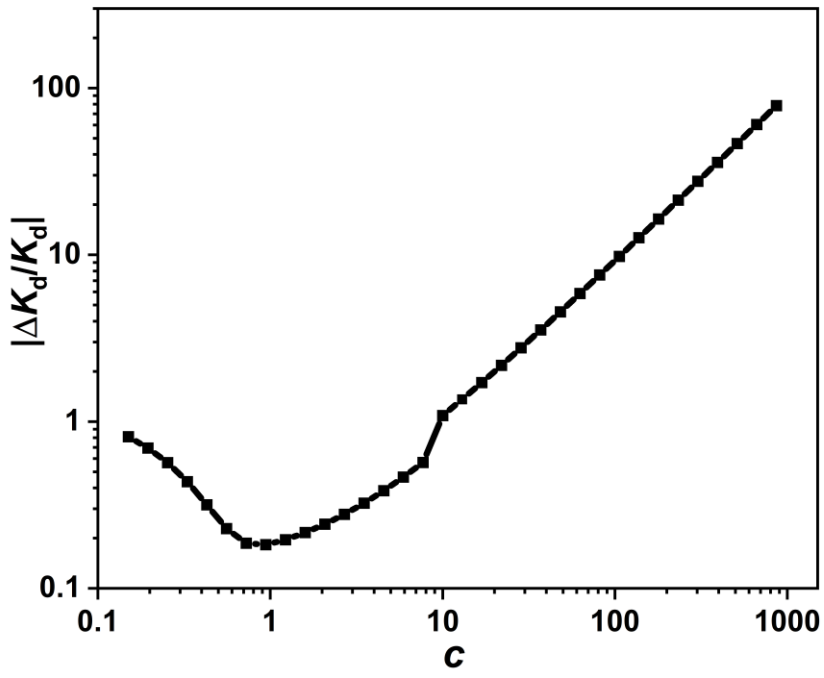


Figure D1.2: General Trends in the Dependence of $|\Delta K_d/K_d|$ on c-Value Under Realistic Experimental Conditions (For $-0.5 < \Delta R < 0$)

The parameters A, B, Θ, Λ, and Ω were calculated using typical values for $|\Delta Q_i/Q_i|$, $|\Delta(\Delta H°)/\Delta H°|$, $|\Delta L_0/L_0|$, and $|\Delta\Gamma/\Gamma|$, consistent with those in **Figure D1.1**, and held constant across the entire c-value range. For $|\Delta(\Delta H°)/\Delta H°|$, we assume $|\Delta(\Delta H°)/\Delta H°| = \exp(-4.6c)$ for $c < 10$ and $|\Delta(\Delta H°)/\Delta H°| = 0.01$ for $c \geq 10$. The dependence of $|\Delta K_d/K_d|$ on c was calculated using **Eq (D38)** for $0.1 < c < 10$ and **Eq (D30)** for $10 \leq c < 1000$. Data are presented on a log-log scale to enhance the clarity of details for low c-values. Notably, the general trend of $|\Delta K_d/K_d|$ on the c-value (for $-0.5 < \Delta R < 0$) obtained here highly agrees with the trend obtained for $0 \leq \Delta R < 0.5$ (**Figure 5.2** in the main text).



Appendix D2. Supporting Information for "A Browser-Based Tool for Assessing Accuracy of ITC-Derived Parameters: K_d, $\Delta H°$, and n"

Note D2.1: Nonlinear Regression and Bootstrapping Overview in ITC Data Analysis

This note introduces two complementary methods for analyzing ITC data: nonlinear regression and Monte Carlo bootstrapping. Nonlinear regression is the standard approach for fitting binding models to calorimetric data, offering a direct means of extracting key thermodynamic parameters such as K_d, enthalpy change $\Delta H°$, and binding stoichiometry n. However, its built-in assumptions and analytical approximations can lead to misleading error estimates, particularly when data are with considerable random/systematic errors. To address these limitations, bootstrapping provides a simulation-based alternative that does not rely on restrictive statistical assumptions. By comparing these two approaches, this note aims to clarify not only how each method can be implemented, but also when and why one may offer advantages over the other in practical ITC data analysis.

D2.1.1. Nonlinear Regression in ITC

In a typical ITC experiment, a ligand is incrementally titrated into a solution containing a macromolecule, and the heat released or absorbed during each injection is recorded. These integrated heats reflect the binding progress and can be interpreted using a theoretical binding model, most commonly the one-site binding model. This model includes three key parameters: K_d, $\Delta H°$, and n. Because these parameters are embedded in nonlinear relationships, nonlinear regression must be used. This approach involves iteratively adjusting parameter values to minimize the discrepancy between observed and predicted heats, typically through least-squares optimization algorithms.

In this study, the model used to describe the cumulative heat per mole of injectant (**Eq (5.7)** in the main text) is inherently nonlinear in all three fitted parameters. The equation takes into account the total ligand concentration after each injection, the initial concentration of the macromolecule, and the K_d, cumulative ΔH ($\Delta H_{cum,i}$), and n as parameters to be optimized. Because of its nonlinear structure, fitting this equation requires numerical routines such as "scipy.optimize.curve_fit", which efficiently searches for parameter values that minimize the sum of squared residuals.

D2.1.2. Limitations of Conventional Nonlinear Regression

Although widely used, conventional nonlinear least-squares regression comes with important limitations that may be overlooked. It assumes that the variance of residuals remains constant across all data points — a condition known as homoscedasticity. In ITC, however, this assumption often fails, as measurement errors tend to grow with increasing injection number or heat magnitude. Furthermore, standard regression routines only account for the uncertainty in heat measurements, disregarding experimental variability in ligand or macromolecule concentrations. Perhaps most critically, the commonly reported confidence intervals for fitted parameters are derived from local linear approximations of the model, which can be invalid in highly nonlinear regimes or when data are sparse. These issues can lead to overconfident and potentially misleading conclusions, especially when the binding is weak, the noise of signal is high, or the titration curve does not span from baseline to saturation.

D2.1.3. Monte Carlo Bootstrapping

To overcome these limitations and obtain more reliable estimates of parameter uncertainty, Monte Carlo bootstrapping offers a valuable alternative. Rather than relying on analytical approximations, this method simulates the impact of experimental errors by generating many

synthetic datasets. The process begins by fitting the model once to obtain predicted heat values and the corresponding residuals. New datasets are then created by resampling these residuals and adding them back to the predicted heats. Each synthetic dataset is refitted independently, and the entire process is repeated hundreds or thousands of times (depending on the user-defined iteration number). This generates an empirical distribution for each fitted parameter.

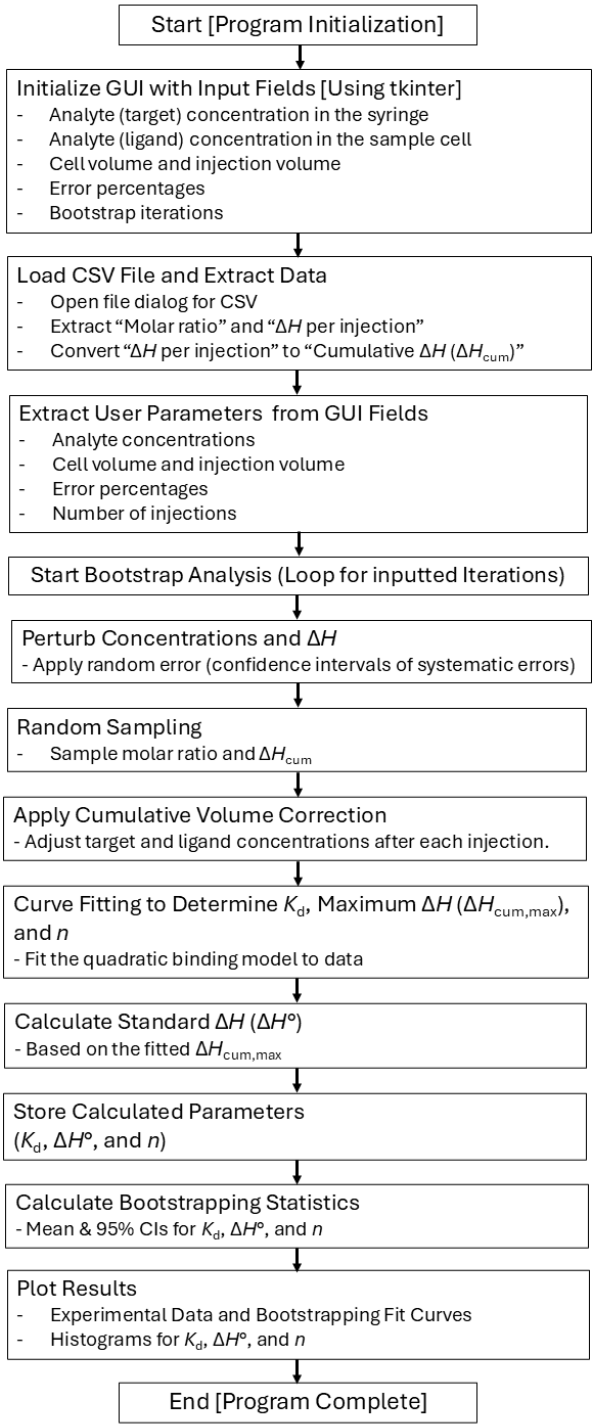
From this distribution, one can directly extract confidence intervals, for instance, the 2.5th and 97.5th percentiles provide a 95% confidence interval for the K_d value. Unlike conventional nonlinear regression methods, bootstrapping makes no assumptions about the underlying error structure and remains robust even when residuals are heteroscedastic, the model is highly sensitive to initial values, or the dataset is small. In practice, bootstrapped confidence intervals tend to be wider but more realistic, offering a better reflection of the actual uncertainty inherent in the experiment.

In summary, nonlinear regression remains a practical and widely used method for fitting ITC data and extracting binding parameters. However, the default confidence estimates produced by this approach can be unreliable when its underlying statistical assumptions are not satisfied. Monte Carlo bootstrapping serves as a valuable complement by providing empirically derived uncertainty estimates that better reflect the structure and variability of real experimental data.

Figure D2.1: Detailed Computational Algorithm for the ACI-ITC Program

We developed a user-friendly web application for calculating the Accuracy Confidence Interval (ACI) of ITC-derived parameters, including K_d, $\Delta H°$, and n. The process began with the design of a detailed computational workflow (**Figure D2.1**). This workflow was implemented in a Python program that encapsulates the computational algorithm illustrated in **Figure D2.1.** The

Python code was then deployed on a web server, resulting in the creation of the ACI web app, which is accessible at https://aci.sci.yorku.ca.



Note D2.2: Validation for the ACI-ITC Program

To validate the ACI-ITC program, it is essential to demonstrate its ability to produce theoretical results accurately. Since the true values of parameters such as K_d, $\Delta H°$, n, L_0, and T_0, along with their systematic errors, are inherently unknown and uncontrollable in real experiments, synthetic data with accurately defined variables and parameters provide the most reliable validation approach. To this end, we developed a Python-based program called "ITC Data Generator" to simulate ITC binding isotherms (ΔH vs molar ratio), which is publicly available on Figshare (https://doi.org/10.6084/m9.figshare.28503245). Using this program, we generated 100 synthetic binding isotherms based on the following true values: $K_d = 2 \times 10^{-6}$ M (2 μM), $\Delta H° = -40$ kcal/mol, $L_0 = 1 \times 10^{-4}$ M (100 μM), $T_0 = 1 \times 10^{-3}$ M (1000 μM), sample cell volume $V_0 = 100$ μL, injection volume $V_{inj} = 1$ μL, and 20 total injections. To mimic

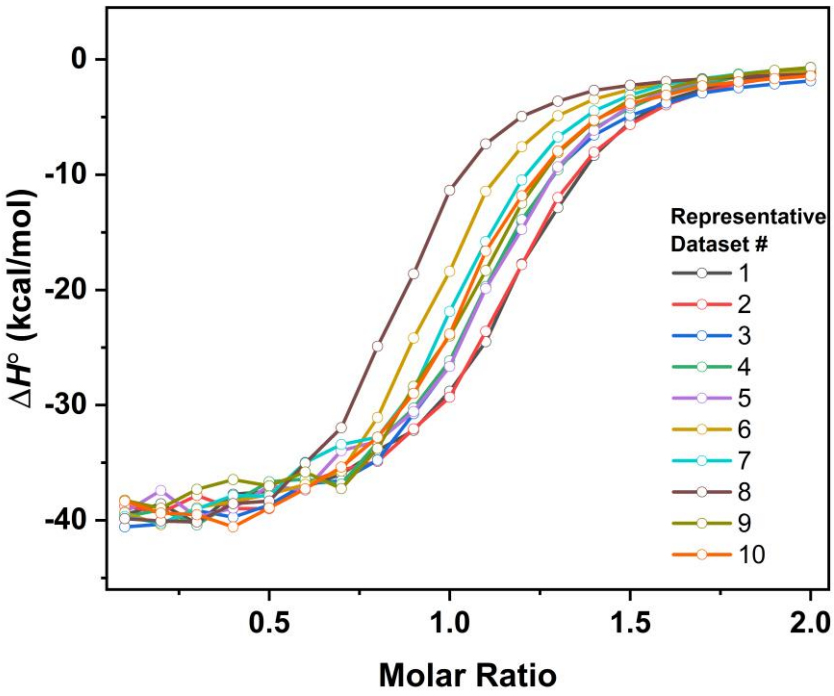


Figure D2.2. Representative ITC binding isotherms generated by the "ITC Data Generator" program. The details of the parameters used in the synthetic data-generating process are explained in Note D2.2.

experimental variability, we introduced controlled errors, including 5% and 7% relative random errors in the ligand and target stock concentrations, respectively, a systematic error in ITC-measured ΔH with a standard deviation of 0.8 kcal/mol, and a 1% relative random error in the generated ΔH values. These conditions enabled realistic simulations for testing the program's capability, and ten representative binding isotherms are shown in **Figure D2.2**.

All synthetic ITC binding isotherms were analyzed using the ACI-ITC program by inputting the same parameters used during data generation to compute the Accuracy Confidence Interval (ACI) for K_d, $\Delta H°$, and n at a 95% confidence level. To validate the ACI-ITC program, it is necessary to assess whether the calculated ACI for each parameter consistently include the true values with a theoretical probability. For K_d and $\Delta H°$, the true values are defined by the input parameters of the "ITC Data Generator" program (2 μM and –40 kcal/mol, respectively), as their physical meanings are well established. However, for n, which reflects corrections for concentration inaccuracies rather than strict stoichiometry,[1, 2] the theoretical value remains uncertain due to its dependence on the interplay between concentration inaccuracies and the errors of measured heat changes. While the data generation process assumes a 1:1 binding model, deviations in n are acceptable within the range of 0.7 to 1.3, which is typically interpreted as single-site binding in ITC experiments. Thus, in this study, we focused on verifying whether the calculated ACI for K_d and $\Delta H°$ encompass their respective true values at or above the theoretical probabilities.

To validate the ACI-ITC program, it is necessary to calculate the theoretical probabilities. At first glance, one might assume that the theoretical probabilities for both parameters are simply 95%. However, having 100 ACI with a 95% confidence level does not guarantee that 95 or more of these intervals will include the true value. The 95% confidence level refers to the long-term

behavior of the method used to construct the intervals, not the outcome of a specific finite sample (*e.g.*, 100 samples in this case). In practice, due to random variability and sampling error, the number of intervals that include the true value in a particular set of 100 intervals may differ from 95. For example, the observed coverage in a specific dataset might be slightly higher or lower than 95%, depending on the characteristics of the synthetic data and the randomness introduced during its generation and analysis. Thus, while the goal is for approximately 95% of the intervals to cover the true value, validating this requires careful statistical testing and interpretation of the observed coverage across multiple synthetic datasets.

For 100 calculated ACI with a 95% confidence level, the probability distribution of the number of intervals containing the true value follows a Gaussian distribution with a mean (μ) of $100 \times 0.95 = 95$ and a standard deviation (σ) of $\sqrt{(100 \times 0.95 \times (1-0.95))} = 2.18$. Therefore, the

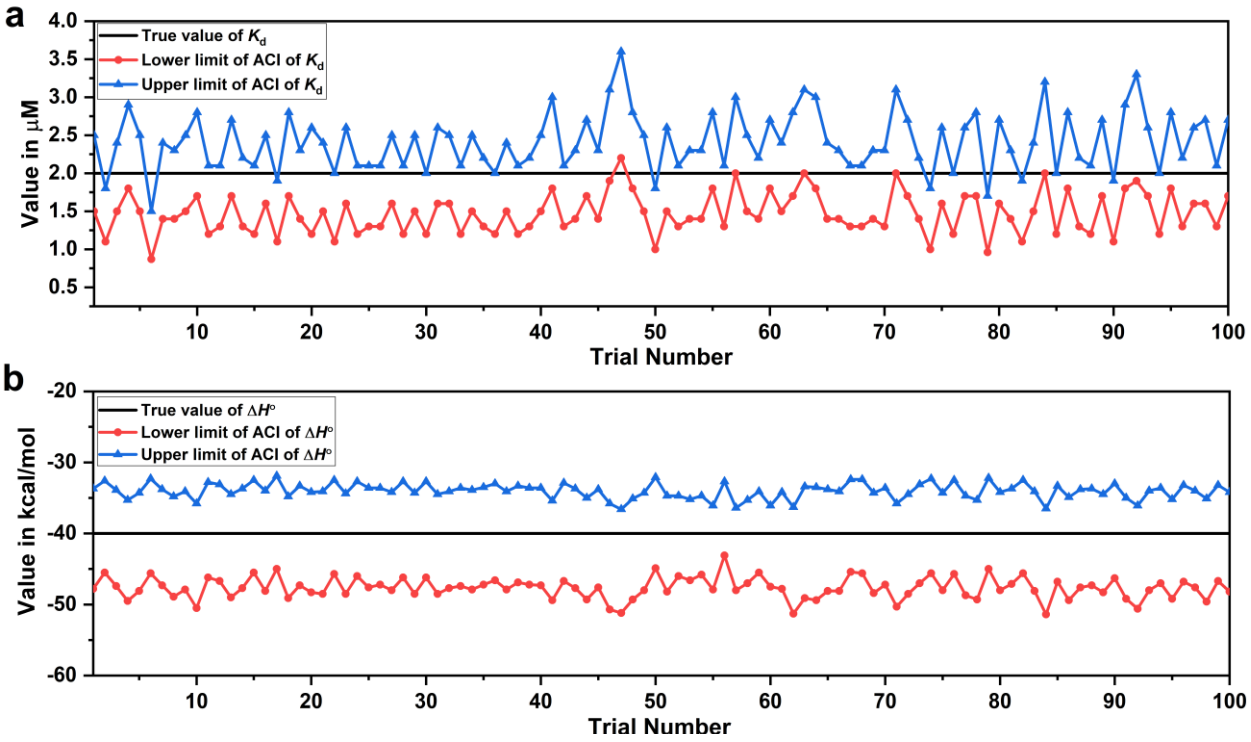


Figure D2.3. The comparison between the ACI-ITC calculated ACI for K_d **(a)** and $\Delta H°$ **(b)** and their corresponding true values.

number of intervals expected to include the true value lies between $95 - 1.96\sigma \approx 91$ and $95 + 1.96\sigma \approx 99$ with 95% probability. To validate the ACI-ITC program, we conclude it is successful if both the calculated ACI for K_d and $\Delta H°$ include the true values (K_d = 2 μM and $\Delta H°$ = −40 kcal/mol) in at least 91 out of 100 intervals.

Figure D2.3 illustrates the comparison between the ACI-ITC calculated ACI for K_d and $\Delta H°$ and their corresponding true values. **Figure D2.3** shows that 91 out of 100 calculated ACI for K_d include the true K_d value of 2 μM, and 100 out of 100 ACI for $\Delta H°$ include the true $\Delta H°$ value of –40 kcal/mol. These results confirm the successful validation of the ACI-ITC program.

Note D2.3: The Application of the ACI-ITC Tool to Retrospective ITC Data

To further demonstrate the generalizability of the ACI-ITC approach for calculating the ACI of ITC-derived parameters, we applied it to a retrospective ITC dataset for the UA-3BP aptamer–uric acid system. The nominal initial concentration of the UA-3BP aptamer in the sample cell ($[\text{UA-3BP}]_0$, nominal L_0) was 61.2 μM, and the nominal initial concentration of uric acid in the syringe ($[\text{uric acid}]_0$, nominal T_0) was 936 μM. For this dataset, where direct assessment of concentration systematic error confidence intervals was impractical, a minimum value of 5% was assigned to estimate the lower limit of the ACI. A comparison of the probability distributions of K_d, $\Delta H°$, and n determined by ACI-ITC and by conventional least-squares (L-S) nonlinear regression (using the Levenberg–Marquardt algorithm) is presented in **Figure D2.4**.

The results in **Figure D2.4** support a similar conclusion to that drawn from **Figure 5.5** in the main text — significant differences often persist between the ACI and the conventionally determined CI. First, accounting for systematic uncertainties in concentrations and measured heats consistently results in a wider 95% ACI compared to the 95% CI from conventional L-S regression, with the most pronounced widening observed for $\Delta H°$. More importantly, the results confirm that uncertainties in concentrations and measured heats have a much greater impact on the determined mean K_d values than on the mean

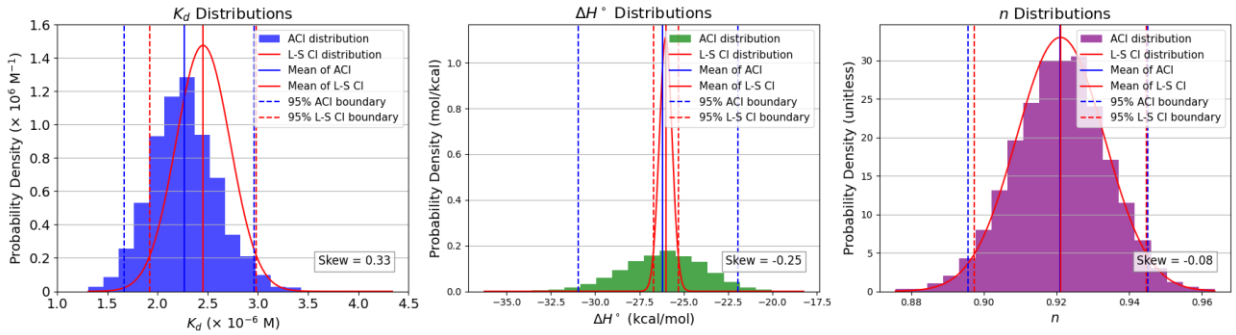


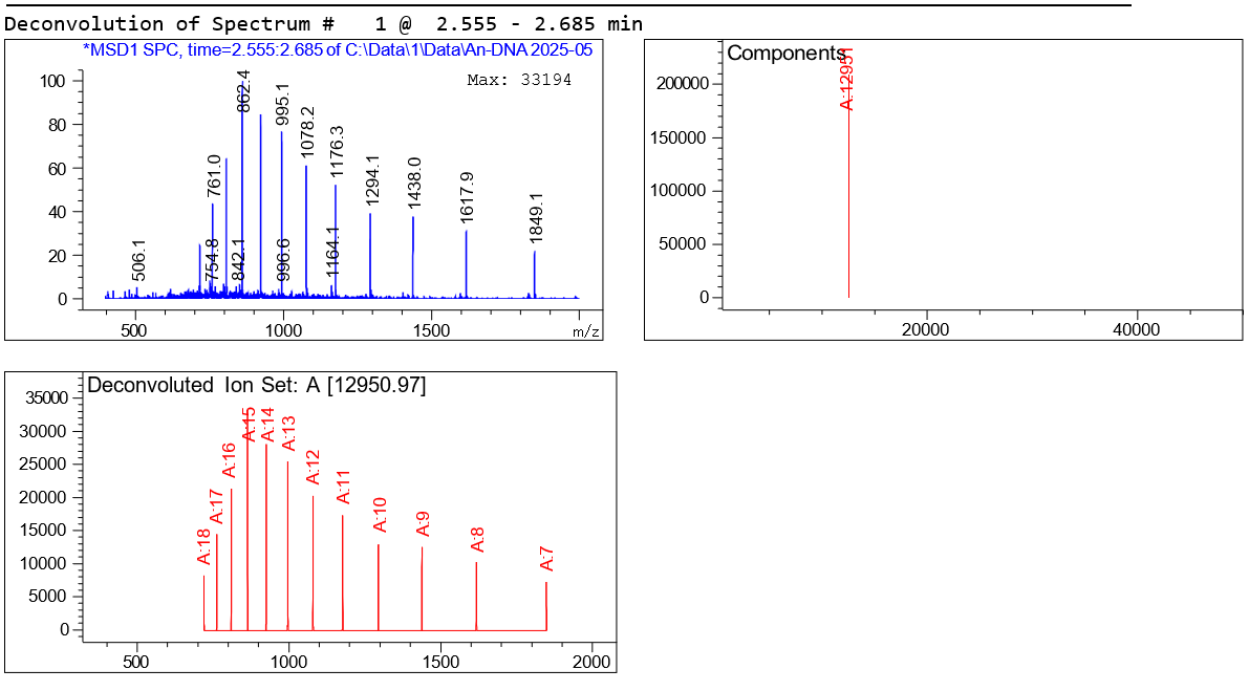
Figure D2.4. Comparison of the probability density distributions for the ACI calculated using ACI-ITC versus the CIs obtained from conventional least-squares (L-S) regression for the UA-3BP aptamer–uric acid binding pair. ACI histograms were generated from 5,000 bootstrap iterations. Gaussian probability distributions were constructed using the mean and standard deviation values obtained from L-S nonlinear regression of the ITC binding isotherms using the Levenberg–Marquardt algorithm. Skewness values are reported in each panel to indicate distribution asymmetry.

values of $\Delta H°$ and n. Specifically, for the UA-3BP aptamer–uric acid interaction, the 95% ACI for K_d spanned 1.7–3.0 μM, with a mean of 2.3 μM, whereas the 95% CI from L-S regression spanned 1.9–3.0 μM, with a mean of 2.5 μM. The mean K_d value determined by ACI-ITC was 8% lower than that determined by conventional regression, while the mean values of $\Delta H°$ and n differed by less than 0.3% between the two methods. It is important to note that only the theoretically narrowest confidence intervals for concentration and heat systematic errors were input here; the true ACI could differ much more substantially from the CI determined by conventional L-S regression.

Note D2.4: Determination of the Purities of Reagents

D2.4.1. Assessment of the purity of Theo2201 aptamer

To assess the purity of the Theo2201 aptamer obtained from Integrated DNA Technologies (IDT, Coralville, IA, USA), a 10 µM solution was prepared in deionized water containing 0.1% formic acid and analyzed by liquid chromatography-mass spectrometry (LC-MS). The UV chromatogram at 260 nm showed a single peak at approximately 2.6 minutes, corresponding to the ssDNA. To evaluate the purity of this peak, the associated mass spectrum was examined (**Figure D2.5**). The MS analysis report (**Figure D2.5**) revealed a single deconvoluted component, indicating that no detectable impurities were present and confirming that the aptamer purity was effectively 100%.



Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	12950.97	206022	100.00

*** End of Report ***

Figure D2.5. Mass spectrometry analysis report for assessing the purity of Theo2201 aptamer.

D2.4.2. Assessment of the purity of theophylline

Theophylline powder (CAS No. 58-55-9) was purchased from Sigma-Aldrich (Oakville, ON, Canada), with a manufacturer-reported purity of ≥ 99% by HPLC. To independently assess its purity, a 20 µM theophylline solution was prepared in deionized water containing 0.1% formic acid and analyzed by LC-MS. The UV chromatogram at 270 nm is presented in **Figure D2.6b**. Since 270 nm is the optimal absorbance wavelength for theophylline and its common impurities (*e.g.*, caffeine, theobromine, and xanthine), any impurity peaks would be expected to appear at this wavelength. As shown in **Figure D2.6b**, by comparing to the result from the negative control (**Figure D2.6a**), only a single prominent peak was observed in the sample, suggesting that the reagent was of high purity.

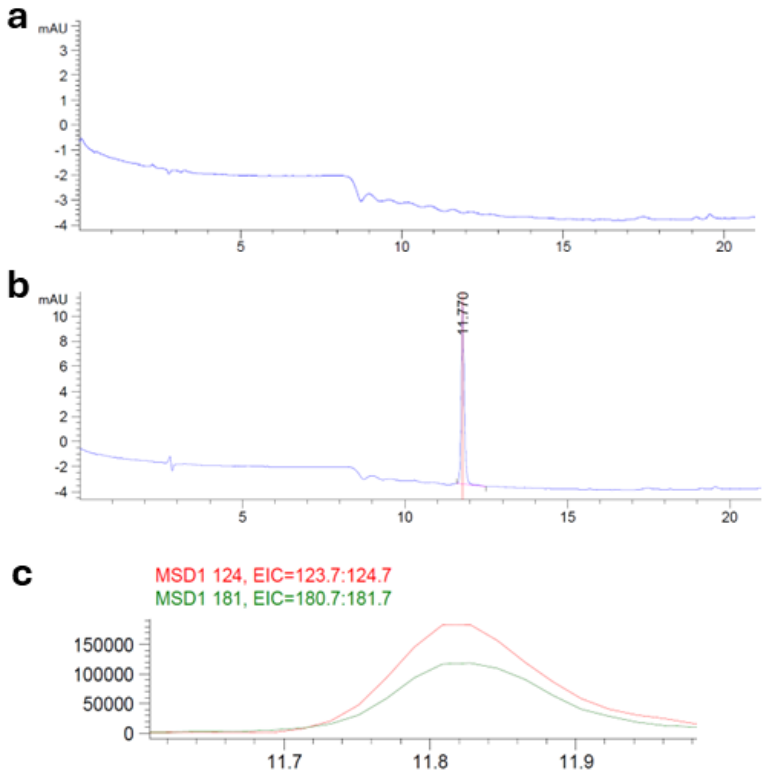
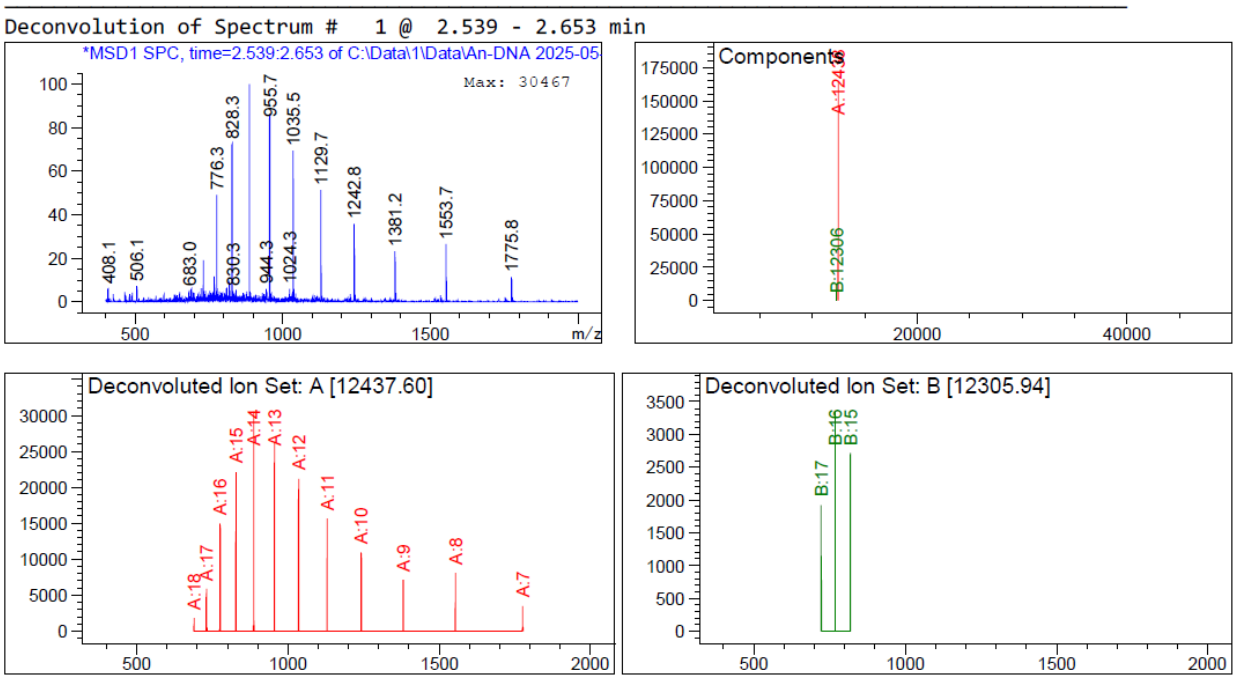


Figure D2.6. LC-MS analysis for assessing the purity of the theophylline sample. **a)** UV chromatogram at 270 nm for the negative control (injection of deionized water containing 0.1% formic acid). **b)** UV chromatogram at 270 nm for 20 µM theophylline solution, showing a single peak at 11.8 minutes. **c)** Mass spectrum corresponding to the peak at 11.8 minutes, used to confirm the identity and purity of the eluted component.

To further assess whether the observed peak contained co-eluting components, such as theophylline and theobromine, which are known to have similar retention times, a more detailed analysis was performed using the corresponding MS spectrum (**Figure D2.6c**). The data show that the single UV chromatographic peak at 11.8 minutes contained ions with m/z values of 181 and 124, which correspond to the protonated molecular ion and a common fragment of theophylline, respectively. These results confirm that no detectable impurities were present within the peak, indicating that the impurity level in the theophylline sample is negligible.

D2.4.3. Assessment of the purity of DA-3BP aptamer

To assess the purity of the DA-3BP aptamer obtained from Integrated DNA Technologies (IDT, Coralville, IA, USA), a 10 µM solution was prepared in deionized water containing 0.1% formic acid and analyzed by LC-MS. The UV chromatogram at 260 nm displayed a single peak eluting at approximately



Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	12437.60	164641	100.00
B	12305.94	6735	4.09

*** End of Report ***

Figure D2.7. Mass spectrometry analysis report for assessing the purity of DA-3BP aptamer.

2.6 minutes, consistent with the expected retention time for the ssDNA aptamer. The mass spectrum corresponding to this peak was further analyzed (**Figure D2.7**). The deconvolution results identified two species with molecular weights of 12,437.60 and 12,305.94 Da, respectively. The major component (12,437.60 Da) accounted for 96% of the total ion abundance, while the minor component (12,305.94 Da) accounted for 4%. Based on this analysis, the purity of the DA-3BP aptamer was determined to be approximately 96%.

D2.4.4. Assessment of the purity of dopamine

Dopamine hydrochloride powder (CAS No. 62-31-7) was purchased from Sigma-Aldrich (Oakville, ON, Canada), with a manufacturer-reported purity of ≥97.5% (HPLC). To independently evaluate its purity, a 100 μM solution was prepared in deionized water containing 0.1% formic acid and analyzed by LC-MS. The corresponding UV chromatograms recorded at 280 nm are shown in **Figure D2.8**. Since 280 nm is the optimal absorbance wavelength for detecting dopamine and its common impurities (e.g., L-DOPA and Catechol), the initial assessment focused on this channel to examine the presence of potential impurities. As shown in **Figure D2.8b**, comparison with the negative control (**Figure D2.8a**) revealed a single prominent peak eluting at approximately 3.6 minutes, suggesting the absence of detectable impurities under these conditions. To identify the species corresponding to the observed peak at approximately 3.6 minutes, the associated mass spectrum was examined (**Figure D2.8c**). The results revealed two dominant ions with m/z values of 154 and 137, corresponding to the protonated molecular ion of dopamine and its characteristic fragment resulting from the loss of ammonia (–17 Da), respectively. These findings confirm that the single UV peak at 3.6 minutes arises solely from dopamine, with no additional detectable impurities under the tested conditions.

Since polymeric oxidation products of dopamine typically exhibit strong absorbance between 350 and 450 nm, the UV chromatograms were further analyzed at 400 nm to evaluate their presence. The resulting chromatograms are shown in **Figure D2.9**. As illustrated in **Figure D2.9b**, a small peak was observed at approximately 2–3 minutes in the dopamine sample, which was not present in the negative control

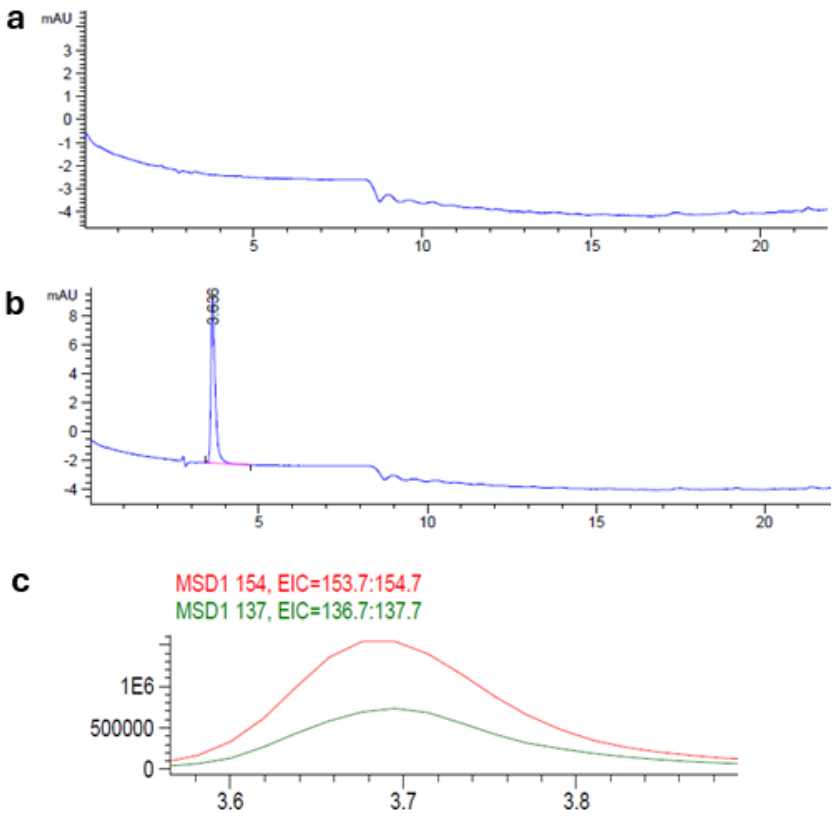


Figure D2.8. Initial assessment of dopamine purity based on UV and MS chromatograms. **a)** UV chromatogram at 280 nm of the negative control (deionized water containing 0.1% formic acid). **b)** UV chromatogram at 280 nm of a 100 µM dopamine solution prepared in deionized water with 0.1% formic acid. **c)** Extracted ion chromatograms for dopamine showing m/z 154.7 (red), corresponding to the protonated molecular ion, and m/z 137.7 (green), corresponding to its characteristic fragment ion. Both ions co-elute at approximately 3.6 minutes, confirming that the observed peak originates from dopamine without evidence of co-eluting impurities.

(**Figure D2.9a**). To identify the species associated with this peak, the corresponding mass spectrum was examined (**Figure D2.9c**). The analysis revealed two dominant ions with m/z values of 239 and 241, which do not match any known dopamine-related impurities or oxidation products. These signals are likely attributable to minor, unidentified contaminants introduced during the LC-MS run.

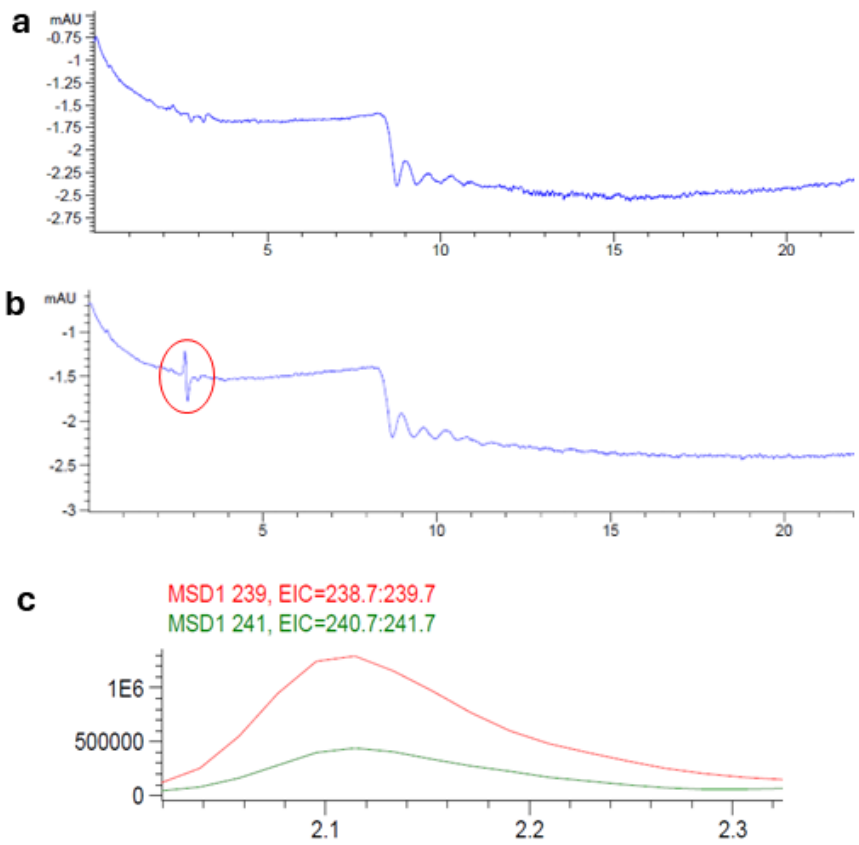


Figure D2.9. Re-assessment of dopamine purity based on UV and MS chromatograms. **a)** UV chromatogram at 400 nm of the negative control (deionized water containing 0.1% formic acid). **b)** UV chromatogram at 400 nm of a 100 µM dopamine solution prepared in deionized water with 0.1% formic acid. **c)** Extracted ion chromatograms for dopamine showing m/z 239 (red) and m/z 241 (green).

Overall, the results presented in **Figures D2.8** and **D2.9** confirm the high purity of the dopamine sample used in the ITC experiments, with no detectable impurities or polymeric oxidation products under the tested conditions.

D2.4.5. Monitoring dopamine oxidation using UV-Vis spectroscopy

According to the results presented in **Figure D2.9**, no apparent dopamine oxidation was observed by LC-MS analysis within one hour of sample preparation. However, the solvent used for LC-MS analysis was deionized water containing 0.1% formic acid, which is highly acidic and may suppress dopamine

oxidation. To reassess the oxidation behavior under conditions relevant to the ITC experiments, I prepared 150 µL of 1.0 mM dopamine in the ITC buffer (20 mM NaxHyPO$_4$, 140 mM NaCl, 2 mM MgCl$_2$, pH 7.4). Since this buffer was incompatible with LC-MS analysis, and because dopamine and its oxidation products have distinct optimal absorbance wavelengths, *i.e.*, 280 nm for dopamine and 350–450 nm for its polymeric oxidation products, the absorbance at 280 nm was monitored using a NanoDrop 1000 spectrophotometer over a period of 90 minutes. This duration approximately matched the total time from dopamine solution preparation to the completion of an ITC experiment. As shown in **Figure D2.10**, no significant change in absorbance at 280 nm (OD_{280}) was observed over the 90-minute period, indicating that dopamine oxidation under ITC experimental conditions was negligible and unlikely to affect the concentration of dopamine during the measurement.

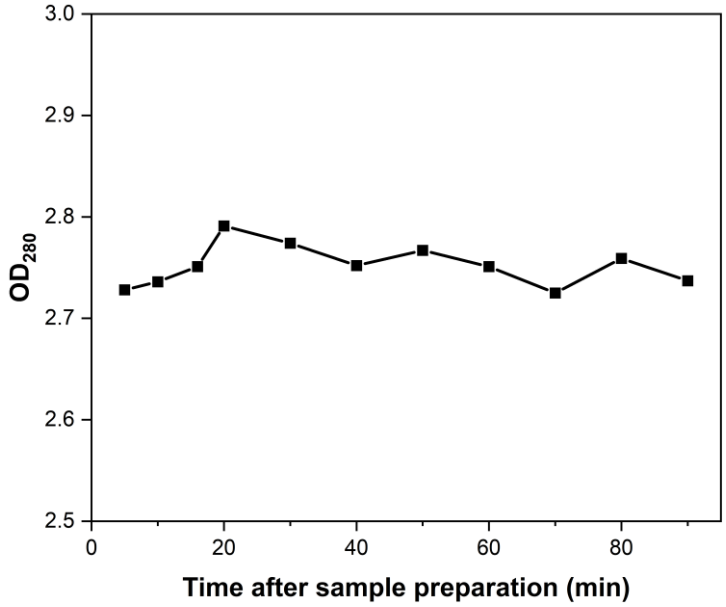


Figure D2.10. Monitoring dopamine oxidation under ITC experimental conditions. The absorbance at 280 nm (OD_{280}) was monitored over a 90-minute period for 1.0 mM dopamine prepared in ITC buffer (20 mM NaxHyPO$_4$, 140 mM NaCl, 2 mM MgCl$_2$, pH 7.4). No significant change in OD_{280} was observed, indicating negligible dopamine oxidation during the typical duration of an ITC experiment.

Note D2.5: Determination of the relative random errors in the analyte concentrations in multiple sample preparations

D2.5.1. Assessment of the relative random error in theophylline concentrations

To assess the relative random error in theophylline concentrations, we independently prepared four aliquots of theophylline solutions from scratch, each with a nominal concentration of 936 μM. The absorbance of these solutions was then measured at 260 nm (OD_{260}) using a spectrophotometer (NanoDrop 1000, Thermo Scientific) with an optical path length of 10 mm. The measured absorbance values and the calculated relative standard deviation (RSD) are summarized in **Table D2.1**.

Table D2.1. Determination of the relative random error for theophylline concentration.

Sample #	**OD_{260}**	**Average OD_{260}**	**SD**	**RSD**
1	6.597	6.3	0.4	6.3%
2	6.179			
3	5.892			
4	6.668			

D2.5.2. Assessment of the relative random error in Theo2201 aptamer concentrations

A stock solution of the Theo2201 aptamer, with a nominal concentration of 2.5 mM determined by OD_{260} measurement using a UV-Vis spectrophotometer, was prepared from a single batch of lyophilized powder obtained from Integrated DNA Technologies (IDT, Coralville, IA, USA). From this stock, five independent 1.0 mL aliquots of 60 μM Theo2201 solutions, matching the nominal concentration used in the ITC experiments, were prepared by dilution. The UV absorbance at 260 nm (OD_{260}) of each aliquot

was measured using a NanoDrop 1000 spectrophotometer. The measurement results are summarized in **Table D2.2**.

Table D2.2. Determination of the relative random error for Theo2201 aptamer concentration.

Sample #	**OD_{260}**	**Average**	**Overall Average**	**SD**	**RSD**
1	21.53	21.63	22.9	0.9	0.04
	21.62				
	21.73				
2	23.00	23.10			
	23.17				
	23.15				
3	23.41	23.56			
	23.66				
	23.62				
4	23.29	23.34			
	23.34				
	23.40				
5	23.09	23.19			
	23.17				
	23.30				

Based on the results presented in **Table D2.2**, the relative random error in the concentrations of the Theo2201 aptamer solutions prepared from the same stock was determined to be 4%. However, to account for additional uncertainty arising from potential batch-to-batch variability, a minimum practical

value of 5% was assigned as the standard deviation of the confidence interval (CI) for the relative systematic error in aptamer concentration when calculating the ACI of the fitted parameters.

D2.5.3. Assessment of the relative random error in dopamine concentrations

To evaluate the relative random error in dopamine concentration, five independent dopamine solutions were freshly prepared from powder, each with a nominal concentration of 7.02 Mm. The absorbance at 280 nm (OD_{280}) for each solution was measured using a NanoDrop 1000 spectrophotometer with a 10 mm optical path length. The measured absorbance values and the calculated RSD are summarized in **Table D2.3**.

Table D2.3. Determination of the relative random error for dopamine concentration.

Sample #	**OD_{280}**	**Average**	**Overall Average**	**SD**	**RSD**
1	12.42	12.70	14.8	1.6	0.11
	13.04				
	12.64				
2	15.44	15.38			
	15.39				
	15.30				
3	15.70	15.34			
	15.02				
	15.29				
4	13.66	13.71			
	13.52				
	13.95				
5	17.09	16.80			

	16.82				
	16.49				

D2.5.4. Assessment of the relative random error in DA-3BP aptamer concentrations

A stock solution of the DA-3BP aptamer, with a nominal concentration of 1.7 mM determined by OD_{260} measurement using a UV-Vis spectrophotometer, was prepared from a single batch of lyophilized powder obtained from IDT. From this stock, five independent 100 µL aliquots of 45 µM DA-3BP solutions were prepared by dilution. This lower concentration was selected to avoid saturating the NanoDrop 1000 spectrophotometer, as the 450 µM solutions used in the ITC experiments exceeded the instrument's linear absorbance range. The absorbance at 260 nm (OD_{260}) for each aliquot was measured, and the results are summarized in **Table D2.4**.

Table D2.4. Determination of the relative random error for DA-3BP aptamer concentration.

Sample #	**OD_{260}**	**Average**	**Overall Average**	**SD**	**RSD**
1	17.79	17.74	17.7	0.5	0.03
	17.69				
	17.75				
2	17.95	18.09			
	18.11				
	18.20				
3	17.98	18.02			
	18.06				
	18.02				

4	17.09 17.03 17.13	17.08			
5	17.72 17.72 17.82	17.75			

As shown in **Table D2.4**, the relative random variability among the independently prepared DA-3BP aptamer solutions from the same stock was calculated to be 3%. Nevertheless, to account for possible additional uncertainty due to batch-to-batch variation, a conservative estimate of 5% was applied as the standard deviation of the CI for the relative systematic error in aptamer concentration when computing the ACI of the fitted parameters.