Received: 31 July 2021 | Revised: 8 November 2021 | Accepted: 9 November 2021

DOI: 10.1002/pro.4240

TOOLS FOR PROTEIN SCIENCE

Analysis of protein–ligand interactions from titrations and nuclear magnetic resonance relaxation dispersions

Emil Björklund[1] | **Anna du Rietz**[2] | **Patrik Lundström**[1]

[1]Division of Chemistry, Department of Physics, Chemistry and Biology, Linköping University, Linköping, Sweden

[2]Division of Molecular Surface Physics and Nanoscience, Department of Physics, Chemistry and Biology, Linköping University, Linköping, Sweden

Correspondence
Patrik Lundström, Division of Chemistry, Department of Physics, Chemistry and Biology, Linköping University, SE-581 83 Linköping, Sweden.
Email: patrik.lundstrom@liu.se

Abstract

We present PLIS, a publicly available, open-source software for the determination of protein–ligand dissociation constants that can be used to characterize biological processes or to shed light on biophysical aspects of interactions. PLIS can analyze data from titration experiments monitored by for instance fluorescence spectroscopy or from nuclear magnetic resonance relaxation dispersion experiments. In addition to analysis of experimental data, PLIS includes functionality for generation of synthetic data, useful for understanding how different parameters effect the data in order to better analyze experiments.

KEYWORDS

dissociation constant, protein–ligand interactions, relaxation dispersion, titration

1 | INTRODUCTION

Protein–ligand interactions are central for processes such as signaling, membrane transport, metabolism, muscle contraction, immune defense, and therapeutics.[1] In addition, the analysis of protein–ligand binding can answer mechanistic questions, such as whether binding proceeds via any of the two limiting cases induced fit or conformational selection.[2] Protein–ligand complexes may be studied by co-crystallizing the protein and ligand to map out the interaction surfaces,[3] and this static picture may be complemented by studying the dynamic equilibrium of binding and dissociation. Important methods of the latter kind are fluorescence spectroscopy,[4] ultraviolet/visible spectrophotometry,[5] circular dichroism spectroscopy,[6] surface plasmon resonance,[7] isothermal calorimetry,[8] and nuclear magnetic resonance (NMR) spectroscopy.[9] Usually, titrations are necessary for determining dissociation constants but sometimes, for example, in some applications of NMR spectroscopy, they can be calculated from experiments performed at a single ligand concentration.

There is no shortage of software for analyzing protein–ligand interactions probed by various techniques in terms of dissociation constants. It has been increasingly popular to publish software as open source. Some benefits are platform flexibility, and that experienced users can tailor or extend the functionality. Examples of open-source software are Anabel for analysis of SPR data[10] and pytc for ITC data.[11] For data derived from titrations probed by, for example, fluorescence or absorption spectroscopy, perhaps the most common software is Prism (Graphpad Inc.) that is not open-source. Prism provides the user with a vast array of binding models and options for how to perform the fit and present the results. For certain methods to determine dissociation constants, such as NMR relaxation dispersion, users are largely left to themselves to write suitable fitting routines.

Herein, we present PLIS (protein ligand interaction software) that is an easy to use open-source tool for the determination of dissociation constants probed by titration experiments or by NMR relaxation dispersions. Binding can be probed by measurements of signals that either

originate from the protein or from a competitor for the ligand.[12] The appearance of graphs representing the data and fits thereof can be tailored and images can be exported in different formats. A key feature of PLIS is a simulation tool. We deliberately chose to limit the functionality compared to for instance Prism (Graphpad Inc.) in the interest of keeping the software intuitive and easy to use. However, additional binding models and other extensions or modifications to PLIS may be added by users who demand them.

2 | THEORY

Binding equilibria and the associated dissociation constants for a protein (P) that may bind n ligands (L) are given by

$$PL \rightleftarrows P + L, \quad K_d^1 = \frac{[P][L]}{[PL]}$$
$$\vdots$$
$$PL_n \rightleftarrows P + L_{n-1}, \quad K_d^n = \frac{[PL_{n-1}][L]}{[PL_n]}$$

Determination of dissociation constants from titration experiments typically involves measurement of a signal that is sensitive to the binding state of a protein as a function of the concentrations of ligand and protein. The signal is modeled as the population weighted average of the signals of the unbound state and the various bound states. For n-site binding, the signal (s) is

$$s = \sum_{k=0}^{k=n} p_k s_k$$

where p_k and s_k are the population and signal of state PL_k, respectively. For methods such as fluorescence spectroscopy and absorption spectroscopy the signal itself scales linearly with concentration, that is, $s_k \propto [PL_k]$. Other signals, such as the chemical shift for a protein–ligand system in fast exchange, do not show such a concentration dependence. This must be considered when fitting the data. The population of state k is in turn given by

$$p_k = \frac{[PL_k]}{\sum_{k=0}^{n} [PL_k]}$$

and can be calculated from the dissociation constants and the total concentrations of protein and ligand

$$[P]_{tot} = \sum_{k=0}^{k=n} [PL_k], [L]_{tot} = [L] + \sum_{k=1}^{k=n} k[PL_k]$$

The concentrations of the free ligand is calculated by solving the following equation

$$[L] = [L]_{tot} - \frac{[P]_{tot}}{1 + \sum_{k=1}^{n} k \prod_{i=1}^{k} \frac{[L]}{K_d^i}} \left(\sum_{k=1}^{n} k \prod_{i=1}^{k} \frac{[L]}{K_d^i} \right)$$

numerically or analytically if possible. From there it is trivial to calculate the concentration of the free protein as

$$[P] = \frac{[P]_{tot}}{1 + \sum_{k=1}^{n} \prod_{i=1}^{k} \frac{[L]}{K_d^i}}$$

and the concentrations of the various bound states as

$$[PL_k] = \frac{[PL_{k-1}][L]}{K_d^k}$$

The dissociation constants are finally determined by optimizing their values to minimize the expression

$$\chi^2 = \sum_i \frac{\left(s_i^{exp} - s_i^{calc}\left([P]_{tot,i}, [L]_{tot,i}, s_P, s_{PL}, ..., s_{PL_n}, K_d^1, K_d^2, ..., K_d^n \right) \right)^2}{\sigma_i^2}$$

where s_i^{exp} and s_i^{calc} are the experimental and calculated signals for data point i, respectively, and σ_i is the uncertainty. The Levenberg–Marquardt method[13,14] is the minimization algorithm used in PLIS.

2.1 | Titrations in the presence of a competitor probe

A way to reduce the number of fitting parameters is to instead monitor the signal of a competitor (C) to the protein that only binds one ligand. This removes all parameters corresponding to the signals of all partially bound protein states at the expense of the introduction of only one additional parameter, the dissociation constant of the competitor (K_d^C).

$$CL \rightleftarrows C + L, \quad K_d^C = \frac{[C][L]}{[CL]}$$

Furthermore, the dissociation constant of the competitor can be measured in a separate experiment. An additional

benefit is that the method does not depend on the amino acid composition of the protein, for example, presence of aromatic residues. The equation that must be solved to obtain the concentration of the free ligand is modified to

$$[L] = [L]_{tot} - \frac{[C]_{tot}[L]}{K_d^C + [L]} - \frac{[P]_{tot}}{1 + \sum_{i=1}^{n} i \prod_{j=1}^{i} \frac{[L]}{K_d^i}} \left(\sum_{i=1}^{n} i \prod_{j=1}^{i} \frac{[L]}{K_d^i} \right)$$

whereas the expressions for the concentrations of the various states of the protein remain unchanged.

2.2 | NMR relaxation dispersion experiments

An attractive alternative to titrations is Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiments,[15–17] where the effective transverse relaxation rate (R_2^{eff}) is measured as a function of the repetition rate of refocusing pulses (ν_{cpmg}). In the context of ligand binding, the experiment reports on kinetic, thermodynamical and structural aspects of the binding reaction. The kinetics and thermodynamics are described by the pseudo-first order rate constants $[L]k_{on}$ and k_{off} or alternatively the exchange rate constant $k_{ex} = [L]k_{on} + k_{off}$ and the population of the bound state $p_{PL} = [PL]/[P]_{tot}$. The data is fitted to the (corrected) Carver-Richards equations[18,19] or to numerical solutions to the Bloch-McConnell equations.[20] The dissociation constant for one site binding is calculated as

$$K_d = (1 - p_{PL}) \left(\frac{[L]_{tot}}{p_{PL}} - [P]_{tot} \right)$$

Alternatively, if the experiments are performed at more than one concentration of ligand and/or protein, the dissociation constant and the total concentrations of ligand and protein can be included as fitting parameters, using the following expression for the free ligand[21]

$$[L] = \frac{1}{2} \left\{ -K_d - [P]_{tot} + [L]_{tot} + \sqrt{\left(K_d + [P]_{tot} - [L]_{tot} \right)^2 + 4[L]_{tot} K_d} \right\}$$

3 | RESULTS AND DISCUSSION

3.1 | The PLIS interface

The PLIS interface is divided into three areas (Figure 1). The data is displayed in a table in the top left corner and the results of fits are shown in the bottom left corner. The right side of the main window is used to display graphs of experimental and fitted data.

3.2 | Import of data

PLIS facilitates three modes. "*Titration (normal)*" implies that the signal is concentration dependent, for example, fluorescence or absorption spectroscopy. "*Titration (NMR)*" is used for methods where the signal does not depend on concentration, for example, the NMR chemical shift in the fast exchange limit. "*CPMG RD*" is used for relaxation dispersion data.

Data sets can be imported from text files consisting of three columns separated by white space. For titration data, the columns correspond to the measured signal, the ligand concentration and the sample volume, respectively. Since accurate measurement errors rarely are available in a single titration series, it is not possible to enter these here. Instead, a common error of unity is assumed, implying that χ^2 is scaled inappropriately if the true errors are unknown. It is however possible to average several series in PLIS and thus obtain uncertainties from the standard deviations of the averaged data points. For CPMG relaxation dispersion data, it is common practice to include duplicate data points for error estimation. Thus, here the three columns are the repetition rate of the refocusing pulses, the effective transverse relaxation rate and its uncertainty. In addition of importing data sets, empty data set can be created where content is added manually or pasted from a spread sheet.

3.3 | Curve fitting

Currently, the software supports binding models with one to four binding sites in the absence or presence of a competitor for titration data and one binding site for CPMG relaxation dispersion data (Table 1). Data is fitted by first choosing the desired model and then optionally entering initial estimates for the parameters and deciding whether to keep certain parameters fixed rather than optimized. If the fit converges, a graph corresponding to the fitted data is displayed and the dissociation constants and their estimated uncertainties are shown. The default method for estimating uncertainties in the model parameters is the jackknife.[22] Monte Carlo simulations[23] can be used if errors in the data points are known.

A drawback of Levenberg–Marquardt minimizations is that they may converge to local rather than global minima. Fits that are unsatisfactory for this reason are often recognized by visual inspection. It is then possible to

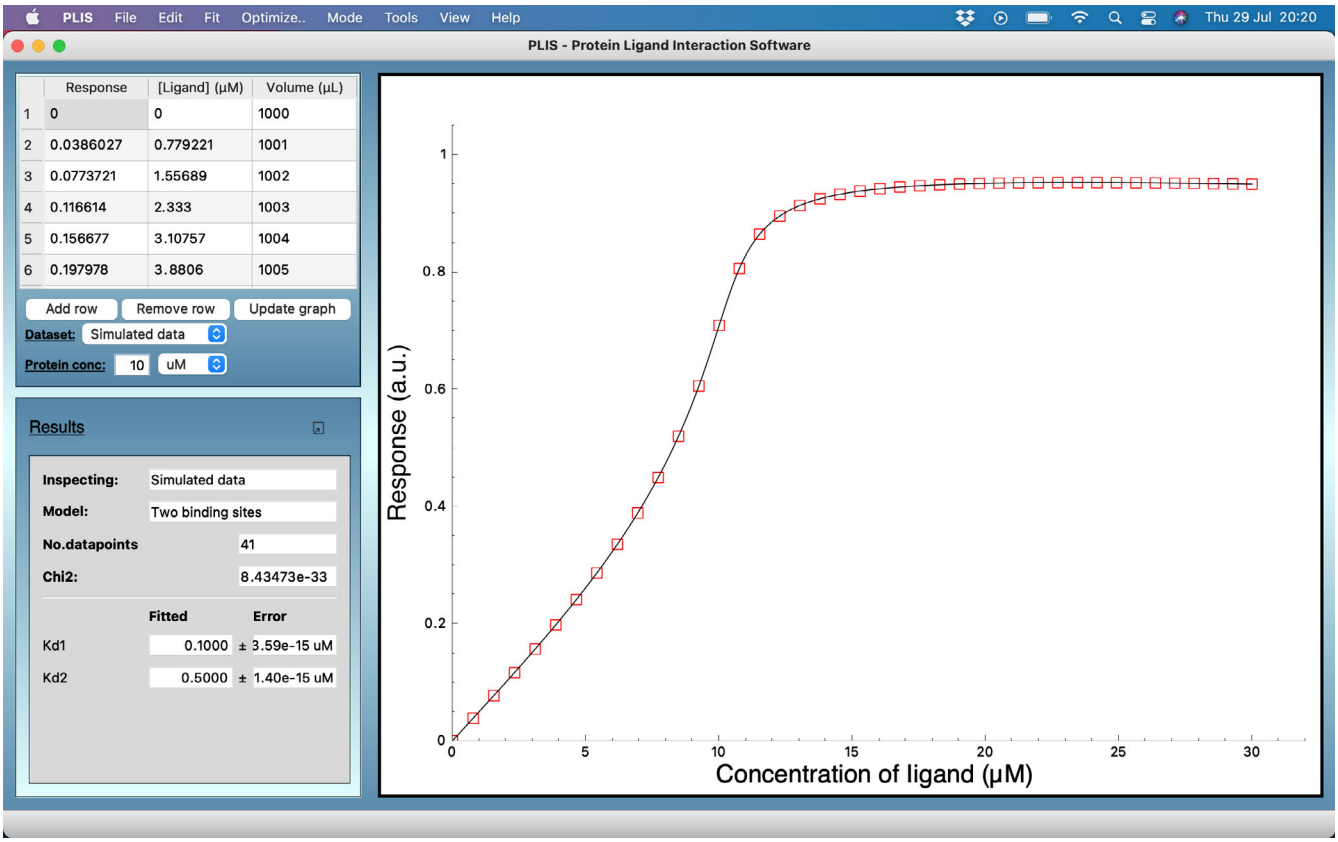

FIGURE 1 The PLIS interface. In the area at the top left, the data is represented as an editable table. Below, the results of fits are shown. The main area is dedicated to graphical representations of experimental and fitted data

TABLE 1 Binding models in PLIS

No. binding sites	Optimized parameters	
	Signal of protein monitored	**Signal of competitor monitored**
1	K_d, $s(P)$, $s(PL)$[a]	K_d, K_d^C, $s(C)$, $s(CL)$
2	K_d^1, K_d^2, $s(P)$, $s(PL)$, $s(PL_2)$	K_d^1, K_d^2, K_d^C, $s(C)$, $s(CL)$
3	K_d^1, K_d^2, K_d^3, $s(P)$, $s(PL)$, $s(PL_2)$, $s(PL_3)$	K_d^1, K_d^2, K_d^3, K_d^C, $s(C)$, $s(CL)$
4	K_d^1, K_d^2, K_d^3, K_d^4, $s(P)$, $s(PL)$, $s(PL_2)$, $s(PL_3)$, $s(PL_4)$	K_d^1, K_d^2, K_d^3, K_d^4, K_d^C, $s(C)$, $s(CL)$

[a]For titration data. For CPMG relaxation dispersions, the optimized parameters are $R_{2,0}$, k_{ex}, p_{PL}, and $\Delta\omega$, from which K_d is calculated.

optimize the fit using the current parameters as guides for initial estimates. Optimization can be done iteratively, and the user decides, aided by the appearance of the fitted curve and χ^2, whether to accept or reject the results of the optimization. It is also possible to precede the curve fitting by a grid search to avoid local minima.

As an example of data fitting in PLIS, we show the results of titration of Ca^{2+} into the C-terminal lobe of calmodulin (Tr2C) in the presence of the competitor O,O′-bis(2-aminophenyl)ethyleneglycol-N,N,N′,N′-tetraacetic acid (BAPTA) in Figure 2. The data was fitted to a model with two Ca^{2+} binding sites for Tr2C and one for the probe BAPTA. The dissociation constants of 0.11 and 1.9 μM are in good agreement with previously reported values, considering differences in sample conditions.[24] However, reliable results are also dependent on accurately determined concentrations of protein and competitor. To keep the fitting simple, we have not included the possibility to include concentrations as adjustable parameters during fitting. We find it just as easy and more robust to manually tweak the protein concentration and then refit the data.

As a second example of data fitting, as well as a demonstration of the data simulation tool of PLIS, we

simulated CPMG relaxation dispersion data for the binding reaction $P + L \rightleftharpoons PL$ and subsequently fitted the data (Figure 3). Since the data was generated using a dissociation constant of 20 μM and realistic concentrations of protein and ligand, the fitted value of 19.9 ± 0.5 μM demonstrates that accurate dissociation constants can be determined from relaxation dispersion data in favorable cases. Among the factors that may occlude an accurate analysis are a narrow sweet spot of the exchange rate constant, ligand independent conformational exchange, uncertainty in the ligand concentration and poor signal-to-noise ratios. As a partial remedy, dispersions may be recorded at several magnetic fields and/or ligand concentrations and fitted together.[21,25]

To illustrate the effect of experimental noise on reliable analyses, we compared fitting two synthetic data sets with different noise levels (see Figure 4 for details). As expected, the fitted dissociation constants are equal to the ones used to generate the data if no noise is added. For the simulated data set with added noise, the fitted

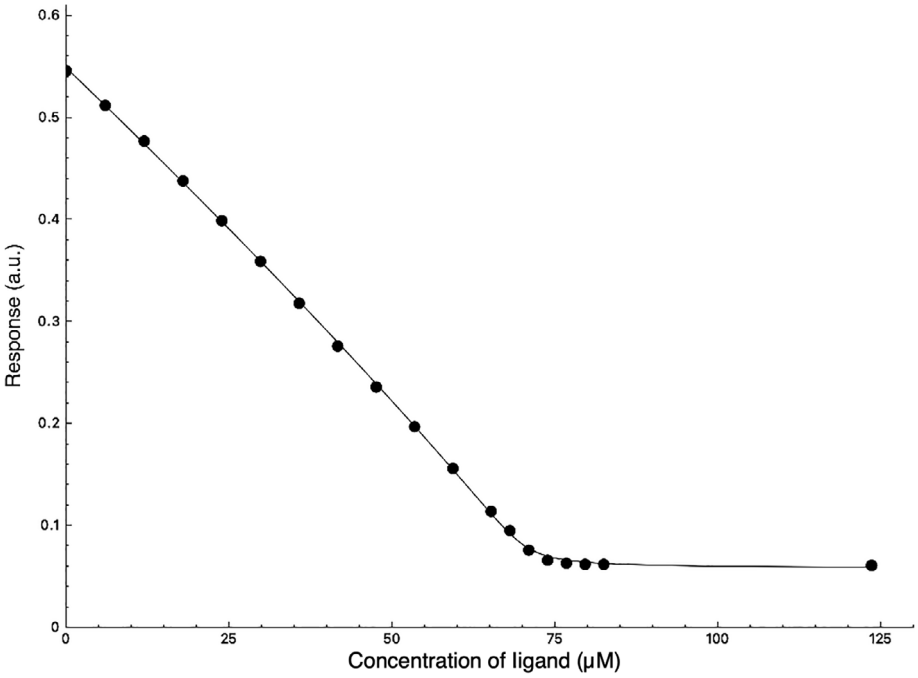

FIGURE 2 The competitor O,O′-bis (2-aminophenyl)ethyleneglycol-N,N,N′, N′-tetraacetic acid (BAPTA) was added to the C-terminal lobe of calmodulin in 20 mM HEPES, pH = 7.5. The concentrations of protein as well as BAPTA was 35 μM. $CaCl_2$ was titrated in and the absorbance at 254 nm was measured

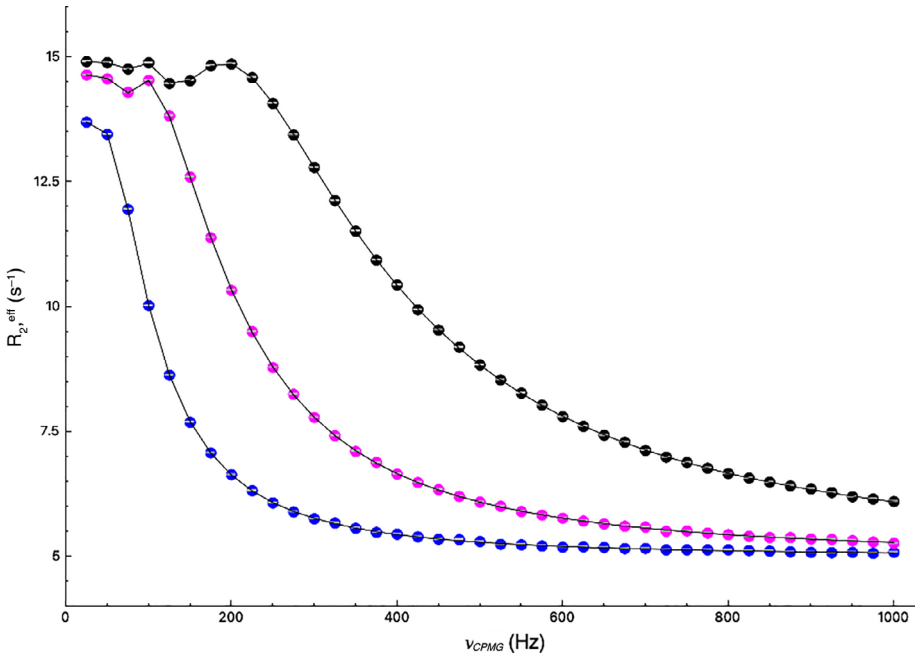

FIGURE 3 Relaxation dispersions for simulated data. The data was generated using a total protein concentration of 500 μM, a total ligand concentration of 26.1 μM and a dissociation constant of 20 μM, corresponding to a population of the bound state of 5.0%. The global exchange rate was set to 200 s^{-1} and $\Delta\omega$ for the three datasets were set to 600 (blue), 1,200 (magenta), 2,400 s^{-1} (black) corresponding to 1.2, 2.4, and 4.8 ppm, respectively, for ^{15}N at 18.8 T. The solid lines correspond to global fits to the Carver Richards equations[18,19]

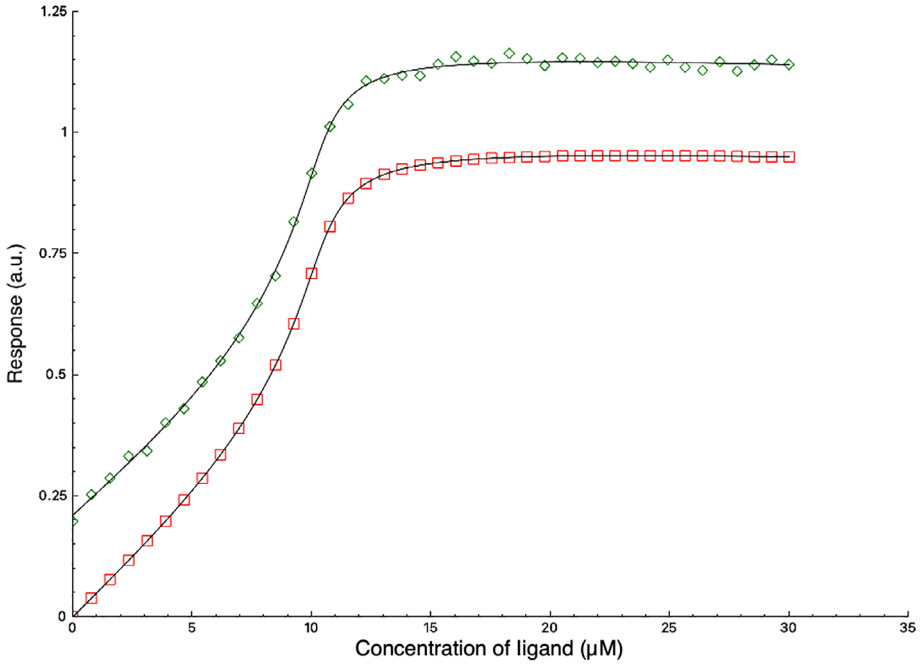

FIGURE 4 The simulation tool of PLIS. Two data sets for two-site binding with dissociation constants of 0.10 and 0.50 μM and an initial protein concentration of 10 μM were generated. For the first data set (red symbols), no noise was added and for the second data set (green symbols), the RMSD of the noise was set to 0.01 a.u. The solid lines correspond to fits to a two-site model

dissociation constants were 0.10 ± 0.05 and 0.41 ± 0.05 μM, respectively, showing that this noise level is incompatible with determination of accurate dissociation constants. Another application of the simulation tool is to learn how different choices of parameters affect the appearance of data, which is useful for analyzing experimental data.

4 | CONCLUDING REMARKS

In conclusion, we have developed PLIS, a tool for the determination of dissociation constants from titrations probed by various methods, in the absence or presence of a competitor, or from NMR relaxation dispersion experiments at one or several ligand concentrations. This functionality is not found in competing software. Another useful feature of PLIS is the possibility to simulate data for various binding models, with or without noise. PLIS is deliberately designed to be intuitive rather than overwhelming. The software is publicly available and open-source freeware that can be extended or modified under the terms of GNU GPL.

5 | DISTRIBUTED FILES

Binaries for 64-bit versions of Windows and macOS as well as source code can be downloaded at https://github.com/PINT-NMR/PLIS/. PLIS is written in C++ and is an open-source software released under the terms of the GNU General Public License (http://www.gnu.org/licenses/). PLIS was compiled in Qt Creator 4.14.1 using libraries from Qt 5.15.2 open source license (https://www.qt.io/download-open-source#section-2) and QCustomPlot (http://www.qcustomplot.com/).

ACKNOWLEDGMENTS

We thank Dr. Cecilia Andresen and Hugo Gustavsson for helpful suggestions and Felicia Larsson for designing the PLIS logotype.

AUTHOR CONTRIBUTIONS

Emil Björklund: Conceptualization (equal); software (lead). **Anna du Rietz:** Investigation (equal); methodology (equal). **Patrik Lundström:** Conceptualization (equal); investigation (equal); methodology (equal); project administration (equal); software (equal); supervision (equal).

ORCID

Patrik Lundström https://orcid.org/0000-0002-4874-5381

REFERENCES

1. Alberts B, Johnson A, Lewis J, et al. Molecular biology of the cell. 6th ed. New York, NY: Garland Science Taylor & Francis Group, 2015.
2. Hammes GG, Yu-Chu Chang Y-C, Oas TG. Conformational selection or induced fit: A flux description of reaction mechanism. Proc Natl Acad Sci U S A. 2009;106:13737–13741.
3. Hassell AM, An G, Bledsoe RK, et al. Crystallization of protein–ligand complexes. Acta Cryst D. 2005;63:72–79.

4. Mocz G, Ross JA. Fluorescence techniques in analysis of protein-ligand interactions. Methods Mol Biol. 2013;1008:169–210.
5. Siligardi G, Hussain R, Patching SG, Phillips-Jones MK. Ligand- and drug-binding studies of membrane proteins revealed through circular dichroism spectroscopy. Biochim Biophys Acta. 2014;1838:34–42.
6. Rodger A, Marrington R, Roper D, Windsor S. Circular dichroism spectroscopy for the study of protein-ligand interactions. Methods Mol Biol. 2005;305:343–364.
7. Patching SG. Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. Biochim Biophys Acta. 2014;1838:43–55.
8. Damian L. Isothermal titration calorimetry for studying protein-ligand interactions. Methods Mol Biol. 2013;1008:103–118.
9. Maity S, Ravi Kumar Gundampati RK, Kumar TKS. NMR methods to characterize protein-ligand interactions. Natural Product Commun. 2013;14:1–17.
10. Krämer SD, Wöhrle J, Rath C, Roth G. Anabel: An online tool for the real-time kinetic analysis of binding events. Bioinform Biol Insights. 2019;13:1–10.
11. Duvvuri H, Wheeler LC, Harms MJ. pytc: open source python software for global analyses of isothermal titration calorimetry data. Biochemistry. 2018;57:2578–2583.
12. André I, Linse S. Measurement of Ca^{2+}-binding constants of proteins and presentation of the CaLigator software. Anal Biochem. 2002;305:195–205.
13. Levenberg K. A method for the solution of certain non-linear problems in least squares. Quart Appl Math. 1944;2:164–168.
14. Marquardt D. An algorithm for least-squares estimation of nonlinear parameters. SIAM J Appl Math. 1963;1963(11):431–441.
15. Carr HY, Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. Phys Rev. 1954;94:630–638.
16. Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. Rev Sci Instrum. 1958;29:688–669.
17. Allerhand A, Thiele E. Analysis of Carr-Purcell spin-echo NMR experiments on multiple-spin systems. II. The effect of chemical exchange. J Chem Phys. 1966;45:902–916.
18. Carver JP, Richards RE. General 2-site solution for chemical exchange produced dependence of T2 upon Carr-Purcell pulse separation. J Magn Reason. 1972;6:89–105.
19. Davis DG, Perlman ME, London RE. Direct measurements of the dissociation-rate constant for inhibitor-enzyme complexes via the $T_{1\rho}$ and T_2 (CPMG) methods. J Magn Reson Ser B. 1994;104:266–275.
20. McConnell HM. Reaction rates by nuclear magnetic resonance. J Chem Phys. 1958;28:430–431.
21. Furukawa A, Konuma T, Saeko Yanaka S, Sugase K. Quantitative analysis of protein–ligand interactions by NMR. Prog Nucl Magn Reson Spectrosc. 2016;96:47–57.
22. Mosteller F, Tukey JW. Data analysis and regression. Reading: Addison-Wesley Publishing Company, 1977.
23. Press WH, Flannery BP, Teukolsky SA, Vetterling WT. Numerical recipes in C: The art of scientific computing. Cambridge: University Press, 1988.
24. Linse S, Helmersson A, Forsén S. Calcium binding to calmodulin and its globular domains. J Biol Chem. 1991;266:8050–8054.
25. Sugase K, Dyson HJ, Wright PE. Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature. 2007;447:1021–1025.

How to cite this article: Björklund E, du Rietz A, Lundström P. Analysis of protein–ligand interactions from titrations and nuclear magnetic resonance relaxation dispersions. Protein Science. 2022;31:301–7. https://doi.org/10.1002/pro.4240