

Generating reliable kinetic data for protein-ligand interactions

James Delproposto, Research Associate, University of Michigan

Introduction

The Community Structure Activity Resource (CSAR, www.csar-dock.org) group is developing a database of high quality protein-ligand structures and the corresponding binding affinities. The data will be provided from in-house experiments and community collaborations. The proteins are generally well-studied structures that have been targeted in drug discovery projects. Current projects include CDK2, LpxC, and urokinase. The drug-like ligands for each of these targets consist of several series of compounds with a wide range of affinities.

The Octet® RED system has dramatically increased the data we can produce in a short time, turning it into the workhorse of the CSAR project. *Note: The Octet RED system referred to in the application note has been replaced by the Octet RED96 system. To learn more about the Octet platform, visit www.fortebio.com.*

Approach

Current docking and scoring models rely on experimental data, which is typically incomplete: crystal structures of protein-ligand complex may exist, but affinity data is not available. Affinity data for a series of compounds might be published, but the crystal data may be poor quality or incomplete. Variations between experimental conditions can create differences between published data. Without high-quality datasets, training these models becomes difficult if not impossible. Our goal is to create complete high-quality data sets to assist computational scientists with their goals.

To populate the affinity database, three label-free methods are used: the Octet RED system, a thermal denaturation assay, and isothermal titration calorimetry (ITC). These produce both complementary and overlapping data. The methods use dissimilar technologies to independently produce affinity values (K_D) with exactly the same compound, protein, and buffer. This helps to reduce experimental error and gives modelers more confidence in the accuracy of the data.

Methods

The Octet RED system determines rate constants, k_a and k_d , which are used to calculate a K_D , defined as k_d/k_a . We determine K_D for six different compounds in two hours using six concentrations for each compound and 0.1 mg of protein. After scouting runs to determine optimal concentrations of compound, pre-made compound plates allow us to complete up to four runs in a single day. This throughput has dramatically increased the data we can produce in a short time, turning the Octet RED system into the workhorse of the CSAR project.

The thermal denaturation assay is performed on the Thermofluor® platform. It is a high-throughput method that detects changes in fluorescence as the protein is heated and melts. Ligand added to the protein stabilizes it and increases the melting temperature. The difference between the melting temperatures of the protein and the protein-ligand complexes allows calculation of an affinity value. A run takes about two hours, and uses a comparable amount of protein and compound to the Octet RED system. The K_D values, however, must be calculated using the Van't Hoff equation. Some of the variables used for the equation are difficult to obtain and must be estimated, therefore the affinity value calculated from the melting temperature difference is an estimation at best. The dye used to create the fluorescent signal also has the potential to interact with the protein in unforeseen ways. An article recently published indicated that the 1,8 ANS dye binds allosterically to CDK2, causing a conformation change. This unexpected challenge invalidates our current Thermofluor data and requires us to repeat experiments with a new dye.

The third method we use is ITC. It has a long history of being a reliable method and gives affinity data very similar to that produced by the Octet RED system. ITC has a number of advantages and disadvantages. It is the slowest of the three methods, but unique in that it can provide accurate enthalpy of binding values. Each run uses several milligrams of protein that must be saturated with 2 to 3 times the molar equivalent of compound. In addition, it can measure only a very limited range of affinities. For the CDK2 project, the affinities of only

Compound	Octet RED system (average)			ITC (average)		ΔG , Calculated average	
	K_D (M)	k_a (1/Ms)	k_d (1/s)	K_D (M)	ΔH (kcal/mol)	ITC	Octet RED system
CS1	3.32E-05	3.57E+04	1.09E+00		-7.52		-6.11
CS2	1.84E-06	2.22E+05	3.81E-01	4.23E-06	-8.76	-7.33	
CS3	8.08E-07	3.35E+05	2.59E-01		-17.70		-8.31
CS4	2.12E-05	7.19E+04	1.29E+00	6.03E-06	-7.95	-7.12	
CS9	5.83E-05	2.35E+04	2.58E-01		-12.70		-6.50
CS10	9.10E-07	7.97E+04	6.66E-02				-8.24
CS11	2.80E-07	2.81E+05	8.40E-02	1.41E-06	-11.75	-7.98	
CS13	9.65E-07	6.95E+04	3.94E-02				-8.21
CS14	1.43E-06	6.43E+04	7.70E-02				-7.97
CS15	5.57E-07	1.13E+05	3.67E-02				-9.00
CS16	7.73E-08	2.94E+05	1.69E-02				-9.70
CS17	5.39E-08	1.22E+05	3.40E-03				-10.16
CS18	9.61E-07	1.00E+05	8.23E-02				-8.21
CS19	2.29E-07	1.75E+05	2.01E-02				-9.06
CS20	7.77E-07	8.59E+04	5.62E-02				-8.33
CS241	2.87E-07	175033.3	4.97E-02				-8.92
CS242	5.28E-07	3.21E+04	1.73E-02		-6.57		-8.56
CS244	1.79E-07	9.73E+04	1.28E-02				-9.20
CS245	2.29E-07	9.72E+04	8.20E-03				-9.06
CS246	5.26E-07	1.14E+05	5.92E-02	1.94E-07	-14.99	-9.15	
CS247	2.26E-08	9.18E+04	2.30E-03				-10.40
CS248	1.06E-06	7.84E+04	7.97E-02				-8.15
CS260	9.86E-08	1.32E+05	1.33E-02	1.79E-07	-10.70	-9.20	
CS261	2.51E-07	3.97E+04	1.00E-02	1.28E-06	-3.39	-8.04	
CS262	4.73E-08	6.49E+04	2.61E-03				-10.00

Table 1: Data obtained for CDK2 binding to compounds of molecular weight ranging from 193 Da to 509 Da. All experiments were performed at T = 298K.

six of 27 compounds could be used, due to issues with ligand solubility and affinity range. The enthalpy alone was determined for an additional four compounds. This data can be combined with the affinity data from the Octet RED system to determine the Gibbs free energy and entropy of binding.

CONCLUSION

Consistent results are the most important aspect of using the Octet RED system. Data must be consistent between runs as well as between different methods. The Octet RED system data have been very consistent with ITC. In cases where it has been possible to use all three methods on a compound, the data produced by each method are similar (as a general rule, an affinity within 3-fold of one determined via another method is considered identical) and this is shown in Table 1. Some of the inconsistencies in our data may be due to compound solubility issues. The Octet RED system requires far lower

ligand concentrations than ITC, and thus is less affected by compound solubility issues. Thermal denaturation assays are often used for compounds with low solubility, but still require saturation of the protein. The Octet RED system can produce usable data at concentrations 10–100 fold lower than ITC and Thermofluor, helping ensure accurate compound concentration in solution.

The results presented in this application note were included in the first data set for the CSAR database. Additional data sets for CDK2-Cyclin A and Lpxc can be found at www.csardock.org. CSAR will continue to create high-quality protein-ligand structures and affinity datasets to aid the computational chemistry community.

Thermofluor is a trademark of Johnson and Johnson. The ITC measurements were made with the Nano ITC Low Volume system from TA Instruments.



www.fortebio.com

FortéBio
47661 Fremont Boulevard
Fremont, CA 94538
888.OCTET-75 or 650.322.1360
fortebio.info@moldev.com

FortéBio Analytics (Shanghai) Co., Ltd.
No. 88 Shang Ke Road
Zhangjiang Hi-tech Park
Shanghai, China 201210
salesops.china@moldev.com

Molecular Devices (UK) Ltd.
660-665 Eskdale
Winnersh Triangle
Wokingham, Berkshire
RG41 5TS, United Kingdom
+44 118 944 8000
uk@moldev.com

Molecular Devices (Germany) GmbH
Bismarckring 39
88400 Biberach an der Riss
Germany
+ 00800 665 32860