Applications of biophysical methods in small-molecule modulators targeting protein function

Wang, Wenjia

DOI:
10.33612/diss.156123664

Chapter 1

Introduction

Drug discovery is known as a complicated yet essential process to discover new classes of medications based on the knowledge of biological targets [1]. Most of the protein targets that are selected play a crucial role in metabolism or signalling pathways of human diseases or pathology [2]. In this context, their definition as "druggable" means their structure is capable of binding to small molecules [3], peptides [4] or antibodies. Here, we mainly focus our attention on small molecule inhibitors. Until now, discovered small molecules could be classified as agonists, antagonists, inverse agonists, modulators, enzyme activators, inhibitors, ion channel opener or blockers based on their modulation of the function of their targeted proteins [5]. Structure-based drug design (SBDD) is one of the powerful methods widely used, that takes advantage of the knowledge and basic understanding of the three-dimension structure of a protein, or protein complex to design, synthesise small molecule anchor and predict or determine their binding affinity to the targets [6]. With the development of computational methods, SBDD has been accelerated by reducing the number of iterations required, and increasing numbers of novel structures are now obtained from X-ray crystallography, nuclear magnetic resonance (NMR) and Cryo-electron microscopy (Cryo-EM) [7].

The idea to make use of the protein 3D atomic structure to guide the design of new molecules is derived from the 1960s when the first protein was solved by X-ray crystallography [8]. In the late 1990s, fragment-based drug design (FBDD) evolved based on the development of SBDD. The main principle underlying FBDD is the efficient screening of chemical group using small libraries of low-molecular-weight sample fragments [9]. While

fragments typically bind with a low potency, they can form productive interactions with the protein target, and thereby provide good starting points for further development.

The process of FBDD is interactive and often proceeds through multiple cycles of optimisation before a lead/hit is generated. In general, the first cycle includes target gene cloning, fusion or native protein purification from prokaryotic expression *E.coli* or eukaryotic expression systems like yeast, insect or mammalian cells, and structure determination of the target protein. Then using computer algorithms, fragments of compounds from a library are positioned into a selected region of the protein. These fragments are scored and ranked based on their interaction with the target site. In total, three main principles can be used in scoring. Firstly, virtual screening that small fragments are randomly docked into the region of interest *in silico* and scored based on predicted steric and electrostatic interactions with the surrounding residues. A second category is known as *de novo* design, building up against the constraints of the pocket by assembling small standard fragment pieces [10-12], such as benzene rings, carbonyl groups, amino groups, and so on, which are positioned in the site, scored, then linked *in silico*. The last but not least method is to optimise the known ligands by evaluating potential analogues within the binding cavity [13].

These binding sites are known as "hot spots" of the protein, which relies upon whether the protein can accommodate drug-sized molecules, generate appropriate hydrophobic surfaces and hydrogen bonding sites, and contribute essential interactions between surrounding residues and small molecules [14]. After that, fragments with best scores are tested via an *in vitro* screening system. For instance, successfully fragment screens on checkpoint kinase II have been carried out using *in vitro* biochemical assays [15]. In general, an *in vitro* screening system is a combination of multiple assays including HTS (high-throughput screening), functional assays, ligand selectivity, safety pharmacology screening and biophysical assays. The ability to identify the biological target is heavily dependent on the choice of detection systems and equipment employed for a given assay. As a structure biology lab, use multiple biophysical methods such as NMR, surface plasmon resonance (SPR), thermal shift assay (TSA), X-ray crystallography, Cryo-EM, together with more recent ones including microscale thermophoresis (MST) and mass spectrometry (MS) [16,17] to identify fragment hits.

The second cycle consists of multiple rounds of structure determination of the target proteins in complex with the promising hits from the first cycle, ideally one with at least micro-molar inhibition *in vitro*. It reveals the actual interaction pose of a moiety of the fragments that can be optimised to increase potency. Additional cycles include synthesis of the optimised lead, structure determination of the new complex. After several cycles, the optimised compounds usually show significant improvement in binding and specificity for the target.

In Chapter 2, a review was written in the first year of my PhD. At that moment, I was unclear about the importance and meaning of structure-based drug design (SBDD). IL17A is an essential protein in autoimmune processes; its aberrant expression results in multiple abnormal symptoms shown in autoimmune-related diseases, for instance, rheumatoid arthritis, asthma. Until now, the most successful therapy mAbs called Cosentyx (Secukinumab) was provided by Novartis. In the aspect of SBDD, linear peptide inhibitor HIP (5HHX) and peptide 18-1(5VB9) were suggested to inhibit IL17A/IL17RA interaction by reducing IL8 release in HEK293T. In 2016, researchers from Pfizer reported three small-molecules, including two macrocycles as IL17A/17RA inhibitors. They were shown to occupy the wide full central cavity of IL17A homo-dimer and their binding resulting significant outward movement of N-terminus of IL17A, changing the protein-protein interaction(PPI) surface of IL17A/IL17RA. In the deposited structures, the bound HIP at N-terminus was implied to contribute more possibility for ligand binding between two IL17A monomers. Based on their contributions, we took the use of IL17A/63Q (5HI5) structure to design a fundamental scaffold as detailed later in my thesis.

In Chapter 5, we selected IL17A as an example target to design a scaffold based on the structure determination of apo and a known artificial macrocycle ligand-protein complex (IL17A/63Q). After analysis of the binding mode, four essential hydrogen bonds were revealed to forming a conserved H-bonding pattern, which is used as pharmacophores to screen different MCR scaffolds. Subsequently, once the compound synthesised with a selected scaffold, we made use of a combination of two biophysical methods to explore whether the concept works: initial differential scanning calorimetry (DSF) screening cross-validated by a thermophoresis (MST) assay. We selected a potential compound MD46 with measured binding affinity as 49.70nM, 10-fold better than that of 63Q in

our assay. To facilitate chemical optimisation, we visualised MD46 stereoisomers in the IL17A apo structure by molecular docking. After three rounds of random docking, three binding poses showed a conserved hydrogen bond with Trp67B as we predicted and another hydrogen bond with Glu95A instead of Leu97A. Further experiments are on-going to obtain the crystal structure of MD46 single isomers to further confirmed the *in silico* structure-based scaffold design is a useful concept in new class binder discovery.

The efforts to design and screen provided essential information for us to successfully discover a new series of compounds with an experimentally determined binding affinity in the nM range in Chapter 6. We found a new class of ligands acting as IL17A binders, which could be considered as a good starting point to investigate anti-IL17A small-molecule antagonists. Two complex structures of them have been determined by X-ray crystallography. Both ligands were visible close to the central homodimer pocket, resulting in a stabilised and extended α-helix that is known to interact with IL17RA. Additionally, they partially overlapped with the published inhibition peptide 18-1 (5H5Y) from Novartis, which has been implicated as a peptide inhibitor affecting IL17A/IL17RA interaction. During preparing this manuscript, I continued to work on optimizing these IL17A ligands based on the structure-activity relationship (SAR) to facility chemical optimisation. Until now, more than 40 compounds from this series have shown various binding affinity after optimisation from the initial hit. The first hit of this series has demonstrated inhibition efficacy in IL17A/IL17RA based cell assay from our collaborators. In all, it is a useful application case to illustrate the general flowchart of structure-based drug discovery. Unfortunately, more experiments will be needed and, after discussion with my supervisors Prof. Dr. Alexander Dömling and Prof. Dr. M.R. (Matthew) Groves, we decided not to show the chemical structures of our lead compounds, which is my main project of PhD in the thesis until appropriate steps have been made to protect this discovery. It will be published in the future.

In Chapter 3, we described a study on bovine Carbonic anhydrase II (bCAII), which is the first crystal obtained during my PhD and provided me an opportunity to learn X-ray structure determination techniques. It started purification from the commercial resource. After optimising purification and crystallisation condition, apo bCAII crystal diffraction datasets were collected to 1.8Å. To understand how copper affects probe binding and

labelling, we designed a set of ligand-directed diazo transfer probes. Seven ligands were soaked into apo bCAII crystals, and their conformation was determined via X-ray crystallography. All of them revealed a tight and clear binding of benzenesulfonamide ligand moiety of the probes via zinc coordination to the active-site of bCAII. All the ligands were visible in the structure with a clear density for the head group. At the same time, their tails are oriented towards the hydrophobic face of the active site and thereby it is pointing away from the modified lysine residues, the compound was shown to modify lysine residues Lys8, Lys17, Lys168 and potentially also Lys166 from mass spectrometry. Here, we decided to analyse structures with linker length of 0, 1 and 2, giving a strong density signal of tails. We proposed that the N-terminus of bCAII might provide a flexible second lysine modification site. Together, this work shows the use of structural biology, especially X-ray crystallography, to develop small molecule interactors.

In Chapter 4, we analysed the structure of pyridoxal kinase from Plasmodium falciparum (*Pf* Pdxk) complex with AMP-PNP and PL. After comparing *Pf* Pdxk structure with multiple homologues, we found a specific repeat motif region XNXH within *Pf* Pdxk. Although we did not find clear electronic density signal of the motif, its existence did not affect the structural integrity of *Pf* Pdxk. Also, we showed in this study that the repeat sequence might play an essential role in malaria cell cycle based on the discovery of another motif FIxxIIxL (Eh1) on the N-terminus of XMXH motif. This hypothesis is based on the report that the Eh1 motif may recruit WD40 proteins and the E3 ligase system to drive ubiquitin-mediated degradation of *Pf* PdxK. Also, we generated models by molecular docking to explain how *Pf* Pdxk selectively phosphorylated PT3, PT5 and PHME into a potent active anti-malaria drug. Our structures proposed an overlapping of phosphorylating 5-pyridoxal group of the three substrates but suggested a potential steric clash of pyridoxal group of PT3 with AMP-PNP. This might impede the phosphorylation process to form active PPT3. Even though it should be noted that *in silico* modelling to predict the binding mode of a potential ligand cannot provide robust and accurate results, it might offer a feasible strategy in optimising function substituents of lead pro-drug towards to modulator or inhibitor, and further facility experimental validation assay design, in the lack of complex structure. Together, this work presented the combination of X-ray crystallography, and molecular docking which provide an insight of anti-malaria pro-drug design and development.

In Chapter 7, I describe the work performed using Microscale thermophoresis (MST), a biophysical technique to quantify the interactions between biomolecules (protein-protein or protein-small molecule). It is highly sensitive to determine the potential binding affinity of 3-FL, LNnT, and LDFT to the His-TNFR1 receptor. A ligand-dependent binding effect of LNnT with His-TNFR1 was observed with a *Kd* of 900nM, which supported a potential hypothesis hMOs especial LNnT has its anti-inflammatory effects through a direct structure-function relationship with TNFR1.

In all, over the last three decades, the great improvements in biophysical methods produce a significant impact on not only enabling drug discovery on challenging targets, such as membrane protein but also providing the foundation for structure-based drug discovery (SBDD). These have driven significant improvements in the automated pipeline, speed, sensitivity and range of possible measurements, providing high-resolution mechanistic, kinetic, thermodynamic and structural information on compound-target interaction. Also, a multiplicity of biophysical methods combined with computational techniques can be used to define a variety of different modulators of protein from many aspects such as function-form labelling, protein-protein interaction (PPI), pro-drug design and so on. Therefore, it becomes a vital component of drug discovery platforms in many pharmaceutical companies and academic laboratories. Here we use this thesis to present a case of application of X-ray crystallography and biophysical methods in ligand discovery and development in our lab. Below is a summary of the methods that will be mentioned in the thesis.

X-ray crystallography is the most powerful, robust and routine method for providing a detailed atomic picture of a compound binding to its target. Structural information on ligand-target complexes, usually obtained by X-ray crystallography, is considered essential in SBDD, especial for FBDD, to provide insight into chemical optimisation.

Differential scanning fluorimetry (DSF), a widely used biophysical method to measure the temperature at which a unfold protein exposed hydrophobic surface through binding of a fluorescent probe. If a ligand stabilises or destabilises the protein, then there will be a change in the transition temperature (Tm) observed.

Microscale thermophoresis (MST) is a relatively new biophysical methodology that

monitors fluorescence in an infrared laser-heated spot. It is an equilibrium-based method that can detect ligand binding-induced change in thermophoretic mobility. It is affected by size, charge and hydration shell and can be used to estimate *Kd*.

Although these techniques provided precious information, whether the compound binds or not, there remain some challenges during the experiments we have experienced. Ideally, the compound must be soluble enough to determine its binding affinity constant, stable and unaggregated under the given experimental condition. This is hampered if the character of the compound cannot meet the demand of techniques, for instance, in X-ray crystallography a substantial percentage of the binding site must be occupied to identify binding and resolve the protein-ligand co-structure unambiguously. Thus, it leads to the partial occupation of the ligand even in high-resolution structure. That is one of the difficulties impeding the efficacy and accuracy of initial hit discovery.

Similarly, for a robust measurement, all of the technique requires the protein targets to be homogenous and well behaved at relatively high concentration. Therefore, more attention needs to be taken in the sample preparation and experiment design, the critical results analysis and well-interpretation. Overall, although we still face many uncertain challenges, the development of biophysical methods application will play a more and more crucial role in pharmaceutical research.

Reference

1. Liljefors, T., Krogsgaard-Larsen, P. Madsen, U. Textbook of drug design and discovery, CRC Press (2002).

2. Dixon, S.J. Stockwell, B.R. Identifying druggable disease-modifying gene products. Curr Opin Chem Biol 13, 549-55 (2009).

3. Yuan, Y., Pei, J. Lai, L. Binding site detection and druggability prediction of protein targets for structure-based drug design. Curr Pharm Des 19, 2326-33 (2013).

4. Fosgerau, K. Hoffmann, T. Peptide therapeutics: current status and future directions. Drug Discov Today 20, 122-8 (2015).

5. Imming, P., Sinning, C. Meyer, A. Drugs, their targets and the nature and number of drug targets. Nat Rev Drug Discov 5, 821-34 (2006).

6. Mobley, D.L. Let's get honest about sampling. Journal of Computer-Aided Molecular Design 26, 93-95 (2012).

7. Singh, J. et al. Successful shape-based virtual screening: the discovery of a potent inhibitor of the type I TGFbeta receptor kinase (TbetaRI). Bioorg Med Chem Lett 13, 4355-9 (2003).

8. Jaskolski, M., Dauter, Z. Wlodawer, A. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. Febs j 281, 3985-4009 (2014).

9. Erlanson, D.A., Fesik, S.W., Hubbard, R.E., Jahnke, W. Jhoti, H. Twenty years on: the impact of fragments on drug discovery. Nat Rev Drug Discov 15, 605-619 (2016).

10. Wang, R., Gao, Y. Lai, L. LigBuilder: A Multi-Purpose Program for Structure-Based Drug Design. Molecular modeling annual 6, 498-516 (2000).

11. Schneider, G. Fechner, U. Computer-based de novo design of drug-like molecules. Nat Rev Drug Discov 4, 649-63 (2005).

12. Jorgensen, W.L. The many roles of computation in drug discovery. Science 303, 1813-8 (2004).

13. Klebe, G. Recent developments in structure-based drug design. J Mol Med (Berl) 78, 269-81 (2000).

14. Leis, S., Schneider, S. Zacharias, M. In silico prediction of binding sites on proteins. Curr Med Chem 17, 1550-62 (2010).

15. Winter, A. et al. Developing Antagonists for the Met-HGF/SF Protein-Protein Interaction Using a Fragment-Based Approach. Mol Cancer Ther 15, 3-14 (2016).

16. Qin, S. et al. Multiple ligand detection and affinity measurement by ultrafiltration and mass spectrometry analysis applied to fragment mixture screening. Anal Chim Acta

886, 98-106 (2015).

17. Silva-Santisteban, M.C. et al. Fragment-based screening maps inhibitor interactions in the ATP-binding site of checkpoint kinase 2. PLoS One 8, e65689 (2013).