Editorial

Structure-based drug design: aiming for a perfect fit

Rob L.M. van Montfort1,2 and Paul Workman1


Correspondence: Rob L. M. van Montfort (rob.vanmontfort@icr.ac.uk) or Paul Workman (paul.workman@icr.ac.uk)

Knowledge of the three-dimensional structure of therapeutically relevant targets has informed drug discovery since the first protein structures were determined using X-ray crystallography in the 1950s and 1960s. In this editorial we provide a brief overview of the powerful impact of structure-based drug design (SBDD), which has its roots in computational and structural biology, with major contributions from both academia and industry. We describe advances in the application of SBDD for integral membrane protein targets that have traditionally proved very challenging. We emphasize the major progress made in fragment-based approaches for which success has been exemplified by over 30 clinical drug candidates and importantly three FDA-approved drugs in oncology. We summarize the articles in this issue that provide an excellent snapshot of the current state of the field of SBDD and fragment-based drug design and which offer key insights into exciting new developments, such as the X-ray free-electron laser technology, cryo-electron microscopy, open science approaches and targeted protein degradation. We stress the value of SBDD in the design of high-quality chemical tools that are used to interrogate biology and disease pathology, and to inform target validation. We emphasize the need to maintain the scientific rigour that has been traditionally associated with structural biology and extend this to other methods used in drug discovery. This is particularly important because the quality and robustness of any form of contributory data determines its usefulness in accelerating drug design, and therefore ultimately in providing patient benefit.

The idea to use the information from protein structure to guide the design of new therapeutic agents stems from the 1950s and 1960s when the first protein structures were elucidated by X-ray crystallography [1,2]. Notably, the pioneering work of John Kendrew and Max Perutz in solving the crystal structures of myoglobin and haemoglobin explained the oxygen-carrying/storing properties of these proteins and also shed light on the molecular basis of sickle cell anaemia and at the same time on potential treatments for this disease [3]. Similarly, the determination of the amino acid sequence of insulin by Fred Sanger, and its three-dimensional structure by Dorothy Hodgkin, were used to engineer slowly acting synthetic insulins for the treatment of diabetes [2]. These ground-breaking studies exemplify how therapeutic exploitation has always been closely linked to the fundamental understanding of protein structure and function. Importantly, in addition to valuable insights into protein folds, structure-function relationships and molecular evolution, the availability of the three-dimensional structures of therapeutically relevant proteins also allowed the identification and characterisation of potential inhibitor binding sites and formed the foundation for structure-based drug design (SBDD).

In the 1980s, the field rapidly evolved with biotechnology companies such as Agouron Pharmaceuticals and Molecular Discovery Ltd, respectively, pursuing structure-guided programmes aimed at inhibition of thymidylate synthase for the treatment of cancer, and inhibition of viral neuraminidase to combat influenza [1]. Another early and influential example of this approach was the use of the structure of HIV protease [4,5] in the design of four FDA-approved antiviral protease inhibitors (saquinavir, nelfinavir, indinavir and ritonavir) for the treatment of HIV/AIDS [1]. These discoveries in turn followed earlier work on inhibitors of the aspartic protease renin, which converts angiotensinogen into angiotensin I to regulate blood pressure, using molecular modelling based on structures of fungal pepsins [2].
In addition to enabling drug design, the exponential growth over the years in the number of macromolecular structures deposited in the Protein Data Bank (PDB), which at the time of writing stood at 133,920 (Figure 1), also stimulated the development of sophisticated software packages such as GRID [6], LUDI [7], GOLD [8] and GLIDE [9] to facilitate ‘docking’ of inhibitors into their predicted binding sites as well as to computationally analyse inhibitor binding and inform on further enhancement.

In the late 1990s, SBDD continued to evolve and, in particular, was a key enabler of fragment-based drug design (FBDD). This approach has its roots in computational and structural biology research from the 1980s and 1990s with key inputs from basic and applied studies in academia and at companies such Abbott Laboratories, Astex Pharmaceuticals and Vernalis [10-14]. As we write, FBDD has yielded over 30 clinical drug candidates and also three FDA-approved drugs in oncology: the BRAF inhibitor vemurafenib, the BCL-2 inhibitor venetoclax, and the CDK4 inhibitor ribociclib [15]. The main principle underlying FBDD is the efficient sampling of chemical space using small libraries of low-molecular weight (≤250 Da) compounds of low complexity [16], which makes it a complementary approach to the high-throughput screening of large libraries of compounds with lead-like properties including a molecular weight ≤450 Da. Importantly, the small size of fragment libraries and resulting low cost of entry has had the benefit of democratising the use of FBDD in academia and small companies. In addition, FBDD has been incorporated into the tool box of major pharma companies.

Although fragments typically bind with a low potency, they form efficient interactions with the protein target and thus provide attractive starting points for inhibitor design. Nowadays, a plethora of different methods are being used to identify initial fragment hits [17]. They include the traditional biophysical ones such as nuclear magnetic resonance (NMR), surface plasmon resonance (SPR), differential scanning fluorimetry (DSF) — also known as thermal shift assay (TSA), and X-ray crystallography, together with more recent ones like microscale thermophoresis (MST) and mass spectrometry [16,18]. In addition, successful fragment screens have been carried out using in vitro biochemical assays [19]. To optimise a fragment hit into a potent lead molecule, a structure-guided approach is critical. Although there are examples in which fragment hits have been optimised using a combination of biophysical and computational methods [20], we are unaware of FBDD campaigns that have not had some contribution from structural data during the hit-to-lead and/or lead optimisation stages of drug discovery.

An area in which structural biology, and hence SBDD, has struggled for many years is that of integral membrane proteins. Generally, these are very difficult to crystallise due to their large hydrophobic areas that are embedded in the membrane and also because their structural and functional integrity is often dependent on the surrounding membrane. However, advances in sample handling, stabilisation, modification and crystallisation, combined with powerful micro-focus synchrotron beamlines, have enabled the structure elucidation of many membrane proteins,
including therapeutically relevant ion channels and G-protein coupled receptors (GPCRs), and spawned subsequent SBDD and FBDD approaches for those target classes [21].

The first protein structures have been yielded by exciting new developments such as the X-ray free electron lasers (X-FELs), in which slurries of nanocrystals are irradiated with high-energy femtosecond X-ray pulses allowing the collection of a constructive diffraction image before destruction of the crystal [22,23]. Moreover, simulated experiments suggest that X-ray scattering from single molecules can, in principle, be captured, which could remove the need to crystallise a protein in the future [24].

Parallel developments in electron microscopes, such as the Titan Krios, allow the determination of structures of proteins and protein complexes at near-atomic resolution by cryo-electron microscopy (cryoEM). With this technique, the conformation of the protein of interest is again not constrained by a crystal lattice [25]. As this issue of Essays in Biochemistry was being finalised, we were delighted at the announcement that three pioneers of cryoEM, Jacques Dubochet, Joachim Frank, and Richard Henderson, were awarded the Nobel Prize in Chemistry. CryoEM structure determination was originally restricted to high-molecular weight complexes (> 300 kDa), but the cryoEM structure of isocitrate dehydrogenase 1 shows that structure solution for proteins and complexes smaller than 100 kDa is possible [25]. This achievement brings cryoEM into the realm of SBDD for targets that are not amenable to X-ray crystallography because of their size and/or flexibility.

In addition, cryoEM may provide an easier way to obtain insight into the conformation of a drug target as part of a more physiological protein complex, and thus offer crucial information complementary to the inhibitor design efforts on isolated protein targets or protein domains. Indeed, the first cryoEM structures visualising bound small-molecule ligands have already been reported and include ribosome structures with bound antibiotics [26] and the human 20S proteasome structure with a covalently bound substrate analogue (see Figure 2 and [27]). It will be exciting to see how the cryoEM field can capitalise on these achievements and further develop the technology to make it more routinely applicable to lower molecular weight proteins and with throughputs for iterative use that are more consistent with the demanding timescale of a typical drug discovery project.

In this issue of Essays in Biochemistry, we have brought together articles that collectively provide a current overview of the field of SBDD and FBDD, and also those that offer an exciting outlook on new developments that advance and complement the current repertoire of methodologies, and are applicable to target classes that are currently not amenable to current approaches.

The issue kicks off with a contribution from Martin Noble and Jane Endicott [28] who describe the use of protein structural data in understanding and interfering with the regulation of cyclin-dependent kinases (CDKs), as examples of a target class that has benefited greatly from structure-based approaches. Other recent examples include highly potent inhibitors of CDKs 8 and 19 which were enabled by SBDD (Figure 3), and that progressed to drug candidates and chemical probes that have seriously questioned the therapeutic window with such agents [29,30].
Figure 3. Crystal structure of the chemical tool CCT251545 bound to the CDK8-cyclin C complex
The inhibitor CCT251545 is shown in gold, CDK8-cyclin C is displayed in purple and key hydrogen bond interactions are shown in yellow (PDB ID code 5BNJ).
the target protein and of seeing proof that the activity can be enhanced in a manner that is consistent with the structural model [44]. A recent perspective has emphasized the use and abuse of chemical probes in exploring biology and validating targets for cancer therapy, emphasized the importance of selectivity as part of best practice and stressed the danger to research integrity of using poor quality chemical tools [45].

In addition, on the subject of research integrity, to continue to reliably exploit the rapidly increasing wealth of structural data, as a scientific community, we will have to ensure that the structures that are deposited in public databases such as the PDB are of the highest possible quality. It is worrying that despite the availability of sophisticated crystallographic hardware, data processing and structure refinement software, combined with the current high-quality structure validation tools, we notice an increase in the number of publicly available structures that have been poorly refined, or for which the interpretation borders on wishful thinking. We believe that this originates from a lack of training and/or supervision in the computational and mathematical aspects of structural biology and a limited understanding of protein stereochemistry. A rigorous understanding of crystallographic analyses and crystal structure quality within the scientific community remains critical in appreciating and assessing the value of the structural data deposited in the PDB. The same rigour should be applied to cryoEM structures, many of which are currently determined at resolutions between 5 and 3 Å, a range in which macromolecular structures are relatively poorly defined and ligands often only just about visible. Combined with the fall-off in resolution between the core and the more peripheral parts of cryoEM structures, researchers will have to be extremely careful in the modelling of macromolecular interaction partners and allosteric inhibitors binding to the surface of the studied complex.

We argue that the tradition of scientific rigour within the structural biology community should also extend to biophysical data, especially to complicated but heavily used techniques such as SPR and NMR. For example, a recent investigation by our colleagues into what, based on SPR data, was claimed to be a nanomolar inhibitor of the HSP70 molecular chaperone revealed a non-specific mode of action due to aggregate formation [46]. As scientists we are the guardians of our data — whether this is structural, biophysical, a chemical tool, or any other type of data — and we should remind ourselves of the spirit of the seminal Kleywegt and Jones paper on structure quality [47] and avoid the situation where freedom being given leads to liberties being taken.

The important role that SBDD has taken in modern drug discovery of chemical probes and drugs is exemplified by the increasing trend in some organizations, including our own, to bring together within single unified groups or departments the key related technologies of protein production, biochemical assay development, biophysical assays and X-ray crystallography — in order to maximize the efficiencies in the design and use of protein constructs, and synergies between protein structure and function.

We hope that our selection of articles will serve to educate non-expert readers who are looking to gain a sense of the current status and future prospects of structure-based approaches to discovering new drugs and chemical probes, and at the same time prove to be of interest to the aficionados in the field. It should always be remembered that potency and selectivity are not the only essential properties of drugs. The drive towards these key features must be coupled of course to optimising within the same molecule other critical properties such as selectivity, cell permeability, and pharmacokinetic and pharmacodynamic properties — a challenge that requires some compromises on particular individual features. Nevertheless, we hope that readers will be convinced that by aiming for, and often achieving through the approaches discussed here, a sufficiently close-to-perfect fit for the desired target, the application of structure-based approaches will allow us to further accelerate and enhance the discovery of chemical probes for mechanistic investigations and drugs for patient benefit.

Funding
We thank Cancer Research UK for core support (Programme Grant C2739/A22897) and other funders for supporting our research. P.W. is a Cancer Research Life Fellow. We thank many colleagues and collaborators for helpful discussions and are grateful to Ed Morris for providing the cryoEM figure of the 20S proteasome and Yann-Vaï le Bihan for critically reading the manuscript. We thank Yvonne Harman and Cerian Jones for excellent administrative support.

Competing Interests
The authors are employees of The Institute of Cancer Research (ICR) which has research collaborations with a range of commercial partners for which employees as well as ICR may benefit. R.v.M. is a former employee of Astex Pharmaceuticals and P.W. is a former employee of AstraZeneca.
References


