INTRODUCTORY ARTICLE

Perspective in Medicinal Chemistry: Structure-Based Drug Design

Miles Congreve[1,*] and Steve Andrews[2,*]

[1]*Heptares Therapeutics Ltd, BioPark, Broadwater Road, Welwyn Garden City, UK;* [2]*Alzheimer's Research UK Cambridge Drug Discovery Institute, University of Cambridge, Cambridge Biomedical Campus, Hills Road, Cambridge, UK*

Miles Congreve

Steve Andrews

Medicinal chemistry is a heavily data driven discipline in which each iteration of design and synthesis is planned from the previous generation of biological data that has just been generated. As such, the quality of the medicinal chemistry process is absolutely dependent on the quality of the screening information to derive the structure activity relationship (SAR). Most likely in parallel with the binding and/or functional effects of the current round of test compounds against the target of interest, some or all of the compounds from the previous iteration will have been tested in a range of further assays to characterise their properties. These experiments help us to understand factors such as compound solubility and physicochemical behaviour, protein binding and metabolic fate in an *in vitro* setting and activity against key anti-targets such as the hERG channel and cytochrome p450 metabolising enzymes. Using these data and the new test results in the primary *in vitro* assay, the medicinal chemists will decide on a further round of analogues to synthesise in an attempt to produce better derivatives in terms of both their biological profile, but also their drug like properties (termed multi-parameter optimisation or MPO) [1].

The problem with this approach is that there are just so very many options for synthetic targets from any given lead structure. Filtering these down to a sensible set of synthetic targets can be very difficult. If only empirical SAR is available the synthetic processes to make the compounds may become a significant bias in the final selection process. This bias will tend to leave many opportunities for modifications neglected, due to difficulties in the synthetic chemistry. Worse still if the affinity or potency data is giving so called 'flat SAR' (in which the data being generated is largely uninformative for the next round of targets) then the synthetic chemistry will tend to further dominate because, all things being equal, why not make the easy compounds? Unfortunately, in this situation many medicinal chemistry efforts are unsuccessful, unless serendipity saves the day.

In contrast, however, when structural biology has been enabled for a project and there is a good understanding of the precise mode of binding of the compound series to the target protein, Structure-Based Drug Design (SBDD) will generally allow computational and medicinal chemists to choose much better synthetic targets. The molecular design should make sense within the SAR and MPO trajectory, but importantly be optimising the fit of the ligand within the binding pocket. Even when the SAR is flat, this approach can more rationally find a breakthrough to move the process forward. Intriguingly, targets suggested by the SBDD approach would often not be made empirically because of the increased complexity and synthetic challenge which would have ruled them out for consideration. Often these more difficult synthetic targets can lead to a breakthrough in activity that move the project forward and build confidence that a drug candidate can be identified in a reasonable time frame.

SBDD also has a tendency to encourage 'atom by atom' optimisation of a ligand, particularly if the starting point is a fragment hit (FBDD) [2]. The concept of Ligand Efficiency (LE) was introduced in 2004 and serves to normalise binding affinity for molecules of different sizes [3]. This allows us to assess the quality of the ligand during optimisation and encourages a more thoughtful consideration of the contribution of each atom to the potency of a compound. In X-ray crystal structures, hydrogen bonding between the ligand and the protein can usually be readily visualised allowing optimisation of polar contacts or the introduction of new contacts. This allows ligands derived from SBDD to be designed, in best cases, to be relatively polar. Given that lipophilicity is a property now understood to be a major culprit in the attrition of drugs due to toxicological findings, the design of more polar compounds is a clear benefit of the SBDD process [4]. A detailed understanding of molecular interactions is now available to medicinal and computational chemists allowing us to better leverage the X-ray data available for design [5]. As well as these direct non-covalent interactions between a ligand and its binding site, the role of water molecules in forming indirect contacts (through-water H-bonding networks) and of solvation and desolvation of both ligands and binding sites are

*Address correspondence to these authors at the Heptares Therapeutics Ltd, BioPark, Broadwater Road, Welwyn Garden City, UK; E-mail: miles.congreve@heptares.com; and Alzheimer's Research UK Cambridge Drug Discovery Institute, University of Cambridge, Cambridge Biomedical Campus, Hills Road, Cambridge, UK; E-mail: spa26@cam.ac.uk

starting to be rationalised and increasingly considered in SBDD campaigns [6]. The role of water may be particularly important in hydrophobic binding pockets such as those often seen in G Protein-Coupled Receptors (GPCRs) as opposed to solvent exposed clefts more common in enzymes [7].

Despite the impressive progress and the broad applicability of SBDD methods, there is still much to learn about the factors that determine the energetics of binding affinity to a protein target for any given ligand. This makes it very difficult to predict ligand affinity *in silico* using computational methods such as docking and scoring [8]. The flexibility of small molecules in solution and of the protein targets themselves is also important and challenging to incorporate into the drug design process. Indeed, there is much research ongoing in the field of molecular dynamics to address these questions [9]. Another fascinating area is the thermodynamics of binding of ligands which can be much more readily measured today with modern instruments. However, the relative contributions of entropy and enthalpy to binding is often difficult to understand, although some gross trends related to molecular complexity are starting to emerge [10]. Finally, the kinetics of binding of molecules to proteins is another parameter that now can be more easily measured and in best cases optimised by developing a Structure Kinetics Relationship (SKR). As an example, it may be desirable to optimise to a slow off-rate of binding to drive a long-lasting pharmacodynamic effect *in vivo* [11]. Presently the factors that determine on and off-rate kinetics are poorly understood and SKR may at best be an empirical process where drug kinetics are rationalised rather than truly designed. But, for projects where SBDD is being employed, good quantities of pure protein will likely be available giving access to a range of biophysical methods and facilitating the determination of kinetic parameters for the project. As such, SBDD programs are often in the luxurious position to be able to both understand how molecules bind to their protein target and to be able to characterise the biophysical parameters associated with the interaction. In summary, the quality of data on the molecules being synthesised by medicinal chemists is often extremely high for SBDD campaigns, increasing the probability of success of the project.

Overall, the application of SBDD is now broadly accepted as a good investment for high value drug targets and most pharma organisations will aspire to access SBDD and biophysics whenever it is possible. As we move in to a new era of structural biology finding broader and broader applicability to protein families, including for example membrane proteins such as ion channels and GPCRs, there is real promise that one day almost all drug discovery projects can benefit from the approach [12,13]. There will of course always be sceptics who believe that the not inconsiderable investments in SBDD hardware, software and people is unnecessary, but we believe the approach has well and truly come of age and will deliver success for the next generation of drug targets.

REFERENCES

[1] Segall, M.D., Yusof, I., Champness, E.J., Avoiding Missed Opportunities by Analysing the Sensitivity of Our Decisions, *J. Med. Chem.* **2016,** *59,* 4267-4277.

[2] Verdonk, M.L. and Rees, D. C. Group efficiency: a guideline for hits-to-leads chemistry. *ChemMedChem,* **2008,** *3,* 1179-1180.

[3] Hopkins, A. L., Groom C. R. and Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today,* **2004,** *9,* 430-431.

[4] Leeson, P. D. and Empfield, J. R. Reducing the Risk of Drug Attrition Associated with Physicochemical Properties. *Annu. Rep. Med. Chem.,* **2010,** *45,* 393-407.

[5] Bissantz, C., Kuhn, B. and Stahl, M. A medicinal chemist's guide to molecular interactions. *J. Med. Chem.* **2010,** *53,* 5061-5084.

[6] Alvarez-Garcia, D. and Barril, X. Molecular simulations with solvent competition quantify water displaceability and provide accurate interaction maps of protein binding sites. *J. Med. Chem.,* **2014,** *57,* 8530-8539.

[7] Mason, J. S., Bortolato, A., Congreve, M. and Marshall, F. H. New insights from structural biology into the druggability of G protein-coupled receptors. *Trends Pharmacol. Sci.,* **2012,** *33,* 249-260.

[8] Leach, A. R., Shoichet, B. K. and Peishoff, C. E. Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. *J. Med. Chem.,* **2006,** *49,* 5851-5855.

[9] Cozzini, P., Kellogg, G. E., Spyrakis, F., Abraham, D. J., Costantino, G., Emerson, A., Fanelli, F., Gohlke, H., Kuhn, L. A., Morris, G. M., Orozco, M., Pertinhez, T. A., Rizzi, M. and Sotriffer, C. A. Target flexibility: an emerging consideration in drug discovery and design. *J. Med. Chem.,* **2008,** *51,* 6237-6255.

[10] Ferenczy, G. G. and Keserű, G. M. Thermodynamics of fragment binding. *J. Chem. Inf. Model.,* **2012,** *52,* 1039-1045.

[11] Vilums, M., Zweemer, A. J. M., Yu, Z., de Vries, H., Hillger, J. M., Wapenaar, H., Bollen, I. A., Barmare, F., Gross, R., Clemens, J., Krenitsky, P., Brussee, J., Stamos, D., Saunders, J., Heitman, L. H. and IJzerman, A. P. Structure-kinetic relationships--an overlooked parameter in hit-to-lead optimization: a case of cyclopentylamines as chemokine receptor 2 antagonists. *J. Med. Chem.,* **2013,** *56,* 7706-7714.

[12] Katritch, V., Cherezov, V. and Stevens, R. C. Diversity and modularity of G protein-coupled receptor structures. *Trends Pharmacol. Sci.,* **2012,** *33,* 17-27.

[13] Congreve, M., Dias, J. M. and Marshall, F. H. Structure-based drug design for G protein-coupled receptors. *Prog. Med. Chem.,* **2013,** *53,* 1-63.