Structural Biology, Informatics and Fragment-Based Approaches to Lead Discovery

Tom Blundell, FRS, FMedSci, William Dunn Professor of Biochemistry and Chair, School of Biological Sciences, University of Cambridge, UK

The knowledge of three-dimensional structures of protein targets that is now emerging from structural proteomics and targeted structural biology programmes has the potential to accelerate drug discovery. High-throughput structural analyses can be used to explore chemical space through fragment-screening techniques, including isothermal calorimetry (ITC), analytical ultracentrifugation (AUC), thermal shift, surface plasmon resonance (SPR) and nuclear magnetic resonance (NMR). However, high-throughput X-ray crystallography focused on identifying several weakly binding small-molecule fragments from compound libraries consisting of hundreds of small-molecule fragments has huge strengths and is an effective way of defining the chemical space of potential ligands. The high-resolution definition of these binding interactions provides

information-rich starting points for medicinal chemistry. The use of high throughput X-ray crystallography does not end there, as it becomes a rapid technique to guide the elaboration of the fragments into larger molecular weight lead compounds. I will describe such developments not only in industry but also in academia for difficult targets, including multiprotein assemblies

Looking for chemical diversity: drug design and natural compounds

Stefano Forli, The Scripps Research Institute, La Jolla, California, USA

A brief introduction on natural compounds and the relative advantages and disadvantages of using complex molecular structures coming from different living organisms for drug design. In the search for active molecules targeting the most diverse diseases, the screening of compounds isolated by living organisms is a source of diverse and original chemical scaffolds. This approach implies many advantages but is not free from issues. Microtubule stabilizing agents discussed as an example of successful approach still on going.

Automated Docking Screens: A Feasibility Study

John J. Irwin, Dept. of Pharmaceutical Chemistry, University of California San Francisco, jji@cgl.ucsf.edu

Structure-based docking has emerged as the most practical approach to leverage protein structure for ligand discovery, but the technique retains important liabilities that make it challenging to deploy systematically on a large scale. A key problem is that effective docking screens often require expert, hands-on guidance. The lack of a reliable and automatic docking protocol has attenuated the impact of this technique, particularly for biologically-oriented investigators who could most benefit from it. We have therefore created an expert system, DOCK Blaster, to investigate the feasibility of fully automating docking screens. The method begins with a PDB code, sometimes augmented with a ligand structure, and from that information can launch a full screen of large compound libraries. The system uses dictionaries of co-factors, metals, post-translational modifications and solutes to identify, separate and treat the ligand and the receptor. It then prepares the target, launches the docking calculation, and ultimately assesses the results using available experimental data, all automatically. A critical feature is self-assessment, which guides the user as to the anticipated reliability of the automated screening results. We assess performance beginning with the Astex-85 and GOLD-114 benchmarking sets, which span a broad range of target and ligand types. DOCK Blaster recapitulated the crystallographic configuration of the ligand to within 2 Å RMSD 58 % of the time for these targets, compared to 70 to 80 % of the time for expert dockers. Since a goal of the method is library docking, we asked whether the automated protocols could enrich the known ligand from among 100 physically matched decoys, each decoy set different for each target, for the ~200 Astex and GOLD targets. In 29 % of the cases the crystallographic ligand was found among the top 5% of molecules from this small screen. Taking this a step further, we undertook automated screens of the DUD benchmarking set against 38 targets. For 37 % of these targets, enrichment of the known 20-600 ligands, for any given target, was 10 fold within the top 1 %, or better, compared with 57 % for an expert docker. Thus for both geometric fidelity and enrichment DOCK Blaster is somewhat outperformed by an expert. Yet, since it is automated, DOCK Blaster can be scaled independent of human effort. We therefore undertook a study of 7,755 eligible PDB structures having 2.5 Å resolution or better and a small organic ligand bound. In 1,398 cases the re-docked ligand ranked in the top 5% of the property-matched decoys while also re-docking within 2 Å RMSD of the crystal structure, suggesting that unsupervised prospective docking is viable in these cases. DOCK Blaster is freely accessible at http://blaster.docking.org.

JJ Irwin*, BK Shoichet, MM Mysinger, N Huang, F Colizzi, P Wassam, Y Cao, Automated docking screens: a feasibility study, **J Med Chem** (2009), in press.

N Huang N, BK Shoichet* and JJ Irwin*, Benchmarking Sets for Molecular Docking, **J. Med. Chem** (2006) 49(23), 6789-6801.

JJ Irwin and BK Shoichet, ZINC - A free database of commercially available compounds for virtual screening, **J. Chem. Inf. Comput. Sci.** (2005) 45(1):177-82

JJ Irwin, How good is your screening library? (Review), **Curr. Op. Chem. Biol.** , (2006) 10(4), 352-6.

Predicting Drug Off-Targets

Michael J. Keiser, Dept. of Pharmaceutical Chemistry, University of California San Francisco, keiser@gmail.com

Chemically similar drugs often bind biologically diverse targets, yet most drugs are presumed selective at therapeutic concentrations. To investigate this assumption, we compared 3,665 FDA-approved and investigational drugs against a panel of 65,000 ligands organized into hundreds of sets according to the targets they modulate. Chemical similarities between the drugs and ligand sets were measured quantitatively using expectation values. Thousands of unanticipated drug off-targets were predicted of which only a small percentage were known; of these, 30 were experimentally tested. These included the predictions that the transporter antagonist Prozac antagonizes the β1-adrenergic G protein-coupled receptor, that the ion channel inhibitor Vadilex inhibits the 5-HT transporter, and that the enzyme inhibitor Rescriptor inhibits the histamine H4 G protein-coupled receptor. In total, 23 novel off-target drug activities were confirmed in pharmacological assays, with four being potent (< 100 nM). The physiological relevance of one of these, the drug DMT on serotonergic receptors, was confirmed in a mouse knock-out model. The chemical similarity approach used here is systematic and comprehensive, assessing essentially all known drugs against all characterized targets. The approach may find use for illuminating side effects and providing new indications for approved drugs.

MJ Keiser, BL Roth, BN Armbruster, P Ernsberger, JJ Irwin*, BK Shoichet*. Relating protein pharmacology by ligand chemistry. ***Nature Biotech*** **25**, 197-206 (2007).

MJ Keiser, V Setola, JJ Irwin, C Laggner, A Abbas, SJ Hufeisen, NH Jensen, MB Kuijer, RC Matos, TB Tran, R Whaley, RA Glennon, J Hert, KLH Thomas, DD Edwards, BK Shoichet* and BL Roth*, Predicting drug off-targets, **Nature** (2009), accepted for publication.

Abstract:

"Past, Present and Future of Structure-Based Design"

I.D. Kuntz, Department of Pharmaceutical Chemistry, University of California, San Francisco, 94143-0446

I will present an overview of the Structure-Based Drug Design, with a focus on what can be accomplished in an academic setting. I will describe our collaborative experiments using the DOCK software package. Database mining has repeatedly been able to identify compounds in the Available Chemical Directory (ACD) or ZINC chemical database that inhibit enzymes and allosteric systems in the low micromolar range. However, it is only with extensive chemical efforts that nanomolar inhibitors have been developed. Merging combinatorial chemistry strategies with structure-based design principles has had a major impact on our ability to identify subnanomolar inhibitors with the selectivity and pharmaceutical properties to be plausible preclinical candidates.

DOCK-related publications:

Kuntz, I.D., K. Chen, K.A. Sharp, P.A. Kollman 1999. The maximal affinity of ligands. Proc. Natl. Acad. Sci. **96**, 9997.

Lamb, M.L., et al. 2001. Multiple Library, Multiple Target Strategies. Proteins:Structure, Function, & Genetics, **42**, 296.

Lu et al. Proteomimetic Libraries. J. Comb. Chem. **8**, 315-325.

Moustakas et al. 2006 Development of DOCK 5. J. Comput. Aided Mol. Des. **20**, 601-619.

P.T. Lang et al. 2009 DOCK 6: Combining techniques to model RNA-small molecule complexes. RNA 15, 1219-1230.

R. Rizzo, et al. 2006. Estimation of Absolute Free Energies of Hydration using Continuum Methods. JACS., **121** 8033.

Challenges and Opportunities for Computational Docking in Structure based Drug Design.

Arthur J. Olson, Ph.D.
Professor, Department of Molecular Biology
The Scripps Research Institute
La Jolla, California USA

With the rapid increase in atomic resolution structural characterization of biological molecules and the continued Moore's Law growth of computational capabilities, the promise of structure-based rational drug design approaches general practicability. However, there remain important challenges that must be faced in using these data and applying computational technologies in an effective and reliable manner in the drug design pipeline.

In this lecture I will discuss some of the challenges that exist, including; location and characterization of druggable sites on targeted proteins; representation of the flexibility and variation of protein targets, efficient computational screening large libraries of compounds against drug targets, and accurate ranking of ligand- protein interaction for drug candidate discovery and enhancement. I will discuss these challenges in the context of several projects being undertaken in my laboratory, including: resistance-driven design of HIV therapeutics; control of oncogene expression through protein assembly stabilization; and the search for new antibiotic agents.

I will present some of the methodological and technical approaches that we are using and developing to aid our progress in facing the above challenges, including: a new program, AutoDock Vina, to accelerate the process of ligand docking; our AutoLigand program to find and characterize new binding sites; the relaxed complex method to represent protein flexibility in the docking process; massive computational screening in our Internet distributed FightAIDS@Home project; and methods to approach a more accurate representation of the entropic component of the free-energy of binding for improving the ranking of ligand candidates.

The structure-based screening and verification of novel Cdc25 phosphatase inhibitors

Seong Eon Ryu
Department of Bio-Engineering, Hanyang University, Seoul, Korea

Cdc25 phosphatases have been considered as attractive drug targets for anticancer therapy because of the correlation of their overexpression with a wide variety of cancers. Structural investigations of Cdc25 phosphatases have lagged behind the mechanistic and pharmacological studies. So far, only two X-ray crystal structures of the catalytic domains of Cdc25A and Cdc25B have been reported in their ligand-free forms.

The lack of structural information about the nature of the interactions between Cdc25 phosphatases and small molecule inhibitors has made it a difficult task to discover a good lead compound for anticancer drugs. Virtual screening has not always been successful because of the inaccuracy in the scoring function that leads to a weak correlation between the enrichment in virtual screening and binding mode prediction. The characteristic feature that discriminates our virtual screening approach from the others lies in the implementation of an accurate solvation model in calculating the binding free energy between Cdc25 phosphatases and putative inhibitors, which would have the effect on hit rate in enzyme assay.

We have been able to identify five novel Cdc25 phosphatase inhibitors with micromolar activity by means of a computer-aided drug design protocol involving the homology modeling of Cdc25A and the virtual screening with the automated AutoDock program implementing the effects of ligand solvation in the scoring function. Because the newly discovered inhibitors are structurally diverse and reveal a significant potency with IC_{50} values lower than $10\mu M$, they can be considered for further development by structure-activity relationship studies or de novo design methods.

New Ligands for Old Targets for and New Targets for Old Drugs

Brian K. Shoichet, Dept. of Pharmaceutical Chemistry, University of California San Francisco, shoichet@cgl.ucsf.edu

The "Central Dogma" of Molecular Biology has for 50 years focused on information flow from DNA to RNA to proteins, to protein structure, where it abruptly stops. Decoding protein function for structure demands an interface between protein biophysics and the chemical make up of candidate ligands. From a drug discovery perspective, this often implies docking candidate ligands and inhibitors against the structures, looking for complementary fits. This has remained challenging because of our incomplete knowledge of biologically relevant molecules and, certainly, our sampling of drug-like "chemical space". I will describe our efforts to leverage the structures of newly determined G-Protein Coupled Receptors, which have been classic drug targets for over 50 years, to discover new and potent small molecule ligands, using docking approaches.

An older pharmacological logic, now largely displaced by that of Molecular Biology, begins with the identity of small molecule reagents and drugs, and catalogs biological receptors according to their responses to more or less selective small molecules. This chemical view of biology is responsible for much of the receptor typing and sub-typing that we still use today; in its own way it was very successful. In a second approach, we take this inverted view of biology—beginning with chemicals and seeking receptors for them, rather than the reverse—and ask if we can quantify it and thereby comprehensively map biological receptors by the similarities of the ligands that bind to them, rather than by the similarities of the sequences or structures. We have compared sets of ligands annotated for over 1200 targets to one another. A statistically significant similarity score for each pair of ligand sets can be calculated once a model of random similarity is developed, and all the sets may be mapped together based on this similarity. Although no biological information is used in calculating these maps, biologically sensible clusters nevertheless appear as an emergent property. These relationships allow one to predict previously unknown off-target effects for many used drugs and reagents. We have used this "Similarity Ensemble Approach" (SEA; http://sea.docking.org/) to comprehensively look at all known drugs for likely off-targets.

I will present experimental results on predicted new ligands for the GPRC β2 Adrenergic Receptor, using docking, and experimental results on new "off-targets" for over 20 established drugs, using the chemoinformatics view.

MJ Keiser, BL Roth, BN Armbruster, P Ernsberger, JJ Irwin*, BK Shoichet*. Relating protein pharmacology by ligand chemistry. ***Nature Biotech* 25**, 197-206 (2007).

P Kolb, DM Rosenbaum, JJ Irwin, JJ Fung, BK Kobilka* & BK Shoichet*. Structure-based discovery of β_2-adrenergic receptor ligands. ***Proc. Natl. Acad. Sci.* 106**, 6843-8 (2009).

Drug Design Driven by Computational Biophysics and Chemistry

Weiliang Zhu
Drug Discovery and Design Centre, Shanghai Institute of Materia Medica, Chinese Academy of Sciences,
555 Zuchongzhi Road, Shanghai 201203, China
E-mail: wlzhu@mail.shcnc.ac.cn

Functional conformation of target protein, and intermolecular interaction between the protein and small organics are two fundamental issues for computational drug discovery and design. For the first one, although experimentally determined structures of target proteins provide solid base for structure-based drug design, proteins may adopt different functional conformations in the process of bionetwork signal transduction. Therefore, functional conformation should be more useful for drug discovery and design. However, there is no experimental method that can capture instant 3D structure when a protein performs its function. As a complementary approach to experimental technologies, computational methods could be used to explore the functional conformations of physiologically essential proteins. For the second, more kinds of intermolecular interaction, e.g., halogen bonding and cation—π interaction, should be taken into account besides classic ones, e.g., hydrogen bonding, hydrophobic and electrostatic interactions, to improve the accuracy and reliability of computationally predicted and designed results. With some projects carried out in my group as examples[1-3], this presentation demonstrates how computational biophysics and chemistry could be used for the drug discovery and design program.

References

1. Yunxiang Lu, Ting Shi, Yong Wang, Huaiyu Yang, Xiuhua Yan, Xiaoming Luo, Hualiang Jiang, Weiliang Zhu*, **Halogen Bonding — A Novel Interaction for Rational Drug Design**, *J. Med. Chem.*, **2009**, 52, 2854–2862.
2. Zhili Zuo, Oi Wah Liew, Gang Chen, Pek Ching Jenny Chong, Siew Hui Lee, Kaixian Chen, Hualiang Jiang, Chum Mok Puah*, Weiliang Zhu*, **Mechanism of NS2B-mediated activation of NS3pro in Dengue Virus: Molecular Dynamics Simulations and Bioassay Studies**, *J. Virology*, **2009**, *83*, 1060-1070.
3. Lingyan He, Jiagao Cheng, Tao Wang, Caimei Li, Zhen Gong, Hong Liu*, Bu-Bing Zeng*, Hualiang Jiang, Weiliang Zhu*, **Cation–π complexes formed between cyclooctatetraene and alkaline earth metals: Predicted and recorded NMR features**, *Chemical Physics Letters*, **2008**, 462, 45–48.

Application of high throughput crystallography in drug discovery: Fragment-based Screening.

Cheng Yang
Rigaku Americas Corporation, The Woodlands, TX 77381 USA

Over the past few years fragment-based screening has been emerging to become a popular method for designing new leads in drug discovery. In its simplest incarnation, this method involves soaking small (100-300 Da), drug-like, organic compounds into pre-formed protein crystals. As protein crystals are highly porous, fragments that have an affinity for a protein will bind at its active site, while those that do not will remain in the solvent channels. Binding at the active site is determined using X-ray crystallography to detect difference density in the active site. NMR can also be employed to detect binding (SAR by NMR), and sometimes both techniques are used in concert, with initial hits by NMR being further tested by crystallography to get a detailed 3-D map of the interaction. Fragments that bind occur at a low frequency, typically 1 in every 100, so it is important to use a large number of fragments and screen through them as rapidly as possible. high-throughput crystallographic method is the key to assess a large number of fragments in a relatively short time; typically 10-100 compounds can be scanned in a day, depending on the tractability of the protein system. The method can also detect fragments with relatively low affinity, 100 μM-10 mM. In this report, we reveal our practical approach in the Fragment-base screening.