Beilstein-Institut

Bozen 2006, May 15th – 19th, 2006, Bozen, Italy

Multiprotein Systems As Targets for Drug Discovery: Opportunities and Challenges

Tom L. Blundell, Owen R. Davies, Dima Chirgadze, Nick Furnham, Luca Pellegrini and Bancinyane L. Sibanda

Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, U.K.

E-Mail: * tom@cryst.bioc.cam.ac.uk

Received: 28th July 2006 / Published: 5th November 2007

Abstract

In the past twenty years structural biology has come to play a major role in lead optimization and target identification in the process of drug discovery. Only recently, however, has the development of high-throughput methods of structure determination provided a powerful approach to the screening of fragment binding so that structural biology can now contribute directly to lead discovery. Most targets for the new approaches continue to be enzymes, channels or receptors which tend to be "tractable" with deep and well defined cavities that bind a range of ligands. Here we discuss the challenges and opportunities in moving fragment-based approaches to target less-tractable multiprotein systems.

Background

Although drugs have traditionally been identified from natural products through *in vivo* studies, the 1980 s saw a more rational approach through the design of analogues of natural ligands of key enzymes and receptors in critical regulatory pathways. Structural biology began to contribute and soon came into its own in the optimization of leads (Fig. 1); structure-based approaches were adopted throughout the industry.

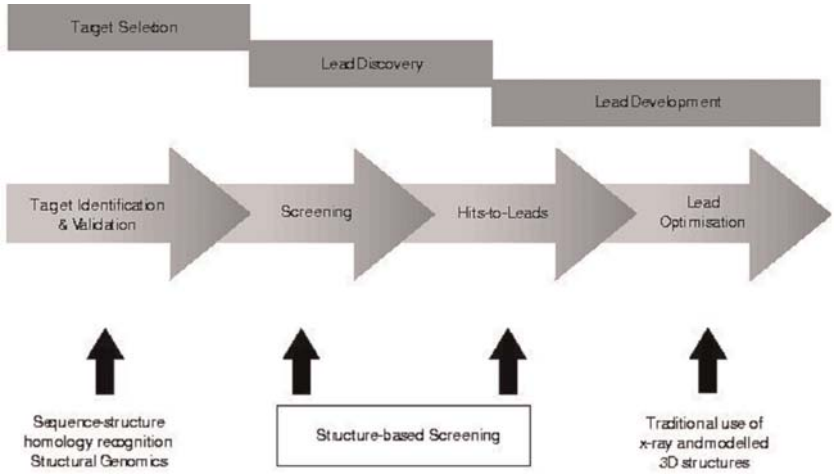

Figure 1. Modern drug discovery generally begins with target selection and progresses through the discovery of a new lead, which is then optimized and developed into a candidate drug. Structural biology has historically contributed most in the final stages of lead optimization. It has played some role through homology recognition and structural genomics to assist target selection. However, its most successful application may be in structure-based screening. Reproduced with permission from Blundell *et al.*, 2006.

But this focus on "rational" approaches was relatively short lived. In the 1990 s most pharmaceutical companies became explored the rapid assay of large numbers of compounds, so called high-throughput screening, which they believed would provide more diverse and useful leads. The adoption of high-throughput screening, which aimed to identify compounds with IC 50 s lower than 10 μM for their target proteins, signalled an intensive search for chemical diversity, in particular for the faster generation of large chemical libraries. Combinatorial chemistry was championed but the search for natural compounds was also accelerated [1 – 4]. Bioassays and systems for collection, storage and analysis of the very large datasets generated were all automated. However the rate of newly registered compounds in clinical trials did not increase in proportion to the exponential increase in investment occasioned by these new automated approaches.

The pharmaceutical industry has responded in two ways. First, it has focused on targets and their homologues that are more tractable, both in terms of the number of drug-like ligands for a target class and the nature of the binding sites [5]. As a consequence we have seen an emphasis on protein kinases and proteinases, and the concomitant development of "focused compound libraries" for particular families. Secondly, the industry has sought to design new drugs that have features of those that have been successful in the market. This has manifested itself in rules, the most significant of which has been the Lipinski "rule of five", in which the size, solubility, flexibility and hydrogen bonding characteristics are "regulated".

Fragment-Based Approaches

During this recent phase it has become increasingly clear that many “hits” deriving from high-throughput screens using bioassays were already close to the molecular weight limit of 500 Da defined by the Lipinski “rule-of-five”. It became apparent that screening very small molecules or 'fragments' might be the way forward. However, most small fragments would not be detected in the usual roboticised assays. The scene was set for the screening using physical methods [6, 7].

In the past few years fragment based approaches have advanced rapidly and drug candidates with their origins in such methodologies are now in the clinic**.** NMR [8, 9]. X-rays (Verlinde *et al.* 1997; Nienaber *et al.* 2000; Blundell *et al.* 2002; Carr & Jhoti 2002), surface plasmon resonance and isothermal calorimetry have all played a role [10 – 12]. However, X-ray crystallography has proved particularly powerful as it allows weak binders to be identified and positioned in the ligand binding site [12, 13]. Indeed such screening provides an efficient sampling of chemical space. Fragments are typically small organic molecules of between 100 and 250 Da, which exhibit low binding affinities (~100 mM to 10 mM) against target proteins and consequently can not be identified by traditional high-throughput screening. Such fragments tend to have high ligand efficiency, i.e. high values for the average free energy of binding per heavy atom [14], and this makes them attractive as start-points for optimization. A lead can be quickly developed from knowledge of how the fragment binds in the active site of the target; Astex Therapeutics quotes an average of 22 weeks for this stage of weak-hit-to-lead [15, 16].

A successful application of such methods has been the Pyramid™ approach developed by Astex Therapeutics in Cambridge where fragment libraries are screened in cocktails using X-ray crystallography [12, 15]. Automated molecular fragment matching and fitting in electron density is achieved by a software procedure called AutoSolve™ which also ranks the candidate fragments in a cocktail. Figure 2 shows a schematic representation of the steps in the Astex Pyramid™ procedure. Fragment hits derived from Pyramid™ are subsequently optimized with carefully designed iterations in order to maintain good ligand efficiency. This process has been carried out against a number of protein targets [16, 17] and including the protein kinase Cdk2, where fragment derived molecules are now in clinical trials [13, 18]**.**

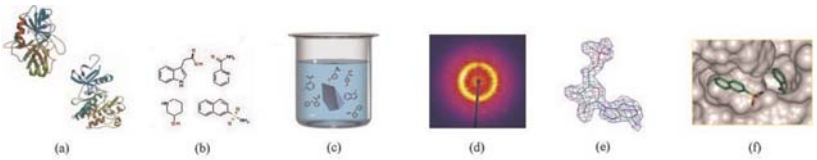

Figure 2. The Pyramid™ system allows lead discovery through a fragment-based approach of molecular fragment matching and fitting. **a.** High resolution target structure determination. **b.** Generation of Astex drug fragment library. Virtual screening used to enrich the library for fragments likely to bind the target. **c.** Drug fragment cocktails used for protein crystal soaks, 4 – 8 compounds per cocktail. **d.** High-

throughput protein/ligand X-ray crystallography. Automated X-ray data collection and analysis. **e.** Electron density analysed by AutoSolve in order to identify bound drug fragment. **f.** Structure-based optimization of hits to leads. Reproduced with permission from Blundell *et al*., 2006.

Multiprotein Protein Targets

One of the great internal contradictions of drug discovery in practice is that whilst most regulatory proteins are components of multiprotein systems, most of the focus in the pharmaceutical industry is on the active sites of monomeric proteins. Is this really sustainable?

Multiprotein complexes that mediate signal transduction and control cell processes such as differentiation, growth and proliferation have proved challenging targets for drug discovery. Many of these have large surfaces of the order of 2000 $Å^2$, for example those involved in receptor recognition and signal transduction (see for example [19]). This is especially true of complexes that are assembled from preformed globular domains. It is difficult to bind a small molecule to these large, relatively flat intermolecular surfaces involved in protein interactions but it is also difficult to disrupt the interaction entirely even if one did.

There have been several approaches taken to overcome these challenges:

1. James Wells and his colleagues [20] have sought to exploit hot spots in protein-protein interaction surfaces; these can provide binding grooves through conformational change or structural adaptivity. Examples include peptidomimetics of IL 2 that bind IL-2Rα [21].
2. Hydrophobic or amphipathic molecules have been designed to form large aggregates that interact with protein-protein interfaces – a specific mechanism of non-specific inhibition (see [22])!
3. Allosteric effects have been exploited. Ligand binding at a site distant from the protein-protein interface can have a profound effect on the assembly of multiprotein complexes. Because the allosteric effectors are often small molecules that bind in distinct and well formed pockets, these provide attractive sites to target and disrupt regulatory protein assemblies, such as those involving the integrin LFA1. Although no inhibitor has been shown to bind to the ligand ICAM1 binding site, allosteric inhibitors have been found which bind to the inactive conformation of the I-domain, thus exploiting the conformational changes in integrins that lead to clustering [23]. One recent, successful initiative has been to target the TNF-α interaction with its receptor [24]. This has led to a candidate molecule that displaces one subunit of the TNF-α trimer by a mechanism that seems to involve adaptive changes to accommodate the ligand between subunits followed by slower dissociation of the trimer, thus affecting the receptor indirectly, probably through allosteric changes.

4. Weakly binding fragments have been tethered to the protein to allow detection of binding sites [20]. A cysteine is introduced close to a puatative small-molecule-binding site and the mutant protein is "interrogated" with a library of disulphide-containing fragments with the objective of binding those that bind to the target and form a disulphide bridge. The conjugates can then be characterized by mass spectrometry.

New approaches are required to this challenging problem. We believe this will need to be a fragment-based approach, perhaps analogous to that described for tractable targets. But further thought is required as to what might comprise tractable protein-protein interfaces; adaptive binding sites are clearly inappropriate for small fragments and preformed pockets are a prerequisite. In this respect, recent analyses of multiprotein systems involved in cell regulation and signalling have identified a large number in which one component involves a flexible or unstructured region of the polypeptide chain (Fig. 5) [12, 25]. Examples are the Xrcc4 dimer in complex with DNA ligase IV (Fig. 3 shows the yeast homologues), in which the linker region between two BRCT domains appears to organize when the complex is assembled [26]. A further example (Fig. 4) involves the complex of the human recombinase, Rad51, and the product of the breast cancer associated gene, BRCA2 [27], which is not only revealing in terms of the nature of the interactions and the molecular origins of cancers associated with mutations in this region of BRCA2, but also offers an encouraging and perhaps more tractable site to target agents that would be helpful during chemo- or radio-therapy. The BRC4 repeat of BRCA2 is completely unstructured in solution but assembles by organization of its sidechains to occupy pockets in the globular structure of the Rad51 partner. We suggest that such proteins forming interactions with a ligand that comprises a continuous region of flexible peptide may be more tractable targets than where complexes are assembled from preformed globular protein structures (Fig. 5) [25]. Other similar systems that are being pursued include Bax, Bak and the BH3-only proteins which bind to Bcl-2 and Bcl-xL by inserting an α-helix into a hydrophobic groove; compounds have been developed that bind in this groove on Bcl-2 and/or Bcl-xL and thereby stimulate apoptosis [28 – 30].

Figure 3. The non-homologous end joining protein XRCC4 binds to a flexible linker between tandem BRCT domains of DNA ligase IV, imposing structure on the linker through the interaction. Here we show the complex of the yeast analogues Lif1 and

Lig4. The Figure shows two views in which the Lif1 is shown red and the Lig4 in blue. Note the extended path of the linker which presumably becomes organized only on complexation

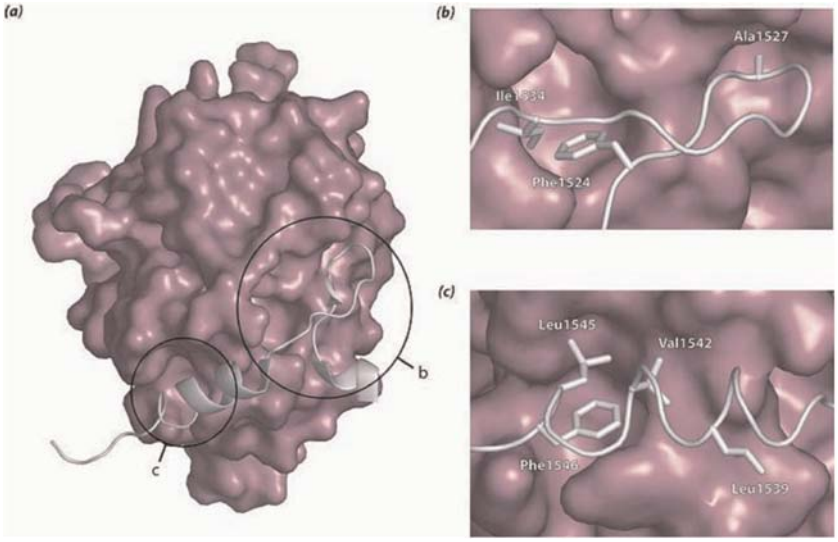

Figure 4. a. Human recombinase RAD51 binds BRC repeats of BRCA2 in an interaction that is essential for function in recombination. Although this is usually essential for normal DNA repair there is an advantage in disrupting recombination during radiotherapy and chemotherapy, which function through the introduction of DNA damage in cancerous cells. Whilst RAD51 independently forms a stable globular structure, only upon interacting with RAD51 does the BRC peptide fold into a defined 3D structure. Closer examination of the interaction (**b** and **c**) shows discrete regions of interaction that may be useful drug targets in disrupting the interaction and thus blocking recombination.

Disruption of multiprotein complexes offers completely new approaches and new targets for therapeutic intervention. However it is an area which is largely avoided by the major pharmaceutical companies, because of the complexity of the systems and the presumed difficulty in making small molecules that disrupt protein-protein interactions. Biotechnology companies have been set up specifically to work in this area (notably Sunesis, see [20]), but commercial pressures have inevitably pulled them back to more conventional targets. In the meantime there is a developing academic interest in disrupting protein-protein interactions, not only in our laboratory but also in many others e. g. the work of A. Hamilton at Yale [31].

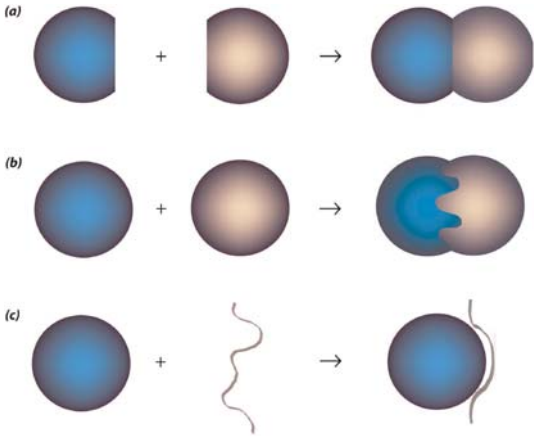

Figure 5. Multiprotein systems usually involve proteins of preformed globular structure that interact with through a discontinuous epitope (a). It has become evident that very often the globular structures adapt through conformational changes upon interaction to form a complex (b). However, it is now recognized that many protein components are natively unstructured and adopt an ordered tertiary structure only upon interaction with another partner (c). In such complexes the interacting polypeptide is a local sequence giving rise to a continuous epitope. After Blundell *et al*., 2006.

Conclusions

Structure-based approaches are now playing major roles in all stages of drug discovery. Structure-based lead optimization is practised widely with large teams of structural biologists recruited into all the large pharmaceutical companies. Drugs now in use and new ones reaching the market provide ample evidence of its usefulness and for most companies structure-guided approaches have become central to developing good drug candidates.

But many key targets for drug discovery are multiprotein systems. The discovery of small molecule antagonists that bind at the large and flat interfaces and disrupt the formation of complexes becomes a real challenge. New approaches and contributions are an urgent priority and there will be key roles in both academia and industry in advancing this process.

Acknowledgements

The authors thank the Wellcome Trust for funding for structural analysis of multiprotein systems. They thank Drs Marko Hyvonen and Nic Harmer for help in producing figures.

REFERENCES

[1] Bailey, N. *et al*.(1997) Solution-phase combinatorial chemistry in lead discovery. *Chimia* **51:**832**-**837.

[2] Spencer, R.W. (1998) High-throughput screening of historic collections: observations on file size, biological targets, and file diversity. *Biotechnol. Bioeng.* **61**(1)**:**61 – 67.

[3] Seneci, P., Miertus, S. (2000) Combinatorial chemistry and high-throughput screening in drug discovery: different strategies and formats. *Mol. Divers.* **5**(2)**:**75**-**89.

[4] Dolle, R.E. (2004) Comprehensive survey of combinatorial library synthesis: 2003. *J. Comb. Chem.* **6**(5)**:**623 – 679.

[5] Hopkins, A.L., Groom, C.R. (2002) The druggable genome. *Nature Rev. Drug Discov.* **1**(9)**:**727 – 730.

[6] Erlanson, D.A., McDowell, R.S., O'Brien, T. (2004) Fragment-based drug discovery. *J. Med. Chem.* **47**(14)**:**3463 – 3482.

[7] Rees, D.C., Congreve, M., Murray, C.W., Carr, R. (2004) Fragment-based lead discovery. *Nature Rev. Drug Discov.* **3**(8)**:**660 – 672.

[8] Shuker, S.B., Hajduk, P.J., Meadows, R.P., Fesik, S.W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**(5292)**:**1531 – 1534.

[9] Fejzo, J., Lepre, C.A., Peng, J.W., Bemis, G.W., Ajay, Murcko, M.A., Moore, J.M. (1999) The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. *Chem. Biol.* **6**(10)**:**755 – 769.

[10] Verlinde, C.L. M.J., Kim, H., Bernstein, B.E., Mande, S.C., Hol, W.G. J. (1997) Antitrypanosomiasis drug development based on structures of glycolytic enzymes. In: *Structure-based Drug Design*. (V.P. Ed.), pp.365 – 394. Marcel Dekker Inc., New York.

[11] Nienaber, V.L., Richardson, P.L., Klighofer, V., Bouska, J.J., Giranda, V.L., Greer, J. (2000) Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nature Biotechnol.* **18**(10)**:**1105 – 1108.

[12] Blundell, T.L., Abell, C., Cleasby, A., Hartshorn, M.J., Tickle, I.J., Parasini, E., Jhoti, H. (2002)High-throughput X-ray crystallography for drug discovery. In: *Drug Design: Special Publication*. (Flower, D.R., Ed.), pp. 53 – 59. Royal Society of Chemistry, Cambridge.

[13] Congreve, M., Murray, C.W., Blundell, T.L. (2005) Structural biology and drug discovery. *Drug Discov. Today* **10:**895 – 907

[14] Hopkins, A.L., Groom, C.R., Alex, A. (2004) Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* **9**(10)**:**430**-**431.

[15] Carr, R., Jhoti, H. (2002) Structure-based screening of low-affinity compounds. *Drug Discov. Today* **7**(9)**:**522 – 527.

[16] Gill, A.L., Frederickson, M., Cleasby, A., Woodhead, S.J., Carr, M.G., Woodhead, A.J., Walker, M.T. *et al.* (2005) Identification of novel p38alpha MAP kinase inhibitors using fragment-based lead generation. *J. Med. Chem.* **48**(2)**:**414 – 426.

[17] Hartshorn, M.J., Murray, C.W., Cleasby, A., Frederickson, M., Tickle, I.J., Jhoti, H. (2005) Fragment-based lead discovery using X-ray crystallography. *J. Med. Chem.* **48**(2)**:**403 – 413.

[18] Congreve MS, Davis DJ, Devine L, Granata C, O'Reilly M, Wyatt PG, Jhoti H. (2003) Detection of ligands from a dynamic combinatorial library by X-ray crystallography. *Angew. Chemie Intl Edn* **42:**4479 – 4482.

[19] Pellegrini, L., Burke, D.F., von Delft, F., Mulloy, B., Blundell, T.L. (2000) Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature* **407**(6807)**:**1029 – 1034.

[20] Arkin, M.R., Wells, J.R. (2004) Small-molecule inhibitors of protein- protein interactions: progressing towards the dream. *Nature Rev. Drug Discov***. 3:**301 – 317.

[21] Tilley, J.W., Chen, L., Fry, D.C., Emerson, S.D., Powers, G.D., Biondi, D., Varnell, T., Trilles, R., Guthrie, R., Mennona, F., Kaplan, G., LeMahieu, R.A., Carson, M., Han, R.J., Liu, C.M., Palermo, R., Ju, G. (1997) Identification of a small molecule inhibitor of the IL-2/IL-2R receptor interaction which binds to IL-2. *J. Am. Chem. Soc.* **119:**7589 – 7590.

[22] McGovern, S.L., Helfand, B.T., Feng, B., Shoichet, B.K. (2003) A specific mechanism of nonspecific inhibition. *J. Med. Chem.* **46:**4265 – 4272.

[23] Kallen, J. *et al*. (1999) Structural basis for LFA-1 inhibition upon lovastatin binding to the CD11a I-domain. *J. Mol. Biol*. **292:**1 – 9 .

[24] He *et al.* (2005) Small-molecule inhibition of TNF-α. *Science* **310:**1022 – 1025.

[25] Blundell, T.L., Sibanda, B.L., Wander Montalvao, R., Brewerton, S., Chelliah, V., Worth, C.L., Harmer, N.J., Davies, O., Burke, D. (2006) Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery. *Phil. Trans. R. Soc. B.* **36:**413 – 423

[26] Sibanda, B.L., Critchlow, S.E., Begun, J., Pei, X.Y., Jackson, S.P., Blundell, T.L., Pellegrini, L. (2001) Crystal structure of an Xrcc4-DNA ligase IV complex. *Nature Struct. Biol.* **8**(12)**:**1015 – 1019.

[27] Pellegrini, L., Yu, D.S., Lo, T., Anand, S., Lee, M., Blundell, T.L., Venkitaraman, A.R. (2002) Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420**(6913)**:**287 – 293.

[28] O'Neill, J., Manion, M., Schwartz, P., Hockenbery, D.M. (2004) Promises and challenges of targeting Bcl-2 anti-apoptotic proteins for cancer therapy. *Biochim. Biophys. Acta* **1705:**43 – 51.

[29] Qian, J., Voorbach, M.J., Huth, J.R., Coen, M.L., Zhang, H., Ng, S.C., Comess, K.M., Petros, A.M., Rosenberg, S.H., Warrior, U., Burns, D.J. (2004) Discovery of novel inhibitors of Bcl-xL using multiple high-throughput screening platforms. *Analyt. Biochem.* **328:**131 – 138.

[30] Real, P.J., Cao, Y., Wang, R., Nikolovska-Coleska, Z., Sanz-Ortiz, J., Wang, S., Fernandez-Luna, J.L. (2004) Breast cancer cells can evade apoptosis-mediated selective killing by a novel small molecule inhibitor of Bcl-2. *Cancer Res.* **64:**7947 – 7953.

[31] Ernst, J.T., Becerril, J., Park, H.S., Yin, H., Hamilton, A.D. (2003) Design and application of an alpha-helix-mimetic scaffold based on an oligoamide-foldamer strategy: antagonism of the Bak BH3/Bcl-xL complex. *Angew. Chemie Intl Edn Engl.* **42:**535 – 539.