

# Fragment approaches in structure-based drug discovery

Roderick E. Hubbard

Received 2 August 2007

Accepted 3 December 2007

Vernalis (R&amp;D) Ltd and University of York, UK. E-mail: rod@ysbl.york.ac.uk

There has been considerable interest recently in what is known as 'fragment-based lead discovery'. The novel feature of the approach is to begin with small low-affinity compounds. The main advantage is that a larger potential chemical diversity can be sampled with fewer compounds, which is particularly important for new target classes. The approach relies on careful design of the fragment library, a method that can detect binding of the fragment to the protein target, determination of the structure of the fragment bound to the target, and the conventional use of structural information to guide compound optimization. In this article the methods are reviewed, and experiences in fragment-based discovery of lead series of compounds against kinases such as PDK1 and ATPases such as Hsp90 are discussed. The examples illustrate some of the key benefits and issues of the approach and also provide anecdotal examples of the patterns seen in selectivity and the binding mode of fragments across different protein targets.

© 2008 International Union of Crystallography  
Printed in Singapore – all rights reserved

## 1. Introduction

There has been considerable interest recently in what is known as 'fragment-based lead discovery' (Erlanson *et al.*, 2004; Rees *et al.*, 2004; Hajduk & Greer, 2007), and a number of drugs entering clinical trials have been discovered using these techniques (Card *et al.*, 2005; Gill *et al.*, 2005; Petros *et al.*, 2006). The major new advance is to begin with much smaller (molecular weight  $M_r < 250$ ) compounds. This has significant advantages in being able to sample a large potential chemical diversity with a small number of compounds. In most other respects the methods are similar to the more traditional use of protein-ligand structures to understand structure–activity relationships (SAR) and to guide optimization of the binding affinity, selectivity and drug-like characteristics of a compound for a particular protein active site.

The ideas of fragment-based discovery developed during the 1990s. Scientists at Abbott pioneered the approach for drug discovery, using NMR to identify fragments binding in the SAR by the NMR approach (Shuker *et al.*, 1996). Although various academic groups had explored systematic binding of solvent molecules to proteins (Mattos & Ringe, 2001; English *et al.*, 2001), the first screening of fragments by crystallography linked to drug discovery was from another Abbott group (Nienaber *et al.*, 2000). These ideas were subsequently developed by various small technology companies (Rees *et al.*, 2004; Card *et al.*, 2005; Blaney *et al.*, 2006) for more focused application in drug discovery.

Most approaches to fragment-based discovery rely on careful design of the fragment library, a method that can detect the binding of the fragment to the protein target, determination of the structure of the fragment bound to the target, and the conventional use of structural information to guide compound optimization. At Vernalis, we have developed an approach that we call SeeDs (structural exploitation of experimental drug startpoints) (Hubbard, Davis *et al.*, 2007).

## 2. The SeeDs approach

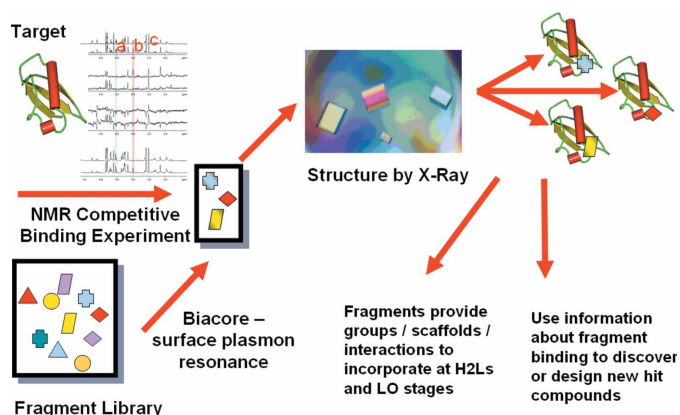
The essential steps are summarized in Fig. 1 and consist of selecting a library of suitable fragments, identifying which fragments bind to the target binding site using NMR or surface plasmon resonance methods, determination of the structure of the fragments binding to the protein by X-ray crystallography, and then using the structural information to guide the evolution of improved hit or lead compounds.

### 2.1. Fragment library design

The design of the fragment library is one of the most critical stages, as in any screening process. Unfortunately, little detail is published about library design. The limited non-proprietary information that is available has been summarized recently (Hubbard, Chen & Davis, 2007). Most strategies for the design of fragment libraries include selection on the presence or absence of either desirable or undesirable chemical functionality and this can often be influenced by the subsequent chemistry to be used to evolve fragments. In addition, the methods used to identify which fragments are binding can place additional constraints on the library in terms of solubility, shape, flexibility or spectral properties. As an example, Fig. 2 summarizes the cheminformatics workflow used to generate an incremental addition to the Vernalis fragment libraries (Baurin, Aboul-Ela *et al.*, 2004). The first two stages can be performed computationally. The major investment comes in visual inspection of the compounds to establish chemical tractability and the extent of physico-chemical quality control required to generate a final library.

### 2.2. Identifying fragments that bind

Fragments bind extremely weakly to a target with useful fragments of  $M_r < 250$  binding with affinities ranging from 10  $\mu\text{M}$  to 10 mM. Measuring such weak binding is a real challenge for most activity and

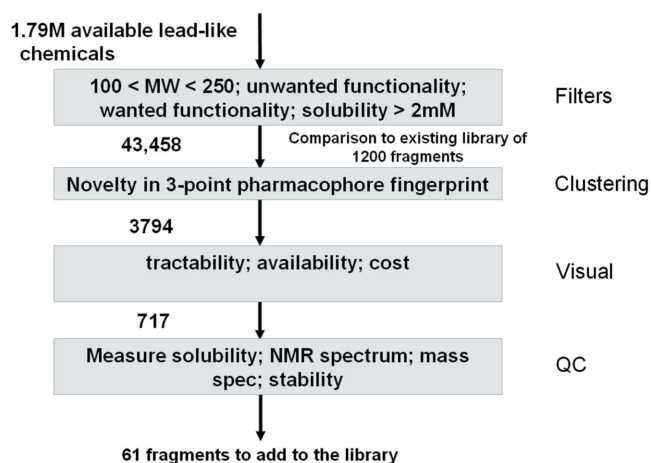


**Figure 1**  
Constructing a fragment library: an example of the pipeline of cheminformatics calculations used to identify suitable compounds to include in a fragment library.

binding assays because of interference by the high concentrations of ligand required. For this reason, biophysical methods have proved popular for identifying binding. Some groups screen using X-ray crystallography (Rees *et al.*, 2004; Blaney *et al.*, 2006), where mixtures of fragments are soaked into crystals of the unliganded protein and difference electron density maps interpreted to identify which fragments have bound. This method immediately provides details of the structure of the fragment bound to the protein. NMR spectroscopy is able to monitor binding at this type of concentration with the added advantage that experiments can readily be configured to check for precipitation, protein unfolding, competitive binding and even mapping where the fragment binds on the protein surface. More recently, there have been developments in surface plasmon resonance (such as the Biacore technology) as a complementary or confirmatory technique which can also directly provide binding affinity data. Each of these methods has benefits and disadvantages (see review by Hubbard, Chen & Davis, 2007).

### 3. Examples of fragments binding to Hsp90 and PDK1

The project teams at Vernalis have used the SeeDs approach to identify potent selective inhibitors for a number of different proteins, including various kinases, ATPases and protein-protein interaction



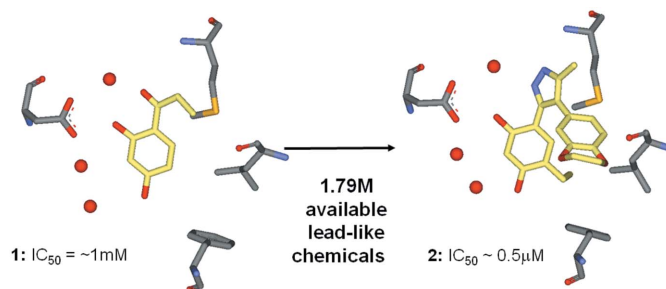
**Figure 2**  
The Vernalis SeeDs process is an example of the steps in fragment-based drug discovery. See text for explanation.

targets. Here, we provide brief information on two recent projects, Hsp90 and PDK1.

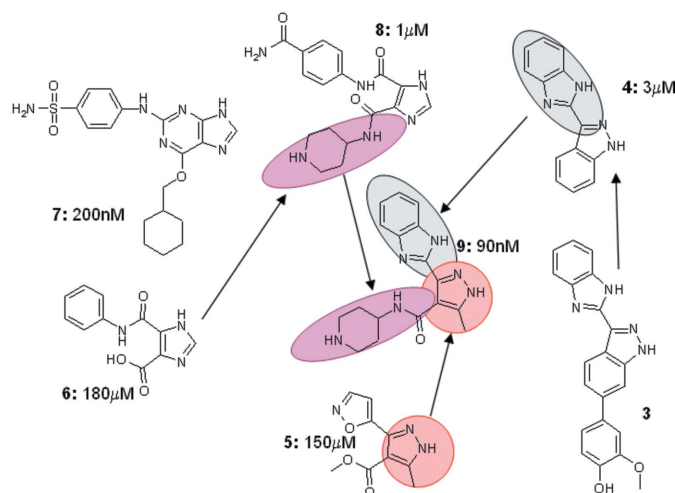
#### 3.1. Hsp90

Hsp90 is a molecular chaperone that stabilizes the final stages of folding of a wide range of proteins in cells under stress. The protein is important for the survival and growth of cancer cells, where client proteins include a number of signalling proteins required for cell growth as well as stabilization of mutated proteins. Hsp90 consists of three domains, with an N terminal domain harbouring ATPase activity crucial for turnover of proteins. A number of natural products and, more recently, synthetic inhibitors have been identified which selectively inhibit this ATPase activity, some of which are entering clinical trials for various cancers (see review by Drysdale *et al.*, 2006). Vernalis scientists have identified various series of Hsp90 inhibitors, including some derived using fragment-based methods.

A fragment library of 790 fragments was assessed for competitive binding to the Hsp90 binding site. Some 17 fragments were identified, one of which (compound 1 in Fig. 3) is a resorcinol. The resorcinol substructure was used to search a catalogue of commercially available compounds (Baurin, Richardson *et al.*, 2004). Following assessment with focused docking (*rDock*, <http://www.ytbl.york.ac.uk/rDock>), a number of compounds were purchased and assayed. From this, compound 2 was discovered which has an  $IC_{50}$  of less than  $0.5 \mu M$ . This is the same compound as identified by a collaborator in a



**Figure 3**  
SAR by catalog: the structure of resorcinol fragment 1 bound to the active site of Hsp90 was used to select compound 2, which formed the initial lead series that generated a pre-clinical candidate.



**Figure 4**  
Fragment evolution: combining structural information about the binding of series of compounds from the literature and fragment screening to generate the potent lead compound 9 for the kinase PDK1.

medium-throughput screen (Cheung *et al.*, 2005) which has subsequently been optimized (Sharp *et al.*, 2007) and is in pre-clinical development. This is an example of how fragment screening, followed by 'SAR by catalogue' can rapidly identify good lead compounds.

### 3.2. PDK1

PDK1 is a classical Ser-Thr kinase (Mora *et al.*, 2004). Growth factors stimulate PI3 kinase to modulate the phosphorylation status of phospho-inositol lipids. These recruit PDK1 to the membrane *via* its PH domain where it can regulate the activity of a closely related kinase, Akt, by phosphorylation at position T308. Akt modulates the activity of both growth and metabolic pathways through phosphorylation of GSK3- $\beta$ . In addition, PDK1 is implicated in the activation of more than 20 other kinases. PDK1 is over-expressed in some cancers and recent evidence has emerged that PDK1 is important for tumour growth (Collins *et al.*, 2005).

Fig. 4 summarizes how information derived from fragments and literature compounds was combined to derive compound 9, a potent PDK1 inhibitor. The structure of compound 3 [a known Chk1 inhibitor (Kania *et al.*, 2001)] bound to Chk1 suggested compound 4, which binds to PDK1 with an  $IC_{50}$  of 3  $\mu M$ . Compounds 5 and 6 were two out of the 80 or so fragments identified by the SeeDs process that bind to PDK1. Compound 7 is a CDK2 inhibitor which also has PDK1 activity (Davies *et al.*, 2002). The structure of 7 bound to PDK1 showed that the cyclohexyl ring lies on an essentially hydrophobic surface in PDK1. The database of available compounds (Baurin, Aboul-Ela *et al.*, 2004) was searched for compounds that contained the 5-carboxy pyrazole core of compound 7 and a hydrophobic ring and which docked successfully into the PDK1 structure (*rDock*). One of the subsequent hits was optimized to give compound 8 which has an  $IC_{50}$  of 1  $\mu M$  for PDK1. The crystal structures of compounds 4, 5 and 8 bound to PDK1 were determined and suggested compound 9 which is the combination of the benzimidazole from 4 with the core pyrazole of 5 with the addition of the piperidine from 8. Compound 9 binds with an  $IC_{50}$  of 90 nM. Subsequent optimization has given higher-affinity compounds, some of which are tolerated *in vivo* and show signs of appropriate pharmacodynamic marker effects in xenograft models.

## 4. Challenges for fragments

Fragment methods are now firmly established as a useful approach to discovering novel hit compounds for some classes of target. It should be stressed that fragments are intrinsically no different from any other hit compound; they are just small and weak binders. The main issues are therefore as follows.

(i) Design of the library of fragments. Most fragment libraries to date have evolved from analysis of known drug compounds and/or selection from the commercially available compounds using appropriate diversity and physico-chemical properties (reviewed by Hubbard, Chen & Davis, 2007). There is scope for further design and synthesis of novel fragments; for example, by considering the three-dimensional shape and distribution of functionality.

(ii) Detecting such weak binding compounds. There is real scope for improving the methods for detecting which fragments are binding to a particular target. Ideally, assays should be available which can accurately determine binding affinities at up to 10 mM. The current biophysical methods (SPR and NMR) can provide such information, but significant improvements are required to improve sensitivity, reduce the tendency for artefacts and the quantity of protein material required for screening.

(iii) Generating sufficient information with which to grow the fragments to more potent compounds. As with any compound optimization, the key requirement is generation of sufficient understanding of structure-activity relationships for the design of compound improvements. If the assay (see above) is accurate and reliable enough, then in principle it should be possible to grow fragments in the absence of direct structural information. Alternatively, if the active site of the protein does not change on binding of fragments, then it is possible to predict the binding mode of fragments with some confidence using docking methods. However, currently for most targets, a crystal structure of the fragment bound to the target is required to successfully evolve the fragments into hit compounds on the scale of the assay.

In conclusion, the early pioneers of fragment methods have used structure-based methods to demonstrate the proof of concept that weak millimolar-binding small compounds can be evolved into high-affinity drug-like molecules, and a number of compounds discovered using these methods are either in or moving towards clinical trials. These successes have stimulated many companies to develop fragment-based screening strategies and the methods are now firmly established as a successful strategy for hit discovery.

Many scientists at Vernalis have contributed to the work summarized in this article. In particular, Ben Davis has led the development of the SeeDs approach, Nicolas Baurin and Ijen Chen have developed the informatics and modelling approaches to fragments, and James Murray, Lisa Wright, Allan Surgenor and Pawel Dokurno have solved many hundreds of crystal structures. The PDK1 project is led by Lee Walmsley and Chris Torrance and the Hsp90 project by Martin Drysdale and Paul Brough.

## References

- Baurin, N., Aboul-Ela, F., Barril, X., Davis, B., Drysdale, M., Dymock, B., Finch, H., Fromont, C., Richardson, C., Simmonite, H. & Hubbard, R. E. (2004). *J. Chem. Inf. Comput. Sci.* **44**, 2157–2166.
- Baurin, N. R., Richardson, C., Potter, A., Jordan, A., Roughley, S., Parratt, M., Greaney, P., Morley, D. & Hubbard, R. E. (2004). *J. Chem. Inf. Comput. Sci.* **44**, 643–651.
- Blaney, J., Nienaber, V. & Burley, S. K. (2006). *Fragment-Based Approaches in Drug Discovery*, edited by W. Jahnke and D. A. Erlanson, pp. 215–245. Weinheim: Wiley-VCH.
- Card, G. L. *et al.* (2005). *Nat. Biotechnol.* **23**, 201–207.
- Cheung, K. M., Matthews, T. P., James, K., Rowlands, M. G., Boxall, K. J., Sharp, S. Y., Maloney, A., Roe, S. M., Prodromou, C., Pearl, L. H., Aherne, G. W., McDonald, E. & Workman, P. (2005). *Bioorg. Med. Chem. Lett.* **15**, 3338–3343.
- Collins, B. J., Deak, M., Murray-Tait, V., Storey, K. G. & Alessi, D. R. (2005). *J. Cell Sci.* **118**, 5023–5034.
- Davies, T. G. *et al.* (2002). *Nat. Struct. Biol.* **9**, 745–749.
- Drysdale, M. J., Brough, P. A., Massey, A., Jensen, M. R. & Schoepfer, J. (2006). *Curr. Opin. Drug Discov. Devel.* **9**, 483–495.
- English, A. C., Groom, C. R. & Hubbard, R. E. (2001). *Protein Eng.* **14**, 47–59.
- Erlanson, D. A., McDowell, R. S. & O'Brien, T. (2004). *J. Med. Chem.* **47**, 3463–3482.
- Gill, A. L. *et al.* (2005). *J. Med. Chem.* **48**, 414–426.
- Hajduk, P. J. & Greer, J. (2007). *Nat. Rev. Drug Disc.* **6**, 212–219.
- Hubbard, R. E., Chen, I. & Davis, B. (2007). *Curr. Opin. Drug Discov. Devel.* **10**, 289–297.
- Hubbard, R. E., Davis, B., Chen, I. & Drysdale, M. J. (2007). *Curr. Trends Med. Chem.* **7**, 1568–1581.
- Kania, R. S., Bender, S. L., Borchardt, A., Braganza, J. F., Cripps, S. J., Hua, Y., Johnson, M. D., Johnson, T. O. Jr, Luu, H. T., Palmer, C. L., Reich, S. H. & Tempczyk-Russell, A. M. (2001). Patent WO 0102369.
- Mattos, C. & Ringe, D. (2001). *Curr. Opin. Struct. Biol.* **11**, 761–764.

- Mora, A., Komander, D., van Aalten, D. M. & Alessi, D. R. (2004). *Semin. Cell Dev. Biol.* **15**, 161–170.
- Nienaber, V. L., Richardson, P. L., Klighofer, V., Bouska, J. J., Giranda, V. L. & Greer, J. (2000). *Nat. Biotechnol.* **18**, 1105–1108.
- Petros, A. M. *et al.* (2006). *J. Med. Chem.* **49**, 656–663.
- Rees, D. C., Congreve, M. & Murray, C. W. (2004). *Nat. Rev. Drug Discov.* **3**, 660–672.
- Sharp, S. Y. *et al.* (2007). *Mol. Cancer Ther.* **6**, 1198–1211.
- Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. (1996). *Science*, **274**, 1531–1534.