Fragment-Based Lead Discovery Using X-ray Crystallography

Michael J. Hartshorn, Christopher W. Murray, Anne Cleasby, Martyn Frederickson, Ian J. Tickle, and Harren Jhoti*

Astex Technology, 436 Cambridge Science Park, Milton Road, Cambridge, CB4 0QA, United Kingdom

Received June 3, 2004

Fragment screening offers an alternative to traditional screening for discovering new leads in drug discovery programs. This paper describes a fragment screening methodology based on high throughput X-ray crystallography. The method is illustrated against five proteins (p38 MAP kinase, CDK2, thrombin, ribonuclease A, and PTP1B). The fragments identified have weak potency (>100 μM) but are efficient binders relative to their size and may therefore represent suitable starting points for evolution to good quality lead compounds. The examples illustrate that a range of molecular interactions (i.e., lipophilic, charge−charge, neutral hydrogen bonds) can drive fragment binding and also that fragments can induce protein movement. We believe that the method has great potential for the discovery of novel lead compounds against a range of targets, and the companion paper illustrates how lead compounds have been identified for p38 MAP kinase starting from fragments such as those described in this paper.

Introduction

Over the last two decades there has been considerable interest in new approaches to drug discovery that offer improvements in the process of identifying new therapeutic agents. Technologies such as high-throughput screening and combinatorial chemistry have taken hold in most pharmaceutical companies and allow significantly larger numbers of compounds to be screened against the target of interest. Despite these developments, the industry has failed to generate the level of productivity that it has strived to achieve.[1] All aspects of the drug discovery process therefore remain the focus of improvements with the application of new technologies. One area that has continued to receive significant interest is lead discovery chemistry as the quality of lead compounds is thought to have a major impact on the attrition rates in drug development. Many groups have reported on the importance of compound libraries for lead generation to become more lead-like rather than drug-like.[2,3] Such an approach takes into account the increase in molecular weight and lipophilicity that typically occurs as a lead molecule is optimized into a potential drug.

More recently, interest has grown in a new approach for lead generation that involves screening libraries of compounds that are significantly smaller and functionally simpler than drug molecules, often referred to as 'fragment-based' discovery.[4,5] This new approach is believed to have many advantages over conventional screening such as more efficient sampling of chemical space using fewer compounds and a more rapid hit-to-lead optimization phase. Fragment-based drug discovery also provides significant challenges, largely due to the fact that fragments typically exhibit low affinity binding (100 μM−mM) and are therefore difficult to detect using bioassay-based screening methods. A variety of alternative biophysical methods have therefore been used to detect the binding of such fragments.[4,6−15] Although fragments often have low affinity, they usually exhibit high ligand efficiency, i.e., high values for the ratio of free energy of binding to the number of heavy atoms.[16] It is important that when a fragment hit is identified, optimization into a useful lead compound is performed with carefully designed iterations consistent with maintaining good ligand efficiency. Here we report several examples where fragment libraries were screened using X-ray crystallography as the primary method for hit detection in a process that we call Pyramid. The limitations and advantages of using X-ray crystallography for fragment screening and the nature of the fragment hits are outlined in these different cases. The accompanying paper describes examples of fragment hit-to-lead optimization for p38 MAP kinase, a key target in inflammatory disease.

High-Throughput X-ray Crystallography

One of the limitations of X-ray crystallography in drug discovery has historically been the time required to obtain crystal structures of a target protein. Even in cases where the target protein structure is known, the time taken to obtain co-complexes with key compounds has often limited the impact of the approach in lead optimization. Developments in technology over the last 5 years have had a dramatic impact on improving the speed of obtaining co-crystal structures, and, furthermore, a growing number of therapeutic targets have had their crystal structures determined. Despite these improvements, a number of additional specific developments have been required in order to use X-ray crystallography as a primary screening technique.

Efficient fragment screening requires the soaking of cocktails of fragments into preformed crystals of the target protein. After collection of the X-ray data, the identification of the active fragment from the cocktail would normally be reliant on manual analyses of the resultant electron density. Such analysis is subjective

* To whom correspondence should be addressed. Phone: +44 (0)-1223 226200; Fax +44 (0)1223 226201; E-mail: h.jhoti@astex-technology.com.

10.1021/jm0495778 CCC: $30.25
Published on Web 12/21/2004

and time-consuming, so a software procedure, AutoSolve, was developed that is capable of rapidly interpreting electron density maps for fragment binding.[17] Another key development has been the implementation of robotic systems that are able to store and automatically manipulate protein crystals, thus removing the need for manual intervention during X-ray data collection.[18,19] Such hardware developments integrated with new software, such as AutoSolve, have transformed X-ray crystallography into a technique that can be used for screening.

Fragment Libraries

Two complementary fragment sets were used in this work. The first set seeks to take advantage of the fact that fragments are small and can be chosen to probe a small number of potential interactions with the protein. Hann[20] has elegantly demonstrated that such fragments can efficiently sample chemical space compared to larger compounds with more functionality. Based on functional groups and scaffolds well represented in drugs, it should therefore be possible to represent drug fragment space with a relatively small number of compounds. The chemoinformatics approach used to construct the drug fragments set is described below.

Targeted sets were also constructed for crystallographic screens against particular proteins or protein classes. Two classes of information were used to construct the targeted sets. First, knowledge of known ligands and their key interactions with proteins has been used to select fragments for acquisition or synthesis, and second, virtual screening has been used to acquire compounds that dock well into the active site of target proteins. The methods used are discussed in more detail below.

In general, the fragments in these sets have molecular weights between 100 and 250 and are relatively simple with few functional groups, making them chemically tractable and suitable for rapid optimization. In a few examples, compounds from the targeted sets are slightly larger or more complex because particular questions were being probed that required greater complexity or because of the constraints of commercial availability. Compounds that were unlikely to be soluble under the experimental conditions used for crystallographic screening were not included in the fragment sets.

Drug Fragment Set. This set was designed to provide a diverse range of compounds that contain ring systems and functionalities often found in known drug molecules. The first stage was to identify a small set of simple organic ring systems that occur in drug molecules so that they are likely to have reduced toxicity liabilities and are amenable to development by medicinal chemistry practices. Several analyses have demonstrated that only a small number of simple organic ring systems (sometimes known as scaffolds or frameworks in other work) occur in many drug molecules.[14,21,22] Low molecular weight ring systems were selected and are shown in Figure 1. In addition to the drug rings a set of simple carbocyclic and heterocyclic ring systems were chosen as the basis for additional fragments. The ring systems chosen are shown in Figure 2.

The virtual library, from which the Drug Fragment Set is chosen, was generated by combining the ring

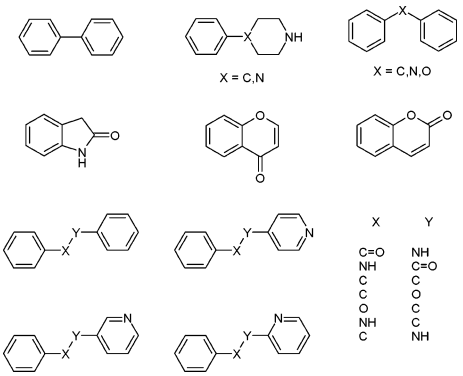

Figure 1. Drug ring systems.

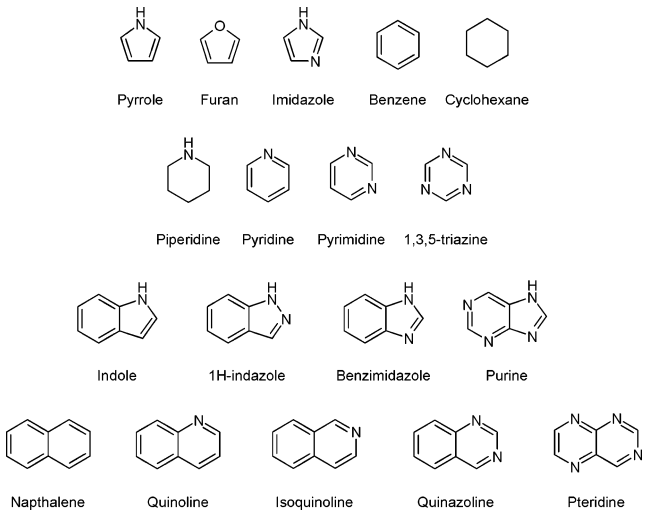

Figure 2. Simple carbocyclic and heterocyclic ring systems.

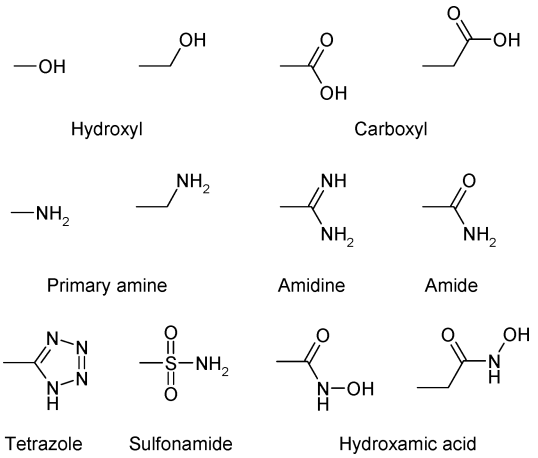

Figure 3. Preferred side chains.

systems shown above with a set of desirable side chains. The side chains used in this process are broken into three categories.

1. Preferred side chains. This set of side chains consists of those that are observed frequently in drug molecules. The side chains are shown in Figure 3.

2. Lipophilic/secondary side chains. The properties of the fragments are further modulated by allowing substitution with a set of secondary side chains. Most of these are lipophilic and are intended to pick up hydrophobic interactions in a protein binding site. The secondary side chains for carbon atoms are shown in Figure 4.

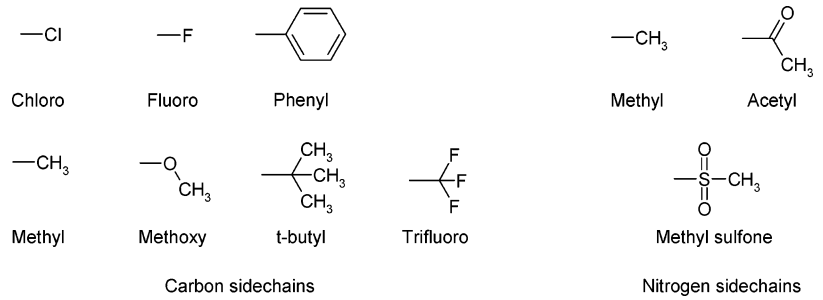

Figure 4. Secondary side chains.

3. N-Substituents. The final possibility for substituting side chains onto a framework is a set of N-substituents. These are also shown in Figure 4.

The virtual library was generated by substituting each of the relevant side chains onto each of the ring systems. Each carbon atom was substituted by the side chains from the preferred group and by those from the secondary side chains. The nitrogen atoms were only substituted from the group of N-side chains. Each ring system was only substituted at one position at a time, with the exception of benzene and imidazole. The latter two were disubstituted at all pairs of positions, with either two preferred side chains or one preferred side chain and one secondary side chain.

The virtual library was generated as SMILES strings, which were searched for in a database of available compounds. A total of 4513 compounds were generated by the virtual enumeration stage of which 401 were available from commercial suppliers. Manual inspection to remove known toxophores and unavailability of some compounds resulted in a final set of 327 compounds (the Drug Fragment Set).

Targeted Sets. Virtual screening using GOLD[23−25] was also used to construct target-specific fragment sets. The molecules were selected using a proprietary virtual screening platform which has been described in detail elsewhere.[26] Briefly, a database called ATLAS (Astex Technology Library of Available Substances) has been constructed from chemical and library suppliers and stored as smiles strings in an Oracle database. This database of over 3.6 million unique chemical compounds can be queried from the Windows desktop using substructure filters and physical property filters (such as molecular weight, clogP, polar surface area, rule of 3 criteria,[27] etc). This querying produces lists of available compounds that satisfy the requirements of the user, which can then be docked against an active site for the protein of interest. CORINA[28] is used to convert the smiles strings into 3D structures. A proprietary version of the program GOLD,[25] was used to perform the docking and a variety of scoring functions were used (i.e., GoldScore[23] and ChemScore[29] both with and without protein based pharmacophores). The docking jobs were run on a Linux cluster. Generally the scoring function that was used in the virtual screening was the one that produced good docking modes for a test set of ligands for which the binding mode was known. In the case of thrombin, a pharmacophore was constructed which rewarded the placement of a lipophilic atom above Tyr228 in the bottom of the S1 pocket (i.e., in a similar position to where the chlorine atom of the ligand is situated in the observed binding mode of compound **4** in thrombin described below). This reward term was added on to either the GoldScore or ChemScore functions depending on which function was being used.

Typically, multiple virtual screens were performed using different protein conformations, to select the best compounds. The conformations were chosen from relevant structures from the PDB.[30] The conformational changes usually related to side chains but sometimes included small backbone movements. The results from the virtual screening were also stored in Oracle and queried using a web-based interface. This interface allows the user to select subsets of compounds for interactive visualization using a variety of filters which include: simple physical properties such as molecular weight; scoring functions such as ChemScore, GoldScore, DrugScore,[31] predefined pharmacophores, etc.; components of scoring functions; steric or electrostatic clashes; the formation of specific hydrogen bonds; and 2D substructure. Molecular visualization is based on the Java molecular visualization program, AstexViewer.[32] Multiple ligands can be interactively visualized in the protein and/or the protein surface, and predicted binding modes can be compared with experimentally observed ones.

The focused kinase set was also constructed to contain motifs that are likely to bind to the ATP binding site of kinases. Analysis of the literature and patents identified a number of scaffolds that were often observed to bind in the ATP binding site of kinases and formed hydrogen bonds with the backbone 'hinge' region.[33,34] Enumeration of these scaffolds with drug-like side chains was used to construct virtual libraries of candidate fragments and available fragments were purchased from chemical suppliers. The set was augmented with compounds from virtual screening. The version of the focused kinase set screened in this work contained 116 compounds.

A focused phosphatase set was constructed that was targeted on the phospho-tyrosine binding pocket of PTP1B.[35,36] Compounds that were expected to contain more than 1 negative charge at normal pH were excluded from the phosphatase set. The literature was analyzed for potential phosphotyrosine and carboxylic acid mimetics, and medicinal chemistry experience was used to search commercial databases for potential fragments. Additionally, where promising fragments could not be accessed from commercial vendors, a small number of fragments were synthesized. To augment the set, virtual screening was performed on the open and closed conformations of PTP1B in order to identify novel scaffolds. The final size of the focused phosphatase set screened in this work contained 264 compounds.

Table 1. X-ray Data Collection Statistics

	P38 compd **1**	P38 compd **2**	CDK2 compd **3**	thrombin compd **4**	thrombin compd **5**	RNaseA compd **6**	PTP1B compd **7**
X-ray source/ detector	Inhouse/ Jupiter	ESRF	ESRF	Inhouse/ Jupiter	Inhouse/ Jupiter	Inhouse/ Jupiter	Inhouse/ Jupiter
space group	P212121	P212121	P212121	C2	C2	C2	*P*3121
resolution (Å)	2.2	2.1	1.8	2.04	2.2	1.8	2.2
R_{sym} (deg)	0.063	0.07	0.082	0.073	0.072	0.039	0.082
completeness (%)	94.4	98.7	96	99.1	97.6	100.0	96.7
a (Å)	45.77	45.76	53.51	70.5	69.39	100.03	88.1
b (Å)	87.10	84.72	72.02	71.51	71.78	32.6	88.1
c (Å)	126.34	126.76	72.47	72.2	71.73	72.32	104.5
β (deg)				100.459	99.87	90.7	

[a] $R_{sym}=\Sigma_h\Sigma_j|I_hj - <I_h>|/\Sigma_h\Sigma_j|I_{hj}|$ where h represents a unique reflection and j represents symmetry related indices.

Cocktailing. To increase the throughput of screening, the compounds were 'cocktailed' together, usually into groups of 4. A computational procedure was devised to minimize the chance of more than one compound in the same cocktail binding to the protein. This also allowed increased shape diversity in each cocktail that would aid the automated interpretation of ligand electron density by AutoSolve. The number of unique ways of partitioning N compounds into cocktails each containing c compounds is given by:

$$\frac{N!}{n!(c!)^n}$$

where $n = N/c$ is the number of cocktails. This number can increase rapidly, for example, there are 10^{860} unique ways of partitioning 400 compounds into 50 groups of 8. Therefore, any algorithm can sample only a small fraction of the possible partitioning space and needs to be extremely efficient to provide good solutions to the problem.

The first step in our approach to this problem is the derivation of a method for chemical similarity. The chemical functionalities were assessed by counting the number of hydrogen bond donors and acceptors, and ionizable groups. The size and shape of the molecule was assessed by counting the total number of non-hydrogen atoms, counting the number of five- and six-membered rings and distinguishing between substitution patterns on di-substituted aromatic rings. The dissimilarity of two molecules is calculated by finding the distance between the feature vectors of the two molecules.

The partitioning into cocktails proceeds from a matrix that describes the dissimilarity of each pair of molecules to be partitioned. The compounds are initially assigned to random cocktails. The cocktail score is calculated by summing the dissimilarity measures, $d(i,j)$, of all pairs of compounds in a particular cocktail, c. The overall score is calculated by adding the scores of all n cocktails.

$$S = \sum_{c=1,n} \sum_{i,j \in c} d(i,j)$$

This function is maximized using a procedure that swaps pairs of compounds in different cocktails. Swaps are accepted if the score improves or does not change; this second step was found to be very useful in preventing premature termination. The search was terminated after 10 000 000 iterations, or after 100 compound swaps that did not yield an improvement in the overall score. Before use, the algorithm was tested on a large partitioning problem that could be solved exhaustively, (i.e., 24 compounds into three cocktails of eight) and was found to rapidly produce solutions at or very close to the global minimum.

Experimental Section

Crystal Soaking. Crystals of p38 MAP kinase (PDB code: 1P38[37]), CDK2 (PDB code: 1HCK[38]), thrombin (PDB code: 1QJ1[39]), PTP1B (PDB code: 1C83[40]) and ribonuclease A (RNaseA) (PDB code: 1AFK[41]) were obtained following published protocols. In all cases the active site of the protein is accessible to compounds soaked into the crystal. The crystallographic screening process involves exposing protein crystals to solutions containing cocktails of between two and eight fragments. The small fragments are expected to have relatively low binding affinities for the protein, and thus a high concentration of compound is required in the soaking solution. The final concentration of individual compounds in the soaking solution ranged from 25 mM to 200 mM. To minimize crystal damage in such harsh conditions the protocols for soaking are optimized for each target prior to the screening experiment. The crystals are typically soaked at room temperature for 1−24 h before being cryo-cooled using liquid nitrogen and subsequently transferred to storage dewars until required for data collection.

X-ray Crystallography. X-ray diffraction data was collected for soaked crystals, either in-house or using synchrotron radiation sources (see Table 1) and processed using D*trek[42] or MOSFLM.[43] In-house data was collected using a Rigaku-MSC Jupiter CCD or an R-AXIS IV image plate detector, mounted on an RU−H3R rotating anode generator equipped with Osmic confocal multilayer 'Purple' and 'Blue' optics, respectively. Data collection in-house was typically 2−6 h, depending on the crystal system. To streamline the process of analyzing X-ray data for potential ligand binding, and to rapidly solve protein−ligand complexes, we invested significant effort into automating the processing, analysis and interpretation of the data. Briefly, this involved an automated refinement and interpretation procedure, starting with a model of the apo-protein (obtained from the PDB structures indicated for each protein) and a set of X-ray diffraction data collected from crystals grown in-house. The starting models were subjected to an automated procedure, in which an initial molecular replacement step was performed using AMORE;[44] the best solution being subjected to rigid-body refinement, followed by restrained positional and isotropic displacement parameter refinement using CNX[45] or REFMAC5.[46] A sigma-weighted difference electron density map was computed, and automatic water placement performed, excluding a defined volume in the binding site. This refinement/water placement process was iterated until the free R-factor at convergence of each refinement job no longer decreased, when the final weighted difference electron density map was computed. The refinement statistics for the complexes described in this paper are shown in Table 2, and the complexes have been deposited with the PDB.

The difference electron density maps were analyzed and interpreted by AutoSolve, which is a procedure that automatically fits ligand structures to difference maps. AutoSolve starts by identifying appropriate regions of unoccupied electron density in the map obtained from the refinement stage. The program CORINA[28] is used to generate a 3D model for each

Table 2. Refinement Statistics for the Protein Ligand Complexes

	P38 compd **1**	P38 compd **2**	CDK2 compd **3**	thrombin compd **4**	thrombin compd **5**	RNaseA compd **6**	PTP1B compd **7**
resolution (Å)	17.65−2.53	50.75−2.16	50−1.7	50.0−2.04	27.2−2.2	19.67−1.9	44−2.23
R_{cryst}[a]	0.19	0.223	0.21	0.21	0.21	0.22	0.27
R_{free}[b]	0.24	0.258	0.25	0.27	0.28	0.27	0.31
RMS deviation bond length (Å)	0.005	0.006	0.014	0.006	0.006	0.008	0.006
RMS deviation bond angle (deg)	0.85	1.19	1.39	1.27	1.29	0.99	1.31
PDB code	1w7h	1wbo	1wcc	1way	1wbg	1wbu	1wax

[a] $R_{cryst}=\Sigma||F_{obs}| - |F_{calc}||/\Sigma_h|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes respectively for reflection h. [b] R_{free} is equivalent to R_{cryst} but is calculated for 5% of the reflections excluded from the refinement.

ligand that was present in the cocktail. Low energy conformations of each ligand structure are then fitted to the unoccupied electron density. The ligand conformer with the best fit to the density is recorded as being that most likely to be bound in the crystal structure. The resulting protein/ligand complexes are then used as the basis for the structure-guided fragment optimization process.

Enzyme Inhibition Assays. p38 Assay. The biological activities for inhibitors of p38α were determined by a radiometric enzyme assay. Compounds were incubated with activated p38α and 5 μg of myelin basic protein substrate (Upstate) in 25 μL of 25 mM HEPES pH 7.4, 25 mM β-glycerophosphate, 5 mM EDTA, 15 mM $MgCl_2$, 70 μM ATP, 1 mM sodium orthovanadate, 1 mM DTT, 10% DMSO, and 0.35 μCi of ^{33}P-γ-ATP (AP Biotech) for 1 h at room temperature. The reaction was stopped by the addition of 30 μL of 2% orthophosphoric acid and transferred to Millipore MAPH filter plates, prewetted with 50 μL of 0.2% orthophosphoric acid. The plates were filtered and washed twice with 200 μL of 0.2% orthophosphoric acid. Incorporated radioactivity was measured by scintillation counting. IC_{50}s were calculated from replicate curves using GraphPad Prizm software, and standard errors and *P* values were within statistically acceptable limits.

CDK2 Assay. Compounds were incubated with CDK2/cyclinA active complex (Upstate) and 3 μg of histone H1 substrate in 25 μL of 20 mM MOPS pH 7.2, 25 mM β-glycerophosphate, 1 mM DTT, 5 mM EDTA, 15 mM $MgCl_2$, 45 μM ATP, 1 mM sodium orthovanadate 2.5% DMSO, and 0.35 μCi of ^{33}P-γ-ATP (AP Biotech) for 5 h at room temperature. Reactions were terminated by the addition of 5 μL of 2% orthophosphoric acid. The plates were filtered and washed twice with 200 μL of 0.2% orthophosphoric acid. Incorporated radioactivity was measured by scintillation counting. IC_{50}s were calculated from replicate curves using GraphPad Prizm software, and standard errors and *P* values were within statistically acceptable limits.

PTP1B Assay. Compounds were incubated with PTP1B and 0.25 mM *p*-nitrophenyl phosphate substrate (Sigma), in 50 mM HEPES pH 6.5, 1 mM DTT, 1 mM EDTA, 0.01% CHAPS, 5% DMSO in 1/2 area 96-well plates. The dephosphorylation of *p*-nitrophenyl phosphate was monitored by following the change in absorbance at 405 nm over 30 min in a Molecular Devices Spectramax plate reader at room temperature. Initial reaction rates were measured and IC_{50}s were calculated from replicate curves using GraphPad Prizm software and standard errors and *P* values were within statistically acceptable limits.

Thrombin Assay. Compounds were incubated with activated thrombin and 14 μM fluorogenic substrate, BOC-Val-Pro-Arg-MCA (Bachem), in 50 mM Tris pH 8, 5 mM $CaCl_2$, 100 mM NaCl, 5% DMSO in 96-well plates. The cleavage of the substrate was followed by monitoring the change in fluorescence at 460 nm (excitation at 365 nm) for 25 min at room temperature on a Packard Fusion plate reader. Initial reaction rates were measured and IC_{50}s were calculated from replicate curves using GraphPad Prizm software, and standard errors and *P* values were within statistically acceptable limits.

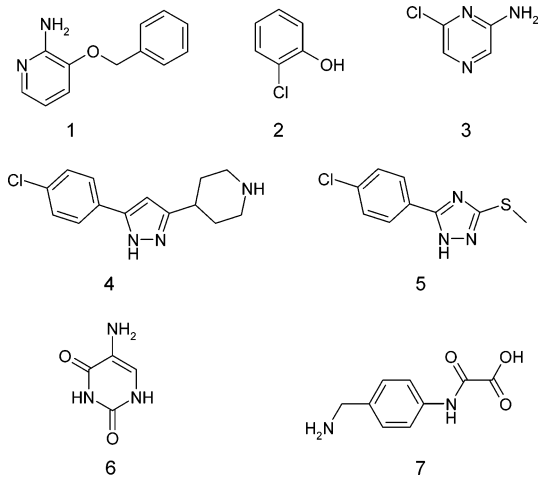

Figure 5. Structures for the Pyramid hits discussed in the text.

Results and Discussion

We have applied our Pyramid fragment screening process to a range of targets including kinases, proteases as well as other enzymes. Here we describe the results from selected examples using p38 MAP kinase, CDK2, thrombin, RNase A and PTP1B.

Example 1: Compounds 1 and 2 Binding to p38 MAP Kinase. P38 MAP kinase is a key modulator in the TNF pathway and as such a potential target for the development of inhibitors of inflammatory disease.[47] Pyramid screening identified several fragment hits such as, compounds **1** and **2**, both of which have clearly defined binding modes but affinities greater than 1mM in an enzyme bioassay (Figures 5 and 6).

The aminopyridine of **1** forms two hydrogen bonds to the hinge region in the ATP binding site of the kinase (Figure 6a). Additionally, the benzyl group binds into an adjacent lipophilic pocket, often utilized by larger p38 kinase inhibitors. The recognition is driven by a combination of the hinge hydrogen bonds and the occupancy of this key lipophilic pocket. Compound **1** is an attractive starting point for medicinal chemistry and the following paper[48] gives details of how this fragment was optimized to nanomolar inhibitors. Figure 6c shows compound **1** superimposed on the structure of an optimized inhibitor described in the subsequent paper. The larger inhibitor causes a substantial shift in the DFG portion (i.e. residues 168−170) of p38 MAP kinase; despite this, the overlay of compound **1** on the large inhibitor is relatively good and shows conservation of the key interactions.

The chlorophenyl portion of compound **2** (Figure 6b) occupies the key lipophilic pocket of the kinase and the

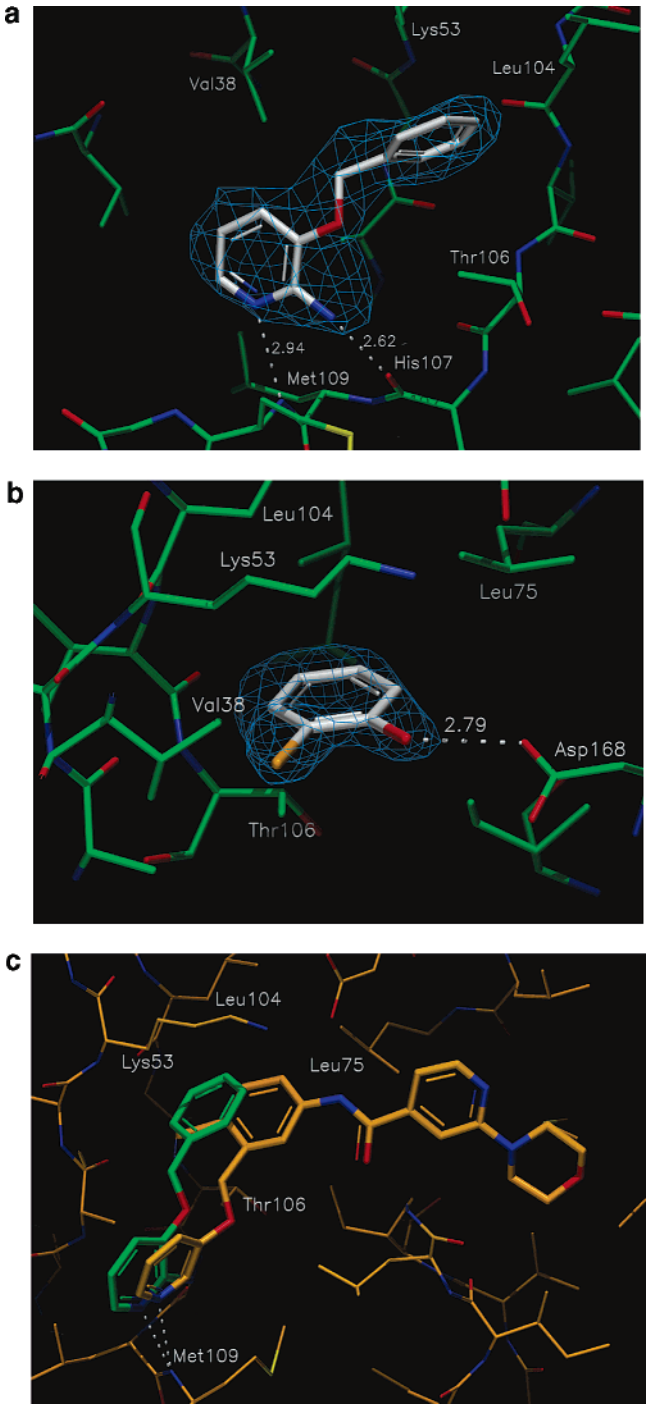

Figure 6. (a) Compound **1** (thick stick) is shown bound to p38 MAP kinase (thin stick). The $F_o - F_c$ electron density map contoured at 3σ is also shown. The maps in this and subsequent figures have been clipped to aid visualization. The fragment binds to the hinge region forming hydrogen bonds with the N−H of Met109 and the C=O of His107. The phenyl ring is bound in a lipophilic pocket formed by the side chains of Lys53, Leu75 (omitted for clarity), Leu104 and Thr106. (b) Compound **2** (thick stick) is shown bound to p38 MAP kinase (thin stick). The $F_o - F_c$ map contoured at 3σ is also shown. The chlorophenyl ring is bound in the lipophilic pocket formed by the side chains of Lys53, Leu75, Leu104 and Thr106. The phenol makes a hydrogen bond to the side chain of Asp168, which has moved over 2 Å to form this interaction. (c) Compound **1** (green) is shown superimposed on a subsequent lead compound (orange) bound to p38 MAP kinase. The latter compound induces a substantial movement in the kinase but the overlap shows that the key fragment interactions are retained in the larger molecule (i.e., hydrogen bond to N−H of Met109 and occupation of the lipophilic pocket defined by the side chains of Lys53, Leu75, Leu104 and Thr106).

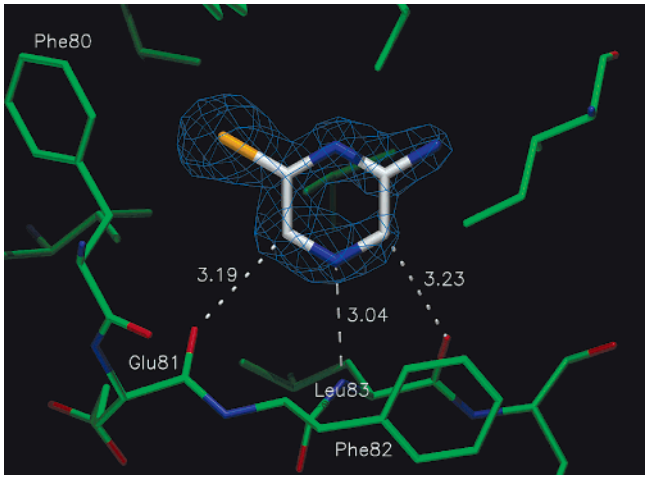

Figure 7. Compound **3** (thick stick) is shown bound to CDK2 (thin stick). The $F_o - F_c$ map contoured at 3σ is also shown. The heterocycle binds at the ATP binding cleft forming a good uncharged hydrogen bond with the N−H of Leu83. There are also strong C−H aromatic contacts with the carbonyls of Leu83 and Glu81. The ligand makes lipophilic contact with Ala31 (omitted for clarity), Phe80, Phe82 and Leu134 (unlabeled, directly behind ligand).

fragment does not form hydrogen bonds to the hinge region of the enzyme. This confirms that it is possible to probe key selectivity pockets around the ATP binding site without interacting with the highly conserved hinge region. The phenolic OH on compound **2** forms a hydrogen bond to the side chain of Asp168, which is an unusual interaction for p38 inhibitors. The Asp 168 side chain has moved over 2.0 Å (in comparison to the compound **1** structure) to form this hydrogen bond and adopts a position not seen in any other public domain p38 complexes. Indeed, only one p38 complex in the public domain (PDB code: 1OUK[49]) shows a (long) hydrogen bond to this side chain. It is our view that the recognition for compound **2** is driven mainly by the lipophilic interactions and this example demonstrates that the fragment screening method is capable of probing key lipophilic interactions, as well as inducing conformational movement as seen with Asp 168.

Example 2: Compound 3 Binding to CDK2. CDK2 is a Ser/Thr protein kinase involved in the cell cycle, and inhibitors of CDK2 are thought to have potential in anticancer therapy.[50−52] Pyramid screening identified several diverse fragment hits for CDK2 such as compound **3**, a pyrazine-based fragment. Compound **3** is clearly defined in the electron density (see Figures 5 and 7) and has a measured affinity of 350 μM in an enzyme assay.

Compound **3** forms only one hydrogen bond with the protein, between the pyrazine nitrogen and the amide of Leu83 in the hinge region of the ATP binding site. The adjacent aromatic C−H's on the pyrazine also form favorable electrostatic interactions with the backbone carbonyl oxygens of Leu83 and Glu81. These "nonclassical" hydrogen bonds with the hinge region are often observed with larger kinase inhibitors containing electron deficient aromatic groups and further exemplify the range of interactions that can be utilized by fragments.[53] Compound **3** also forms additional hydrophobic interactions with the side chains of Ala31, Phe80, Phe82 and Leu134. Given the importance of the observed hydrogen

bonds to many known kinase inhibitors, it is likely binding is mainly driven by the formation of these interactions. From a medicinal chemistry point of view, compound **3**, with its associated binding mode, is an attractive starting point for the design of novel lead compounds, and progress on the design of CDK2 compounds derived from **3** will be reported in due course.

Example 3: Compounds 4 and 5 Binding to Thrombin. Thrombin is a serine protease involved in the blood coagulation cascade, and inhibitors are potentially useful as anticoagulants.[54] Pyramid fragment screening against thrombin identified compounds **4** and **5** as hits (see Figures 5 and 8). The compounds have affinities of 400 μM and approximately 1 mM, respectively, in enzyme bioassays.

Figure 8a shows the chlorophenyl of compound **4** deeply buried in the S1 pocket with the chlorine atom located above the π-face of Tyr228. The detection of an uncharged S1 binder is a key discovery and provides further validation of Pyramid's ability to detect fragments driven primarily by hydrophobic binding to key enzyme pockets. Additional interactions include a hydrogen bond between the pyrrole N−H and the highly flexible side chain of Glu192. The terminal portion of the piperidine group is not visible in the initial electron density map, perhaps indicating greater flexibility for this portion of the molecule. Indeed, plausible positions for this charged terminal portion of the piperidine indicate that it is positioned outside the enzyme cavity. It seems likely the piperidine group is important for increasing the solubility of the compound, rather than forming important interactions with the enzyme. The structure for compound **5** (Figure 8b) is similar with recognition dominated by lipophilic contact in the S1 pocket. There is also an additional hydrogen bond between the N−H on the triazole and the carbonyl of Gly219.

Figure 8c shows compound **4** superimposed on the structure of Factor Xa complexed to a nanomolar inhibitor (PDB code: 1MQ5[55]) and clearly indicates that the chlorophenyl fragment has an almost identical binding mode in the S1 pocket as compared to a larger, high affinity inhibitor. Compound **5** also exhibits an identical binding mode in the S1 pocket. Although there are no thrombin crystal structures in the public domain that have a chlorophenyl moiety in S1, structures have been described in the literature and they are consistent with the binding modes of the chlorophenyl fragments presented here.[56,57] The eventual replacement of benzamidine mimics by uncharged lipophiles in thrombin and Factor Xa inhibitors was an important development in the search for novel anticoagulant agents and could have been accelerated had fragment screening approaches been applied.

Example 4: Compound 6 Binding to Ribonuclease A. Ribonuclease A is a digestive enzyme responsible for cleaving single stranded RNA and contains a flat, extended active site similar to drug targets that are often difficult to prosecute in lead generation. Crystallographic screening of a targeted fragment set identified compound **6**, whose affinity was not measured but is expected to be very weak as exceptionally high concentrations of compound **6** (i.e., 200 mM) were required in order to observe binding in the crystal.

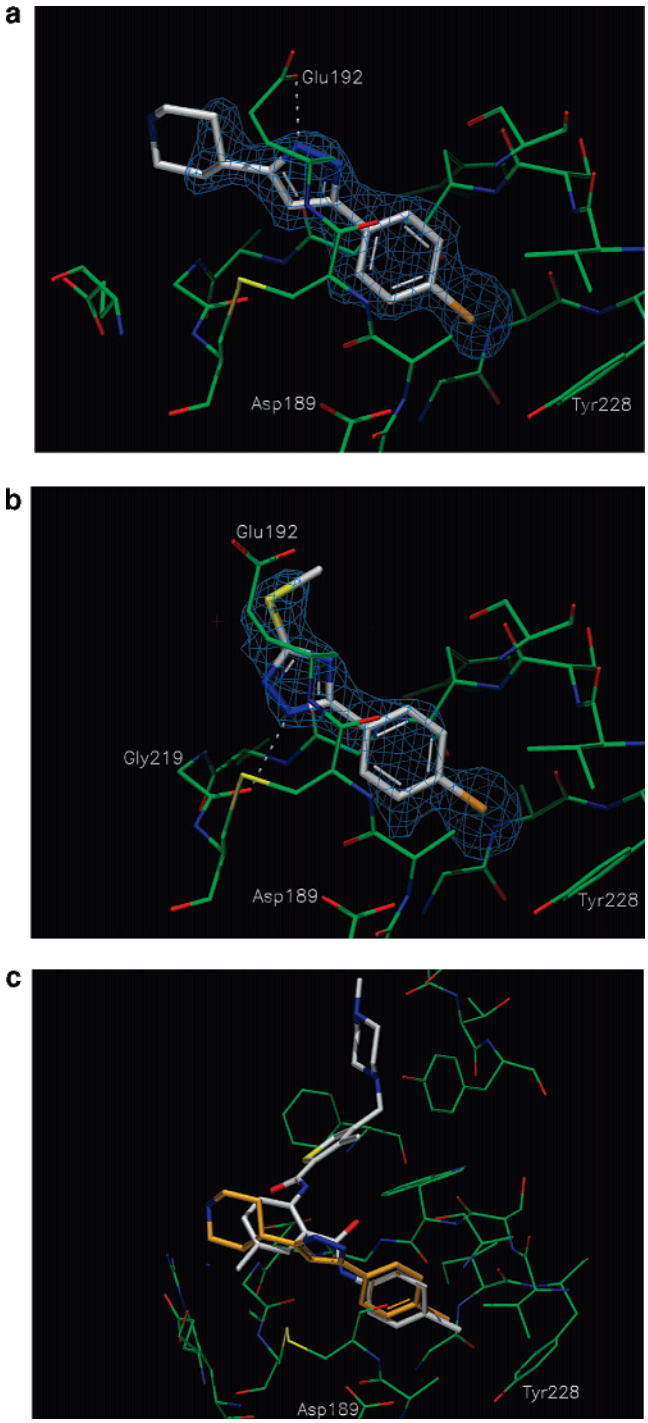

Figure 8. (a) Compound **4** (thick stick) is shown bound to thrombin (thin stick). The $F_o - F_c$ map contoured at 3σ is also shown. The key interactions are the chlorophenyl group buried deep in the hydrophobic portion of the S1 pocket above Tyr228. There is also a hydrogen bond between the N−H of the pyrazole and the flexible side chain of Glu192. The piperidine forms no significant contacts with the enzyme and is disordered, although its presence probably helps solubilize the compound. (b) Compound **5** (thick stick) is shown bound to thrombin (thin stick). The $F_o - F_c$ map contoured at 3σ is also shown. The key interactions are the chlorophenyl group buried deep in the hydrophobic portion of the S1 pocket above Tyr228. There is also a hydrogen bond between the N−H of the triazole and the backbone carbonyl of Gly219. (c) Compound **4** (orange) is shown superimposed on a Factor Xa complex (PDB code: 1mq5) with the protein carbon atoms shown in green and the ligand carbon atoms in white. The overlay of the chlorophenyl group of the fragment is good when compared to the chlorophenyl in the potent Factor Xa ligand.

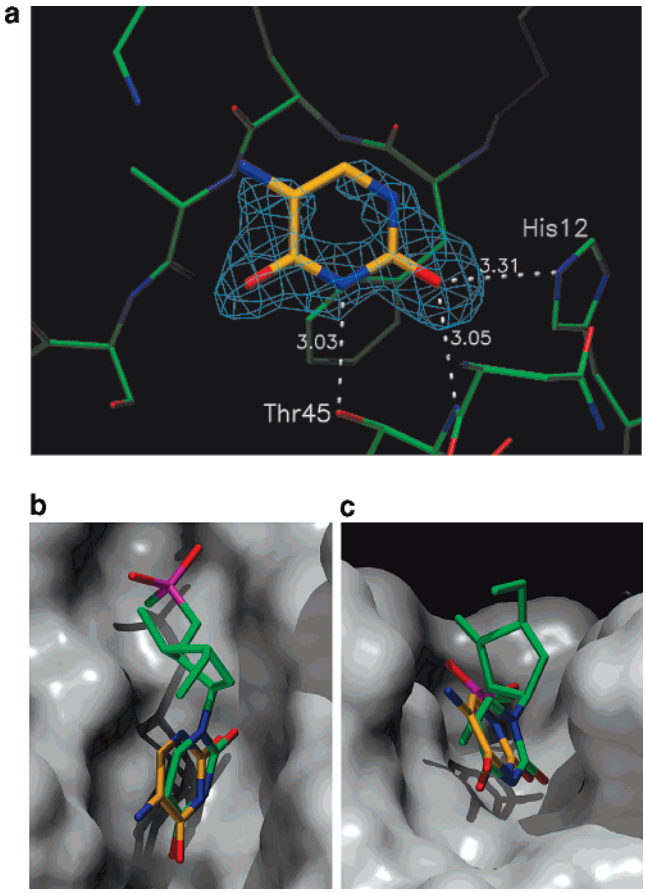

Figure 9. (a) Compound **6** (thick stick) is shown bound to Rnase A (thin stick). The $F_o - F_c$ map contoured at 1.7σ is also shown. The heterocycle forms good uncharged hydrogen bonds with the backbone and side chain of Thr45 and with the side chain of His12. (b and c). Two orientations of compound **6** (orange) superimposed onto a Rnase A ligand in green (PDB file: 1EOS in green). The fragment has good overlap with the relevant key groups in the larger ligand. The solvent accessible surface indicates the open and extended active site for this enzyme.

Compound **6** is observed to bind with low occupancy but is clearly defined in the electron density (Figure 9a) and forms good hydrogen bonds to the backbone NH of Thr45 and the side chains of Thr45 and His12. These hydrogen bonding interactions are also observed during binding of larger, known ligands[58] which provides extra confidence in the interpretation of the binding mode of compound **6** (Figures 9b and 9c). This example highlights the potential of fragment screening for targets with less defined active sites.

Example 5: Compound 7 Binding to PTP1B. Protein tyrosine phosphatase 1B (PTP1B) is a drug target for obesity and type II diabetes.[59] Compound **7** was identified as a hit against PTP1B from our focused phosphatase set, and its structure and binding mode are shown in Figure 5 and Figure 10. The affinity of the compound is 86 μM as measured in an enzyme bioassay. The fragment shows multiple hydrogen bonds to the protein as detailed in Figure 10. PTP1B generally recognizes doubly anionic groups in the phosphate binding pocket so it is significant that compound **7** contains only one anionic functional group. Furthermore, compound **7** induces a significant conformational movement in the enzyme relative to the apo structure

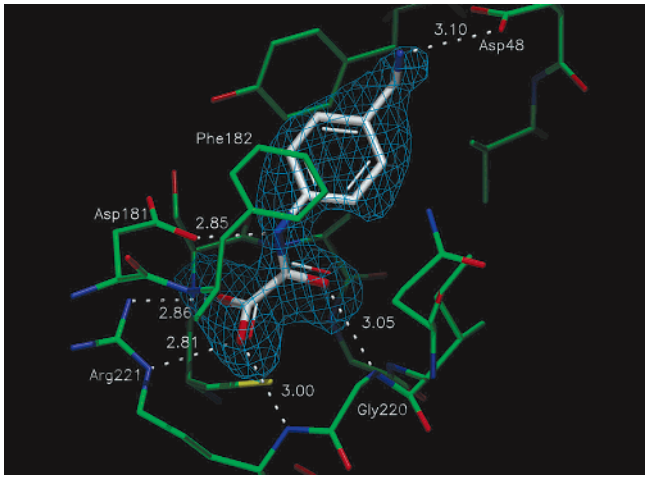

Figure 10. Compound **7** (thick stick) is shown bound to PTP1B (thin stick). The $F_o - F_c$ map contoured at 3σ is also shown. The carboxylate forms a salt bridge with the side chain of Arg221 and also with Arg221's backbone N−H. The ligand carbonyl forms a hydrogen bond with the backbone of Gly220 while the ligand N−H forms a hydrogen bond with the side chain of Asp181. The amine on the ligand also forms a salt bridge with Asp48. The ligand induces a movement of the "WPD loop" of the enzyme allowing the formation of additional interactions (only the final conformation of the "WPD" loop is shown).

with Phe182 of the WPD loop (i.e. residues 179−181 and adjacent residues) moving over 10 Å to make contact with the ligand. It should be noted that this movement of the WPD loop has been observed before with a number of other PTP1B inhibitors including some fragment-sized molecules.[60] Given its reduced anionic nature, compound **7** may prove a useful starting point toward the design of compounds with superior physical properties and pharmacokinetic parameters.

Discussion

These examples highlight the range of protein−ligand interactions that can be identified using a fragment screening approach. We have shown that molecular recognition can be driven primarily by lipophilic interactions (thrombin and p38), uncharged hydrogen bonds (RNase A, CDK2), or charged hydrogen bonds (PTP-1B) as well as more balanced combinations of interactions (e.g., the binding of compound **1** to p38). We have also illustrated that both significant (PTP1B) and more subtle (p38) protein movements can be observed when fragments bind. There is thus no obvious general restriction to the types of interactions that can be observed with fragment screening although the particular molecular interactions or protein movements will be governed by the energetics of the binding event. Typical hit rates observed in the above fragment screening examples range from 0.5 to 10% and reflect the fact that although the fragment libraries contain only 500−1000 compounds, they are able to sample significant chemical space. The efficient sampling is a feature of most fragment-based methods and has been reported by several groups.[6,14,20]

In all of the examples in this study, despite the low binding affinity, the fragment hits are clearly defined in the electron density. This, together with the observation that during the hit to lead optimization phase (a

process that we call 'fragment evolution') the original interactions are maintained, suggests that these first interactions between the fragment and protein are energetically significant. Furthermore, the low affinity of the fragments does not theoretically imply weak binding, as it has been suggested that when ligands bind to proteins, they must overcome a rigid body entropic barrier that is about 3 orders of magnitude in size and is independent of molecular weight.[61,62] This implies that the limited interactions formed by a fragment must be of very 'high quality' to overcome this barrier. We propose that fragments, due to their functional simplicity and low MW, are 'efficient' binders. The challenge in the fragment evolution process is to transform these low affinity binders into potent lead compounds while maintaining high ligand efficiency.

There is significant literature investigating the properties of small molecules that are required to make good lead compounds.[2,3] The now familiar Lipinski's 'Rule of 5' provides a useful framework for developing orally bioavailable drug candidates,[63] and more recently, the term 'leadlike' was introduced for molecules identified from HTS campaigns that are suitable for further optimization.[3] However, all of these studies address the issues facing compounds discovered using conventional bioassay-based screening of drug-size compound libraries. Our experience of screening low molecular weight compounds suggests that a different set of rules may apply to fragments. We recently proposed 'Rule of 3' filters for the construction of fragment libraries[27] in which molecular weight (MWT) is <300, hydrogen bond donors (HBD) ≤ 3, hydrogen bond acceptors (HBA) ≤ 3, and ClogP ≤3. In addition, number of rotational bonds (NROT) ≤ 3 and polar surface area (PSA) ≤ 60 might also be useful limits for fragment selection.

Although there are several technical hurdles to overcome when using X-ray crystallography for fragment screening, such as obtaining suitable protein crystals, it also has key advantages over other techniques. While several other methods may identify low affinity binding fragments, their transformation into potent lead compounds while maintaining optimal ligand efficiency is uniquely enabled by knowledge of the precise binding orientations of the fragment, and subsequent analogues, from X-ray crystal structures. These co-crystal structures can also identify nearby water molecules as well as conformational movement of the protein, both aspects potentially exploitable in drug design. Furthermore, although false-negative hits cannot be discounted, the occurrence of false-positive hits is near impossible using crystallographic screening. This aspect particularly avoids being distracted by fragments that bind in a nonspecific manner, a key issue in low affinity screening. The fragment hits can also be immediately assessed for functional relevance based on whether they bind at an active site or an allosteric pocket by examining the crystal structure. Finally, similar to some other fragment-based methods, the protein may be screened either in an active or inactive state as there is no reliance on a biological readout. This may offer alternative strategies for development of lead compounds such as targeting inactive forms of protein kinases.

We believe that the use of structure-based drug design starting from highly efficient fragments is a promising approach to the development of efficient lead molecules. Such lead molecules may offer the best opportunities for the development of drugs with good physicochemical and pharmacokinetic properties as well as the necessary potency and efficacy. The full potential of fragment-based lead generation using X-ray crystallography for discovering novel lead compounds is likely to emerge in the coming years.

Acknowledgment. The authors gratefully acknowledge the support of many people at Astex Technology who have contributed to some aspects of the work presented in this paper. We thank Owen Callaghan and Miles Congreve for the synthesis of compounds, Andrew Sharff, Lindsay Devine, Rob van Montfort, Marc O'Reilly, Dominic Tisi, Wendy Blakemore, Wijnand Mooij, and Jasmine Tickle for assistance with crystallography, and Nicola Wallace, Lisa Magor, and Debbie Davis for enzyme assays. We acknowledge ESRF Grenoble and SRS Daresbury for access to beamtime and technical assistance.

References

(1) Campbell, S. F. Science, art and drug discovery: a personal perpsective. *Clin. Sci.* **2000**, *99*, 255−260.
(2) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308−1315.
(3) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. The Design of Leadlike Combinatorial Libraries. *Angew. Chem., Int. Ed.* **1999**, *38*, 3743−3748.
(4) Lesuisse, D.; Lange, G.; Deprez, P.; Benard, D.; Schoot, B.; Delettre, G.; Marquette, J. P.; Broto, P.; Jean-Baptiste, V.; Bichet, P.; Sarubbi, E.; Mandine, E. SAR and X-ray. A new approach combining fragment-based screening and rational drug design: application to the discovery of nanomolar inhibitors of Src SH2. *J. Med. Chem.* **2002**, *45*, 2379−2387.
(5) Carr, R. and Jhoti, H. Structure-based screening of low-affinity compounds. *Drug Discovery Today* **2002**, *7*, 522−527.
(6) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. Site-directed ligand discovery. *Proc. Natl. Acad. Sci. U. S. A* **2000**, *97*, 9367−9372.
(7) Nienaber, V. L.; Richardson, P. L.; Klighofer, V.; Bouska, J. J.; Giranda, V. L.; Greer, J. Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat. Biotechnol.* **2000**, *18*, 1105−1108.
(8) Sanders, W. J.; Nienaber, V. L.; Lerner, C. G.; McCall, J. O.; Merrick, S. M.; Swanson, S. J.; Harlan, J. E.; Stoll, V. S.; Stamper, G. F.; Betz, S. F.; Condroski, K. R.; Meadows, R. P.; Severin, J. M.; Walter, K. A.; Magdalinos, P.; Jakob, C. G.; Wagner, R.; Beutel, B. A. Discovery of potent inhibitors of dihydroneopterin aldolase using CrystaLEAD high-throughput X-ray crystallographic screening and structure-directed lead optimization. *J. Med. Chem.* **2004**, *47*, 1709−1718.
(9) Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, G. M., Jr.; Marcotte, P. A.; Severin, J.; Walter, K.; Smith, H.; Gubbins, E.; Simmer, R.; Holzman, T. F.; Morgan, D. W.; Davidsen, S. K.; Summers, J. B.; Fesik, S. W. Discovery of Potent Nonpeptide Inhibitors of Stromelysin Using SAR by NMR. *J. Am. Chem. Soc.* **2004**, *119*, 5818−5827.
(10) Boehm, H. J.; Boehringer, M.; Bur, D.; Gmuender, H.; Huber, W.; Klaus, W.; Kostrewa, D.; Kuehne, H.; Luebbers, T.; Meunier-Keller, N.; Mueller, F. Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. *J. Med. Chem.* **2000**, *43*, 2664−2674.
(11) Lange, G.; Lesuisse, D.; Deprez, P.; Schoot, B.; Loenze, P.; Benard, D.; Marquette, J. P.; Broto, P.; Sarubbi, E.; Mandine, E. Requirements for specific binding of low affinity inhibitor fragments to the SH2 domain of (pp60)Src are identical to those for high affinity binding of full length inhibitors. *J. Med. Chem.* **2003**, *46*, 5184−5195.
(12) van Dongen, M.; Weigelt, J.; Uppenberg, J.; Schultz, J.; Wikstrom, M. Structure-based screening and design in drug discovery. *Drug Discovery Today* **2002**, *7*, 471−478.

(13) van Dongen, M. J.; Uppenberg, J.; Svensson, S.; Lundback, T.; Akerud, T.; Wikstrom, M.; Schultz, J. Structure-based screening as applied to human FABP4: a highly efficient alternative to HTS for hit generation. *J Am. Chem. Soc.* **2002**, *124*, 11874−11880.
(14) Fejzo, J.; Lepre, C. A.; Peng, J. W.; Bemis, G. W.; Ajay; Murcko, M. A.; Moore, J. M. The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. *Chem. Biol.* **1999**, *6*, 755−769.
(15) Maly, D. J.; Choong, I. C.; Ellman, J. A. Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2419−2424.
(16) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430−431.
(17) Blundell, T.L.; Abell, C.; Cleasby, A.; Hartshorn, M.J.; Tickle, I.J.; Parasini, E.; Jhoti, H. High-throughput X-ray crystallography for drug discovery. In *Drug design: special publication;* Flower, D. R., Ed.; Royal Society of Chemistry: Cambridge, UK, 2002; pp 53−59.
(18) Muchmore, S. W.; Olson, J.; Jones, R.; Pan, J.; Blum, M.; Greer, J.; Merrick, S. M.; Magdalinos, P.; Nienaber, V. L. Automated crystal mounting and data collection for protein crystallography. *Struct. Fold. Des.* **2000**, *8*, R243−R246.
(19) Sharff, A. J. High throughput crystallography on an in-house source, using ACTOR. *Rigaku J.* **2004**, *20*, 10.
(20) Hann, M. M.; Leach, A. R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856−864.
(21) Bemis, G. W. and Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* **1996**, *39*, 2887−2893.
(22) Bemis, G. W. and Murcko, M. A. Properties of known drugs. 2. Side chains. *J. Med. Chem.* **1999**, *42*, 5095−5099.
(23) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43−53.
(24) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(25) Verdonk, M. L.; Cole, J. C.; Hartshorn, M.; Murray, C. W.; Taylor, R. D. Improved protein−ligand docking using GOLD. *Proteins* **2003**, *52*, 609−623.
(26) Watson, P.; Verdonk, M. L.; Hartshorn, M. J. A web-based platform for virtual screening. *J. Mol. Graph. Model.* **2003**, *22*, 71−82.
(27) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'rule of three' for fragment-based lead discovery? *Drug Discovery Today* **2003**, *8*, 876−877.
(28) Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic Generation of 3D-Atomic Coordinates for Organic Molecules. *Tetrahedron Comput. Methodol.* **2004**, *3*, 537−547.
(29) Baxter, C. A.; Murray, C. W.; Clark, D. E.; Westhead, D. R.; Eldridge, M. D. Flexible docking using Tabu search and an empirical estimate of binding affinity. *Proteins* **1998**, *33*, 367−382.
(30) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(31) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based scoring function to predict protein−ligand interactions. *J. Mol. Biol.* **2000**, *295*, 337−356.
(32) Hartshorn, M. J. AstexViewer: a visualisation aid for structure-based drug design. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 871−881.
(33) Gum, R. J.; McLaughlin, M. M.; Kumar, S.; Wang, Z.; Bower, M. J.; Lee, J. C.; Adams, J. L.; Livi, G. P.; Goldsmith, E. J.; Young, P. R. Acquisition of sensitivity of stress-activated protein kinases to the p38 inhibitor, SB 203580, by alteration of one or more amino acids within the ATP binding pocket. *J. Biol. Chem.* **1998**, *273*, 15605−15610.
(34) Dreyer, M. K.; Borcherding, D. R.; Dumont, J. A.; Peet, N. P.; Tsay, J. T.; Wright, P. S.; Bitonti, A. J.; Shen, J.; Kim, S. H. Crystal structure of human cyclin-dependent kinase 2 in complex with the adenine-derived inhibitor H717. *J. Med. Chem.* **2001**, *44*, 524−530.
(35) Groves, M. R.; Yao, Z. J.; Roller, P. P.; Burke, T. R., Jr.; Barford, D. Structural basis for inhibition of the protein tyrosine phosphatase 1B by phosphotyrosine peptide mimetics. *Biochemistry* **1998**, *37*, 17773−17783.
(36) Jia, Z.; Ye, Q.; Dinaut, A. N.; Wang, Q.; Waddleton, D.; Payette, P.; Ramachandran, C.; Kennedy, B.; Hum, G.; Taylor, S. D. Structure of protein tyrosine phosphatase 1B in complex with inhibitors bearing two phosphotyrosine mimetics. *J. Med. Chem.* **2001**, *44*, 4584−4594.
(37) Wang, Z.; Harkins, P. C.; Ulevitch, R. J.; Han, J.; Cobb, M. H.; Goldsmith, E. J. The structure of mitogen-activated protein kinase p38 at 2.1-A resolution. *Proc. Natl. Acad. Sci. U. S. A* **1997**, *94*, 2327−2332.
(38) Schulze-Gahmen, U.; De Bondt, H. L.; Kim, S. H. High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design. *J. Med. Chem.* **1996**, *39*, 4540−4546.
(39) Jhoti, H.; Cleasby, A.; Reid, S.; Thomas, P. J.; Weir, M.; Wonacott, A. Crystal structures of thrombin complexed to a novel series of synthetic inhibitors containing a 5,5-trans-lactone template. *Biochemistry* **1999**, *38*, 7969−7977.
(40) Andersen, H. S.; Iversen, L. F.; Jeppesen, C. B.; Branner, S.; Norris, K.; Rasmussen, H. B.; Moller, K. B.; Moller, N. P. 2-(oxalylamino)-benzoic acid is a general, competitive inhibitor of protein-tyrosine phosphatases. *J. Biol. Chem.* **2000**, *275*, 7101−7108.
(41) Leonidas, D. D.; Shapiro, R.; Irons, L. I.; Russo, N.; Acharya, K. R. Crystal structures of ribonuclease A complexes with 5′-diphosphoadenosine 3′-phosphate and 5′-diphosphoadenosine 2′-phosphate at 1.7 A resolution. *Biochemistry* **1997**, *36*, 5578−5588.
(42) Pflugrath, J. W. The finer things in X-ray diffraction data collection. *Acta Crystallogr. D.* **1999**, *55 (Pt. 10)*, 1718−1725.
(43) Leslie, A. G. W.; Brick, P.; Wonacott, A. MOSFLM. *Daresbury Lab. Inf. Quart. Protein Crystallogr.* **2004**, *18*, 33−39.
(44) Navaza, J. *AMoRe*: an automated package for molecular replacement. *Acta Crystallogr. A* **2004**, *A50*, 157−163.
(45) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D.* **1998**, *54 (Pt. 5)*, 905−921.
(46) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D.* **2004**, *D53*, 240−255.
(47) Salituro, F. G.; Germann, U. A.; Wilson, K. P.; Bemis, G. W.; Fox, T.; Su, M. S. Inhibitors of p38 MAP kinase: therapeutic intervention in cytokine-mediated diseases. *Curr. Med. Chem.* **1999**, *6*, 807−823.
(48) Gill, A. L.; Frederickson, M.; Cleasby, A.; Woodhead, S. J.; Carr, M. G.; Woodhead, A. J.; Walker, M. T.; Congreve, M. S.; Devine, L. A.; Tisi, D.; O'Reilly, M.; Seavers, L. C. A.; Davis, D. J.; Curry, J.; Anthony, R.; Padova, A.; Murray, C. W.; Carr, R. A. E.; Jhoti, H. Identification of Novel p38α MAP Kinase Inhibitors using Fragment-Based Lead Generation. *J. Med. Chem.* **2005**, *48*, 414−426.
(49) Fitzgerald, C. E.; Patel, S. B.; Becker, J. W.; Cameron, P. M.; Zaller, D.; Pikounis, V. B.; O'Keefe, S. J.; Scapin, G. Structural basis for p38alpha MAP kinase quinazolinone and pyridol-pyrimidine inhibitor specificity. *Nat. Struct. Biol.* **2003**, *10*, 764−769.
(50) Fischer, P. M. and Lane, D. P. Inhibitors of cyclin-dependent kinases as anti-cancer therapeutics. *Curr. Med. Chem.* **2000**, *7*, 1213−1245.
(51) Knockaert, M.; Greengard, P.; Meijer, L. Pharmacological inhibitors of cyclin-dependent kinases. *Trends Pharmacol. Sci.* **2002**, *23*, 417−425.
(52) Sausville, E. A.; Zaharevitz, D.; Gussio, R.; Meijer, L.; Louarn-Leost, M.; Kunick, C.; Schultz, R.; Lahusen, T.; Headlee, D.; Stinson, S.; Arbuck, S. G.; Senderowicz, A. Cyclin-dependent kinases: initial approaches to exploit a novel therapeutic target. *Pharmacol. Ther.* **1999**, *82*, 285−292.
(53) Pierce, A. C.; Sandretto, K. L.; Bemis, G. W. Kinase inhibitors and the case for CH···O hydrogen bonds in protein−ligand binding. *Proteins* **2002**, *49*, 567−576.
(54) Fenton, J. W.; Ofosu, F. A.; Moon, D. G.; Maraganore, J. M. Thrombin structure and function: why thrombin is the primary target for antithrombotics. *Blood Coagul. Fibrinolysis* **1991**, *2*, 69−75.
(55) Adler, M.; Kochanny, M. J.; Ye, B.; Rumennik, G.; Light, D. R.; Biancalana, S.; Whitlow, M. Crystal structures of two potent nonamidine inhibitors bound to factor Xa. *Biochemistry* **2002**, *41*, 15514−15523.
(56) Sanderson, P. E.; Cutrona, K. J.; Dyer, D. L.; Krueger, J. A.; Kuo, L. C.; Lewis, S. D.; Lucas, B. J.; Yan, Y. Small, low nanomolar, noncovalent thrombin inhibitors lacking a group to fill the 'distal binding pocket'. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 161−164.
(57) Tucker, T. J.; Brady, S. F.; Lumma, W. C.; Lewis, S. D.; Gardell, S. J.; Naylor-Olsen, A. M.; Yan, Y.; Sisko, J. T.; Stauffer, K. J.; Lucas, B. J.; Lynch, J. J.; Cook, J. J.; Stranieri, M. T.; Holahan, M. A.; Lyle, E. A.; Baskin, E. P.; Chen, I. W.; Dancheck, K. B.; Krueger, J. A.; Cooper, C. M.; Vacca, J. P. Design and synthesis of a series of potent and orally bioavailable noncovalent thrombin inhibitors that utilize nonbasic groups in the P1 position. *J. Med. Chem.* **1998**, *41*, 3210−3219.

(58) Vitagliano, L.; Merlino, A.; Zagari, A.; Mazzarella, L. Productive and nonproductive binding to ribonuclease A: X-ray structure of two complexes with uridylyl(2′,5′)guanosine. *Protein Sci.* **2000**, *9*, 1217−1225.
(59) Fischer, E. H.; Charbonneau, H.; Tonks, N. K. Protein tyrosine phosphatases: a diverse family of intracellular and transmembrane enzymes. *Science* **1991**, *253*, 401−406.
(60) Andersen, H. S.; Iversen, L. F.; Jeppesen, C. B.; Branner, S.; Norris, K.; Rasmussen, H. B.; Moller, K. B.; Moller, N. P. 2-(oxalylamino)-benzoic acid is a general, competitive inhibitor of protein-tyrosine phosphatases. *J. Biol. Chem.* **2000**, *275*, 7101−7108.
(61) Finkelstein, A. V. and Janin, J. The price of lost freedom: entropy of bimolecular complex formation. *Protein Eng.* **1989**, *3*, 1−3.
(62) Murray, C. W. and Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J Comput.-Aided Mol. Des*. **2002**, *16*, 741−753.
(63) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3−26.

JM0495778