Fragment-based screening using X-ray crystallography and NMR spectroscopy

Harren Jhoti, Anne Cleasby, Marcel Verdonk and Glyn Williams

Approaches which start from a study of the interaction of very simple molecules (fragments) with the protein target are proving to be valuable additions to drug design. Fragment-based screening allows the complementarity between a protein active site and drug-like molecules to be rapidly and effectively explored, using structural methods. Recent improvements in the intensities of laboratory X-ray sources permits the collection of greater amounts of high-quality diffraction data and have been matched by developments in automation, crystallisation and data analysis. Developments in NMR screening, including the use of cryogenically cooled NMR probes and ^{19}F-containing reporter molecules have expanded the scope of this technique, while increasing the availability of binding site and quantitative affinity data for the fragments. Application of these methods has led to a greater knowledge of the chemical variety, structural features and energetics of protein–fragment interactions. While fragment-based screening has already been shown to reduce the timescales of the drug discovery process, a more detailed characterisation of fragment screening hits can reveal unexpected similarities between fragment chemotypes and protein active sites leading to improved understanding of the pharmacophores and the re-use of this information against other protein targets.

Addresses
Astex Therapeutics, 436 Cambridge Science Park, Cambridge CB4 0QA, United Kingdom

Corresponding author: Williams, Glyn (g.williams@astex-therapeutics.com)

Current Opinion in Chemical Biology 2007, **11**:485–493

This review comes from a themed issue on
Analytical Techniques
Edited by Peter M Fischer

Available online 11th September 2007

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2007.07.010

Introduction

Structure-guided drug design has recently enjoyed a resurgence in drug discovery [1]. In large part this has been due to the adoption of fragment-based screening (FBS) approaches in which sensitive biophysical techniques are used to detect the binding of small molecules, typically less than half the size of those present in traditional screening libraries, to protein targets [2,3]. Structural methods, principally X-ray crystallography or NMR spectroscopy, may then be used to rationalise their binding and direct synthetic modifications to these small scaffolds, rapidly and efficiently increasing their affinity for the target and often allowing other properties, such as selectivity, to be explored at an early stage [3,4•].

In March 2007 there were seven compounds undergoing or approved for clinical trials, whose origins lie in fragment-based screening: ABT-263 (an inhibitor of Bcl-2) and ABT-518 (MMP) from Abbott Laboratories, AT7519 (CDK2) and AT9283 (Aurora kinases) from Astex Therapeutics, PLX204 (PPAR) and PLX4032 (BRAF) from Plexxikon and LP261 (tubulin) from Locus. Three more investigational new drug (IND) applications are expected within the next 12 months: SGX523 (c-Met kinase) from SGX, SNS314 (Aurora kinases) from Sunesis and AT13387 (HSP90) from Astex Therapeutics.

These initial forays have demonstrated some of the predicted advantages of fragment-based drug design, ranging from an ability to generate drug-like leads against difficult targets, to a reduction in the time taken from inception of the project to the filing of patent or IND applications. In future, considerable interest will be focused on whether these fragment-derived drugs also have additional benefits, such as improved metabolic properties and selectivity profiles, which could significantly reduce the risk of attrition in the clinic.

A less obvious benefit has been the ability of FBS to generate multiple starting points for drug design against a given target. Although fragment libraries are small (typically less than 5×10^3 compounds) compared to their high-throughput screening (HTS) equivalents (typically more than 1×10^6), fragments are intrinsically more promiscuous since they lack redundant complexity which in most cases will reduce the probability of binding to a protein active site [5,6••]. Consequently, both hit rates and the number of different chemotypes, which are detected as hits and which may therefore be developed into lead-series, are greater for FBS than for HTS. The greater number of lead-series increases the chances of finding a lead compound which is suitable for pre-clinical development and the larger chemical space explored by fragment hits can be used to provide additional insights into the requirements of protein active sites which may then be leveraged against related targets or used to design novel chemotypes that were not represented in the original library. The ability to use this information depends strongly on the degree of structural understanding of these hits and it is in this area that many of the current improvements in experimental methods are directed.

It has been estimated that loss of rotational and translational entropy opposes binding of a fragment by some 15–20 kJ mol^{-1} which is equivalent to three orders of magnitude in the room-temperature dissociation constant [7]. Thus a fragment which has a dissociation constant of only 100 μM must still make strong, complementary interactions with the protein in order to overcome this barrier. Typically, these interactions are retained in the final, nanomolar-affinity, drug candidate and will contribute more than half of the intrinsic binding energy. Accurate measurement of the binding affinities or relative affinities of fragments therefore allows the strongest interactions to be identified and retained during optimisation. Fragment-based screening generates hits with ligand efficiencies, defined as the binding free energy per ligand heavy-atom [8,9] which are nearly optimal, given the nature of the active site and the chemical space explored by the fragment library. This information provides an early estimate of the likely minimum molecular weight of a lead candidate of given potency, although in some cases this may be an oversimplification [10]. It can also allow the 'druggability' of the protein target to be assessed at an early stage [11] and can provide insight into protein recognition of substrates and cofactors [12] or clues to the functions of less-well characterised targets [13,14].

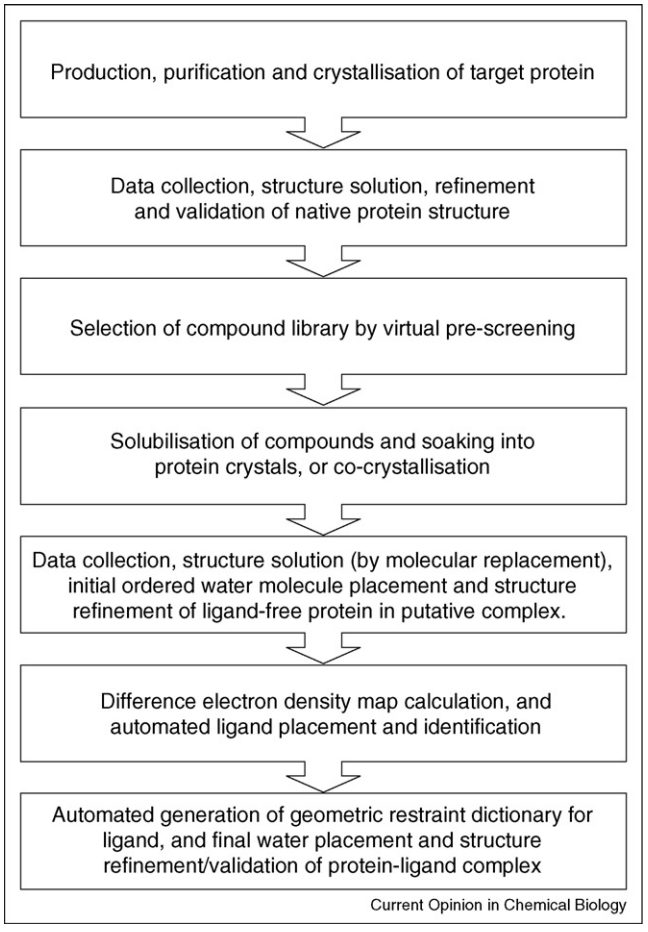
There are many recent excellent reviews of fragment-based drug discovery [15,16]. This article focuses on the most recent developments in the application of X-ray crystallography and NMR spectroscopy to the initial identification of hits — fragment-based screening.

X-ray crystallography

The use of crystallography as a method for fragment screening often involves the preparation and analysis of several hundred crystalline complexes. Innovative developments in several areas have made this a realistic and practical proposition.

The methodology has been well described in the literature [17••,18] and is illustrated in Figure 1. In order to screen a library of fragments by this method, it is necessary to expose the protein to the compounds, and solve the crystal structures of the resulting complexes. The method most frequently described in the literature involves growing crystals of the target protein, and then soaking them in solutions of the fragments, either as single compounds or as cocktails of compounds. It is usually more efficient to expose the crystals to fragment cocktails. The optimum number of compounds in the cocktail is dependent on the target, the likely affinity of the compounds and the expected hit rate, and is usually between 2 and 8, although soaking of up to 100 compounds in one cocktail has been reported [19]. The concentration of each compound in the cocktail is relatively high, in the region of 25–100 mM, as it is expected that affinities for fragments will lie in the high micromolar to low millimolar range [16]. An alternative approach in which a larger library of fragments is first screened using an enzymatic assay, followed by the high-throughput co-crystallisation of hits with the target protein, is well described by Card *et al.* for phosphodiesterase [20••].

Figure 1



Typical flow chart for high-throughput ligand screening using crystallography. Reproduced with the permission of the Royal Society of Chemistry [18].

From protein production to X-ray diffraction

Screening by X-ray crystallography requires the production of large numbers of diffraction quality crystals, either in the absence of compound for soaking experiments, or in the presence of compound for co-crystallisation. It is therefore necessary to be able to reproducibly generate crystals of a similar size and quality on a large scale. This requires optimisation of the processes from protein production to crystallisation.

Progress has been made in protein production, partly due to the efforts of structural genomics initiatives and partly due to the introduction of robotics which has enabled parallelisation of expression trials, increasing the chances