Lead generation and lead optimisation:

the value of linking HT co-structure analysis and HT chemistry

The coupling of High Throughput co-structure analysis with focused library generation is not only proving a powerful general tool in lead optimisation but also increasing the probability of successful discovery of high quality oral development compounds for targets that have been quite difficult for the pharmaceutical industry.

By Dr David U'Prichard

An earlier article[1] identified the lead optimisation process as the major bottleneck of modern high throughput drug discovery, and expertly outlined the virtue of parallel multi-parametric lead optimisation in the drug discovery process. In my experience, it still today typically takes a pharmaceutical company research effort 12-18 months to discover 'sustainable leads' (CTX or IND) at a new target, and about 2-3 years to improve that lead to a development candidate compound that merits the investment required to conduct GLP analytical and safety studies to support the application to conduct clinical experiments in human volunteers (CTX or IND). A sustainable lead for an oral drug development project is a compound with about 100nM affinity for the target protein, 100-fold selectivity for that target compared to other related proteins, acceptable absorption and physicochemical characteristics, and very importantly the molecular weight and structural attributes that medicinal chemists recognise as being a valid starting point for lead optimisation. In fact, investment in a lead optimisation programme typically requires that there are two or three good chemical lead series that have the potential for improvement to development candidate status. As pointed out previously[1], the above definition of sustainable lead as a starting point for lead optimisation varies somewhat from organisation to organisation. Nevertheless, something close to this definition is required to increase the investment in the programme to at least three chemistry teams (10 chemists) and perhaps initially six bioscientists.

Previously in *DDW*, I have described[2] the modern amalgamation of structure-based drug design and high throughput combinatorial chemistry, where structural information obtained from x-ray crystallographic or NMR experiments has been integrated into the design of focused combinatorial libraries ranging from several score to one thousand or more molecules. The aim is to balance rational sculpting of the compound that fits perfectly into the receptive site on the target protein with empirical screening of many compounds that represent a multitude of experiments to optimise the several key parameters mentioned before[1]. In this article, I want to specifically describe the value of obtaining many co-crystals of the target plus different lead compounds to obtain increasingly refined knowledge about the tolerance of the receptive site of the target for structural variation in the lead compounds as optimisation of several properties (eg affinity, solubility and absorption characteristics) is brought into play. I refer to this capability as 'high throughput co-structure analysis', which when allied to high throughput combinatorial chemistry is a powerful driver for not only

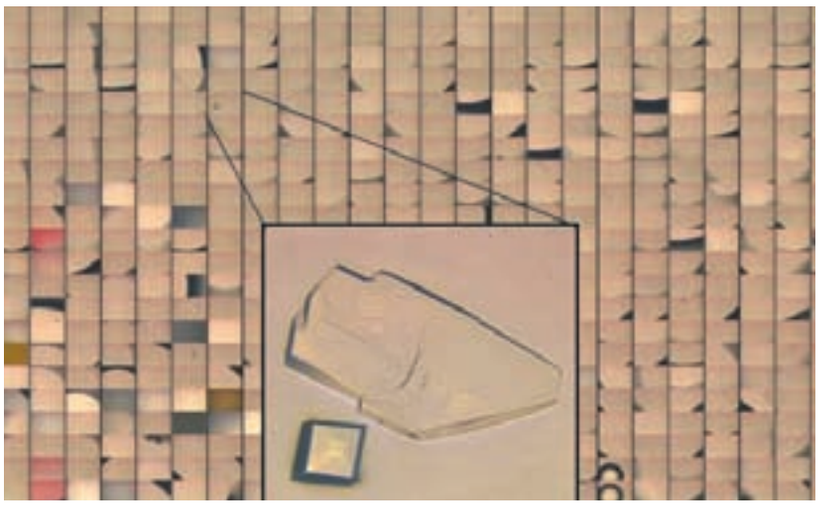
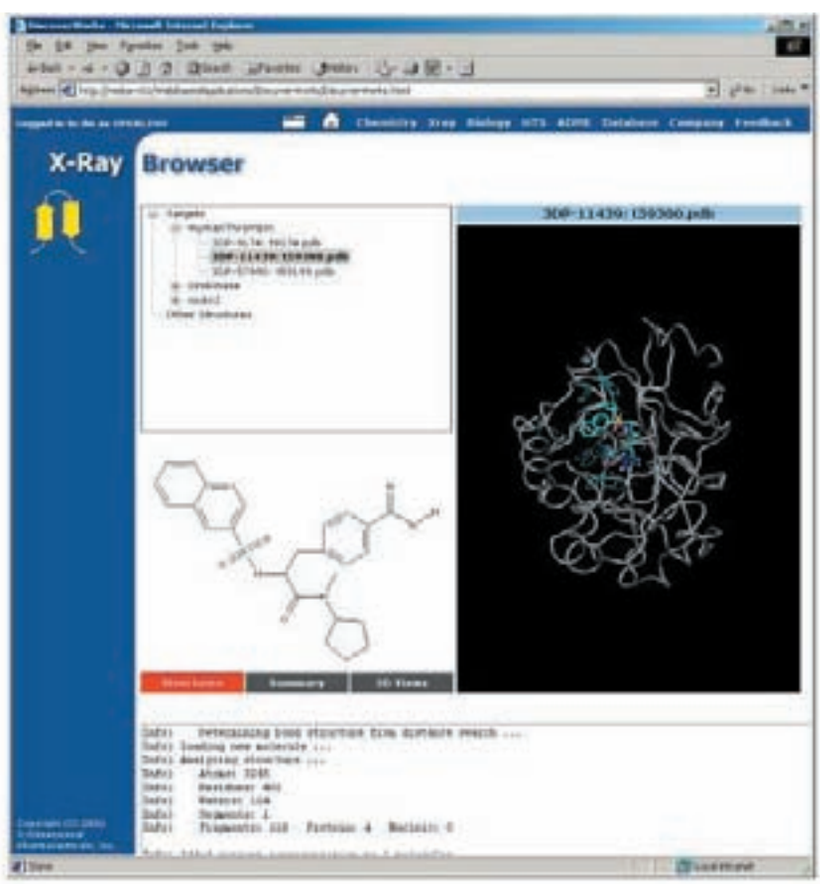


Figure 1
Close-up view of an individual well from a 1536-well crystallisation experiment

Figure 2
Screen from web-based integrated chemistry and x-ray target structure database at 3DP, illustrating a set of co-crystallisation data alongside chemistry data for a compound in 3DP's oral thrombin inhibitor programme



faster, more efficient lead optimisation, but also for increasing the probability of successful discovery of high quality oral development compounds for targets that have been quite difficult for the pharmaceutical industry. At 3-Dimensional Pharmaceuticals, we have successfully developed orally active, potent leads for historically difficult targets such as serine proteases and integrin receptors, and are reaching the same point of success with a difficult protein/protein interaction target. I describe here our efforts to discover oral direct-acting thrombin inhibitors, and antagonists of the interaction of the p53 and hdm2 proteins.

Technological advances supporting high throughput co-structure analysis

1. High throughput crystallisation

Recent miniaturisation of crystallisation trials for x-ray crystallography has been a major technical development. A single plate, containing many wells that examine different conditions for crystallising the target protein, can replace many larger plates, greatly reducing the target protein requirement, allowing for more rapid equilibration and optimisation, and better temperature control, and reducing the readout time. Automated plate imaging simplifies the task of scanning for crystals and precipitates by a single researcher. Artificial intelligence procedures are becoming helpful in analysing the 'grey area' that can occur between precipitate and micro-crystal. A number of companies including 3DP have installed a 1536-well platform for HT crystallisation with desktop image scanning and software that allows blow-up of individual wells and easy identification of promising conditions and crystals (**Figure 1**). Large crystals in wells can be directly mounted and diffraction images taken. Quite powerful laboratory-scale x-ray sources are now available with high-flux optics, which allow for rapid turnaround of target protein structural information to medicinal and computational chemists.

2. Integration of chemical and structural databases

Many attempts are now being made to integrate target protein structural information (x-ray or NMR), especially from co-crystals of the target protein and a small molecule, into chemical compound databases that contain at least 2-D and possibly 3-D information about the structure of the compound, together with theoretical and experimental physicochemical, pharmaceutical and biological activity data. At 3DP, we have developed a web-based database for this task (**Figure 2**), which affords ease of annotation, customisable views, saved orientations of the co-structure, and predictive capacity.

Oral thrombin inhibitors

Many companies including 3DP have designed and made quite large combinatorial libraries for specific families of target proteins, eg proteases, kinases, GPCRs, etc. Structure-based design of focused combinatorial libraries (100-1,000 compounds) for lead optimisation at 3DP involves a 'recombination' of initial combinatorial screening library features using 'structure-based design', especially information from many co-crystallisation experiments.

The discovery and development of orally active, direct-acting small molecule inhibitors of the blood coagulation enzyme thrombin has been challenging for the pharmaceutical industry over the last 20 years. The major difficulty has been to retain the requisite selectively high affinity for thrombin while optimising pharmaceutical properties that would ensure good oral bioavailability and pharmacokinetics in man. At 3DP, we linked our ability to produce focused combinatorial libraries with a structure-based design approach based on more than 130 co-crystal structures of thrombin and 3DP compounds to optimise compounds for potency, selectivity, oral bioavailability, *in vivo* activity and good pharmacokinetics. These optimised compounds have begun to be tested in man, and so far there is an excellent translation of pre-clinical properties.

The structures of thrombin and a number of other serine proteases are known, and at a secondary level appear to be identical with only minor variations (**Figure 3A**). However close examination of the solvent accessible surfaces of these enzymes shows they are in fact quite different (**Figure 3B**). By taking advantage of these structural differences, we designed a highly selective library of compounds at the outset. Detailed knowledge of the structure of thrombin served as the starting point. The catalytic residues were deliberately not targeted to avoid specificity and reversibility issues. The two major interactions areas in the thrombin structure are the basic or guanidine binding site located at the bottom of the S1 pocket and the aryl binding region located adjacent to the S2 or 'scaffold' region (**Figure 4**). A central scaffold was selected. Different P1 pieces were explored by combinatorial chemistry around the central scaffold and counter-screening was routinely performed against several other serine proteases to look for both potency and selectivity advantages. Routine co-crystallisation of inhibitors established that although the P1 and aryl groups from many inhibitors occupied similar areas on the enzyme, these groups projected from many different angles. This, and the opportunity for broad diversity around the central scaffold, allowed for many different series of compounds to be made,

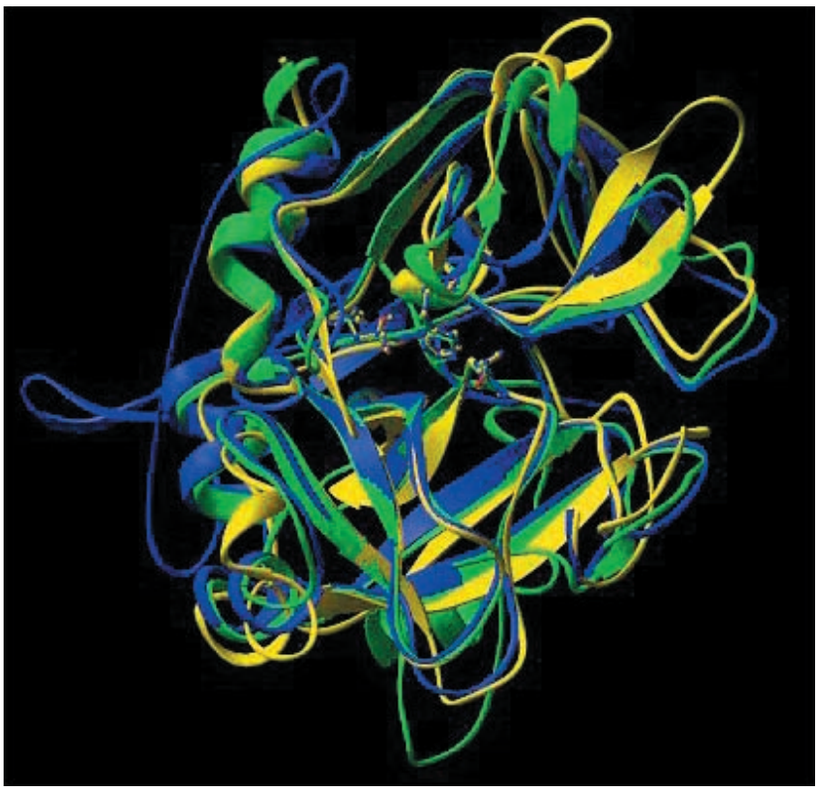
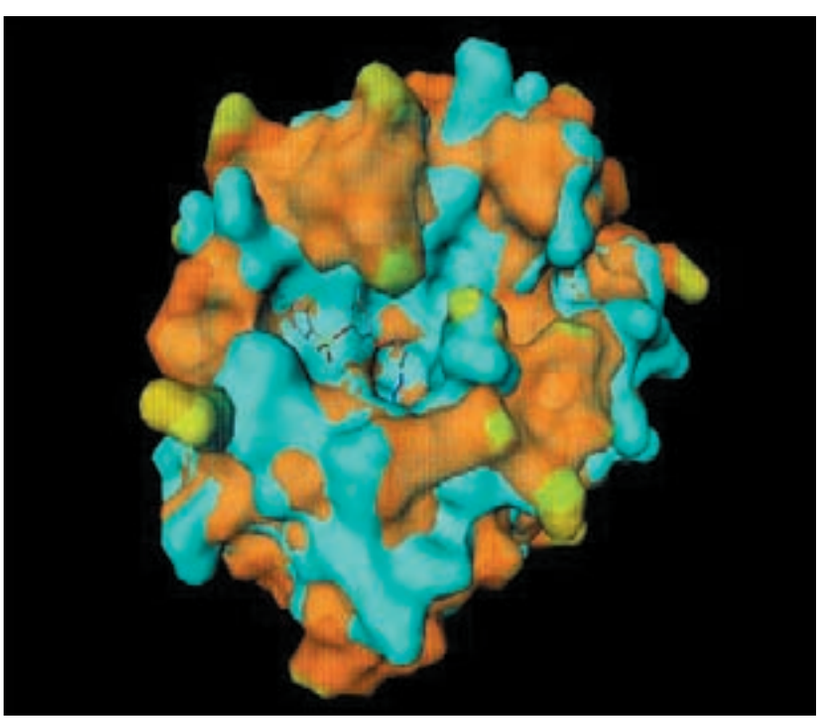


Figure 3A

Structural 'ribbon' diagrams for three serine proteases, Thrombin, Factor Xa and Urokinase. At the secondary structural level serine proteases appear to be identical with only minor variations. GREEN – thrombin; YELLOW – urokinase; BLUE – factor Xa

Figure 3B

Structural 'solvent accessible surface' diagrams for Thrombin and Factor Xa, illustrating significant structural differences. RED – thrombin; BLUE – factor Xa



screened and co-crystallised. Diverse chemical properties were sampled in early combinatorial libraries to allow for later optimisation of pharmacokinetic properties, and even at this early stage of optimisation, potent and highly selective compounds were obtained (**Figure 5**).

Combinatorial exploration of the aryl binding pocket and co-crystallisation of these compounds provided much insight into how both enzyme and inhibitor conformations contributed to the binding of compounds. The enzyme itself undergoes conformational shifts in order to interact maximally with any bound compound. Knowledge of the binding modes of multiple structures from a combinatorial experiment and co-crystallisations allowed the chemists to envisage the possibilities for expanding the range of compounds with varying physicochemical properties to maximise both binding affinity and bioavailability. The plasticity of the aryl binding pocket allowed for a large number of different aryl groups to be accommodated with high affinity. Large numbers of structural examples of libraries allowed for increasingly accurate SAR models and predictions to be made (**Figure 6**). The many co-crystallisations performed and co-structures analysed illustrated the diversity available in inhibitor design (**Figure 7**). An example of the power of multiple co-crystal structures is shown in **Figure 8**.

Early on, absorption, efficacy and pharmacokinetic predictions were experimentally tested, and 3DP thrombin inhibitors was continually being improved to meet a profile of: a novel, non-peptide with a non-covalent mechanism of action; potent and selective thrombin inhibition; efficacy *in vitro* with respect to anti-coagulation in human blood; *ex vivo* efficacy in rabbits (i.v. and p.o.); superior pharmacokinetics in rabbits and dogs; efficacy in dogs. The coupling of co-structure information with *in vivo* dog pharmacodynamic data allowed for systematic improvements to be efficiently obtained late in lead optimisation (**Figure 9**). The leading development candidate 3D-4815 was taken into man, and an excellent PK/PD correlation has been observed (**Figure 10**).

Small molecule inhibitors of the p53/hdm2 protein-protein interaction

In cancer cells that express normal p53, the well-known tumour suppressor gene product, another protein hdm2 (mdm2) is a negative regulator of p53 by inhibiting its expression and by binding to p53, thus exposing the p53 protein to intracellular degradation via an ubiquitination pathway. Thus in normal p53 tumours, a small molecule inhibitor of the binding of hdm2 to p53 should cause a prolonged elevation of intracellular p53 levels, driving

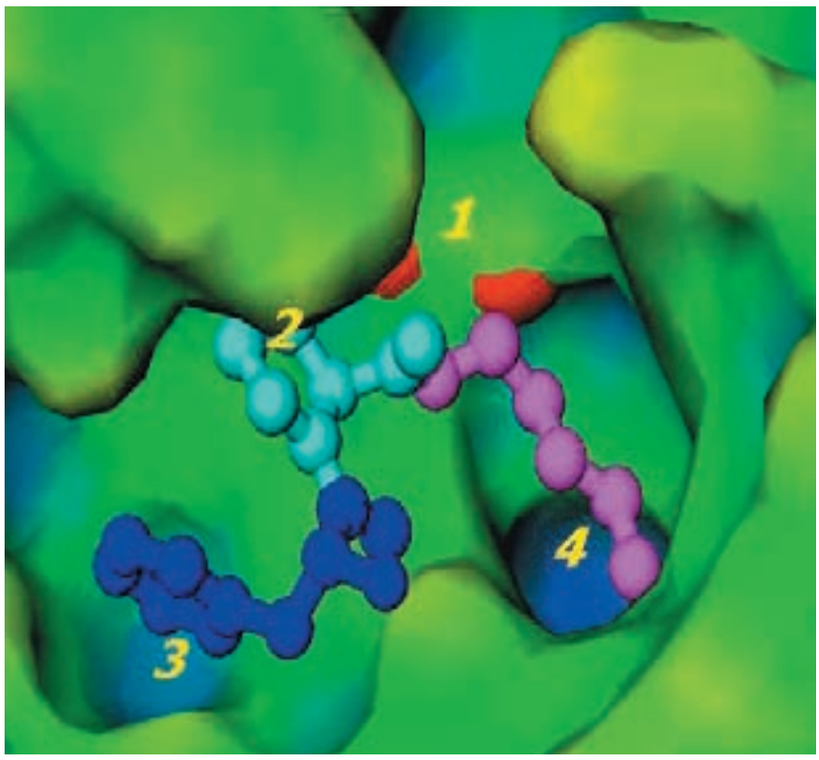
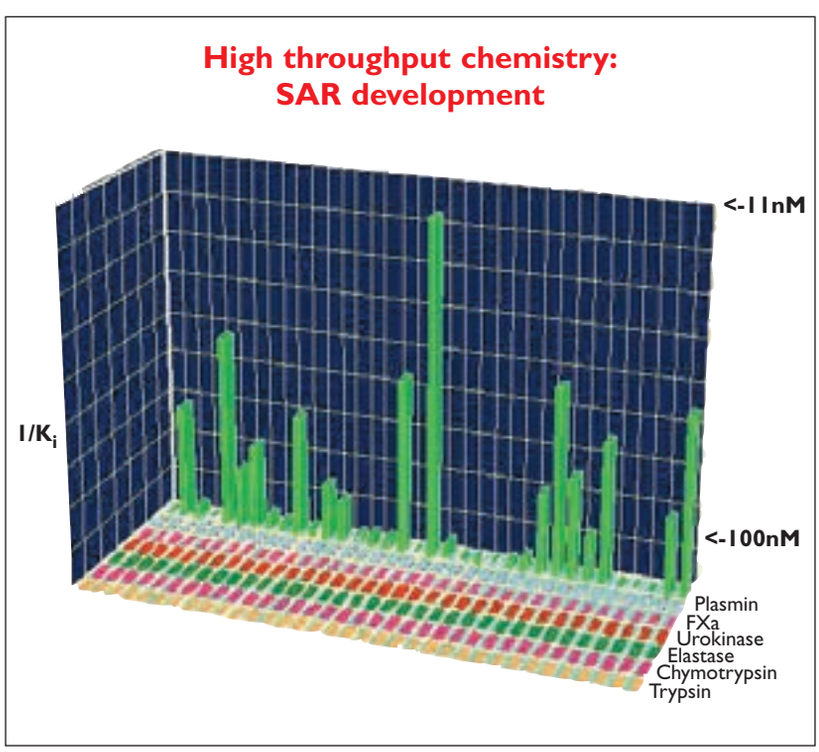


Figure 4
Active site of a serine protease, and the starting point for combinatorial library construction in the 3DP thrombin inhibitor programme. (1) Catalytic residues. (2) 'Scaffold Region' S2 binding site. (3) Aryl binding site. (4) S1 basic group binding site

Figure 5
Screening and counter-screening data from an early focused library of thrombin inhibitor compounds



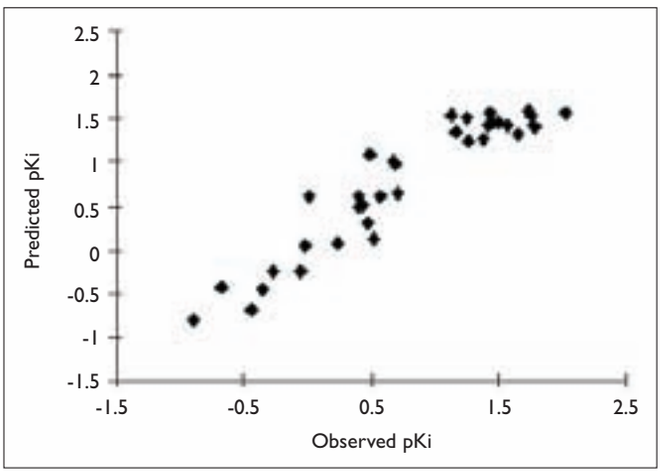


Figure 6
SAR analysis of thrombin inhibitors: correlation of observed and predicted binding constants

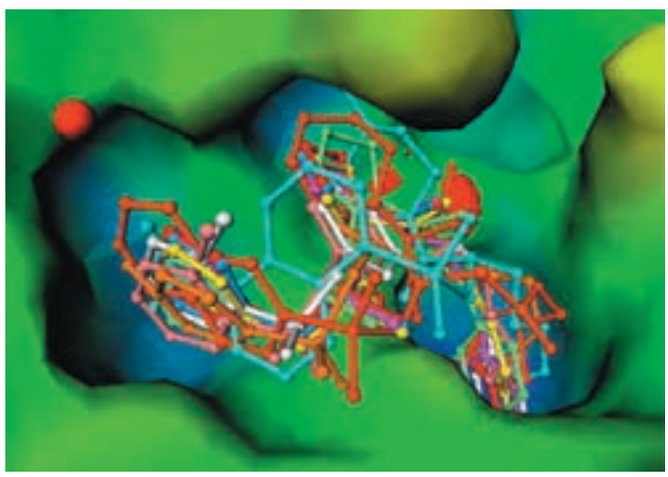


Figure 7
A subset from the more than 130 different thrombin/inhibitor complexes solved shows the diversity available in inhibitor design. The superposition of many different scaffolds facilitates the use of a mix and match strategy for chemistry design work

the tumour cell towards apoptosis rather than division and proliferation, and be very useful as an adjunct to cancer chemotherapy or radiotherapy. This has been a well-recognised and popular target in the pharmaceutical industry for some time, but as with other protein-protein interaction targets, little success has been achieved in finding good small molecule sustainable leads that can be fully optimised. Some general reasons include the lack of an appropriate small molecule binding pocket within the contact surface of the two proteins, and the difficulty of generating precise SAR from weak leads that are observed in a typical ELISA-type HTS.

Combining high throughput co-structure analysis with a non-functional binding affinity screen can powerfully assist lead generation and optimisation for

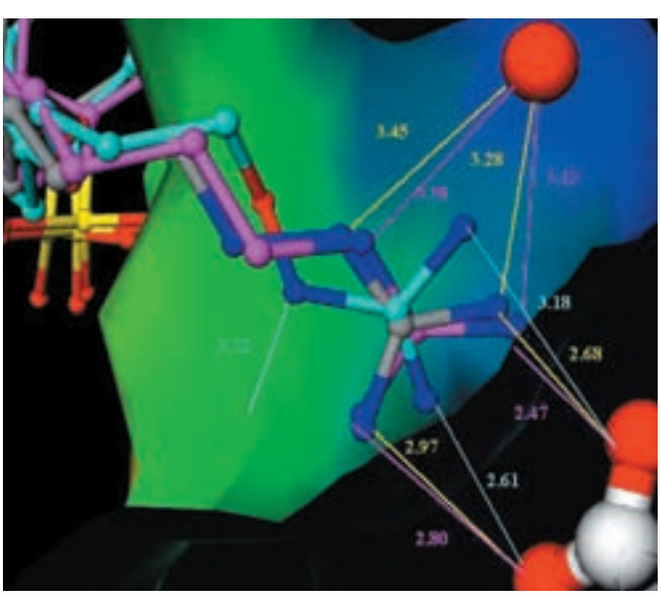


Figure 8
Sometimes minor variations in the hetero atom in a compound were found to have dramatic effects on bioavailability and binding mode. Many of these different binding modes can only be interrogated by x-ray crystallography. The binding affinity may only be slightly altered (all three compounds shown have Ki values between 11nM and 13nM), even though the binding modes are very different

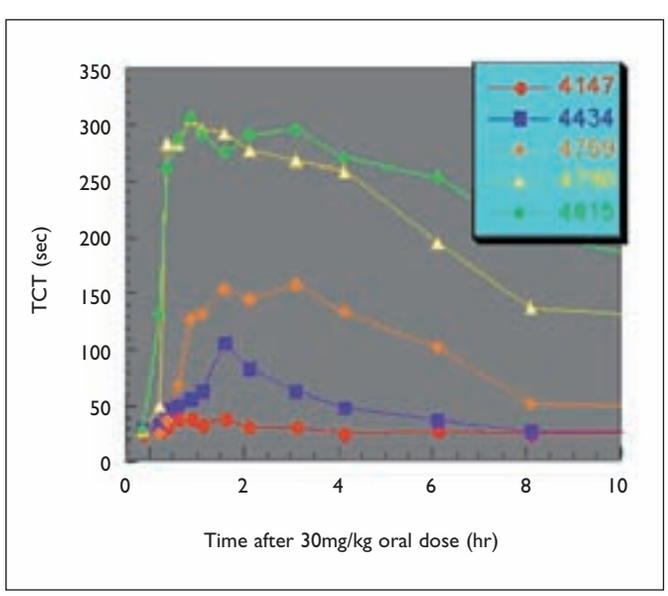


Figure 9
In vivo pharmacodynamic evaluation of lead thrombin inhibitors. X-axis: time (hours) after 30 mg/kg oral dose. Y-axis: thrombin clotting time (sec)

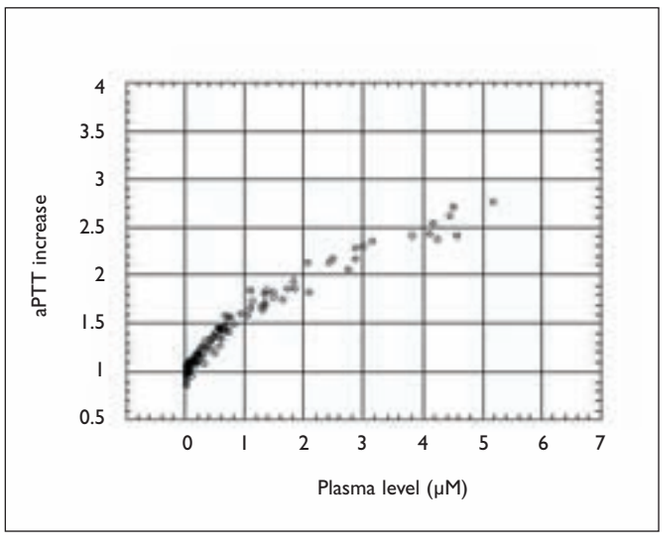


Figure 10
Correlation in human volunteers of pharmacodynamic (Y-axis: increase in alpha prothrombin time (min)) and pharmacokinetic (X-axis: plasma level of drug (μM)) parameters for the thrombin inhibitor 3DP-4815. 25 – 600mg of 3DP-4815 were dosed orally

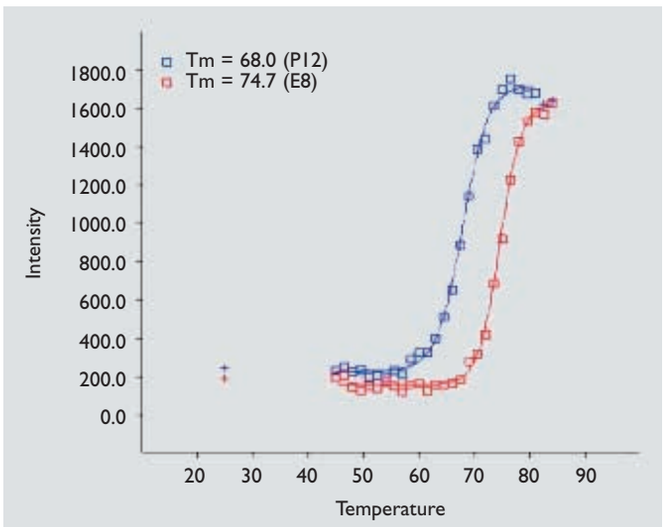


Figure 11
A typical 'hit' response in a thermodynamic screen for compounds binding to hdm2. The melting curve of hdm2 (blue) is shifted in the presence of a small molecule hit (red)

References

1 Baxter, AD and Lockey, PM (2001/2). 'Hit' to 'lead' and 'lead' to 'candidate' optimisation using multi-parametric principles. Drug Discovery World 3 (1), 9-15.
2 U'Prichard, DC (2000). Exploiting the fruits of the human genome: a strategic perspective. Drug Discovery World 1 (2), 12-22.

a protein-protein interaction target. Using a thermodynamic binding screen[2] coupled to x-ray structure determination of co-crystals of lead compounds with a construct of hdm2 protein, at 3DP we have been able to make quite rapid progress with this difficult target. Screening a library of 220,000 compounds yielded a quite respectable 0.14% hit rate (**Figure 11**). Two rounds of combinatorial chemistry and screening iteration produced compounds that were potent enough to be co-crystallised successfully with hdm2, to verify that the screening leads actually bound to the same site on the hdm2 protein as p53 (**Figure 12**). These screening leads exert anti-proliferative and pro-apoptotic effects in a number of tumour cell models by demonstrably preventing the binding of p53 and hdm2 in these cells. Co-crystallisation of a significant number of compounds with mdm2 is rapidly providing a wealth of target structural information to expedite the lead optimisation process.

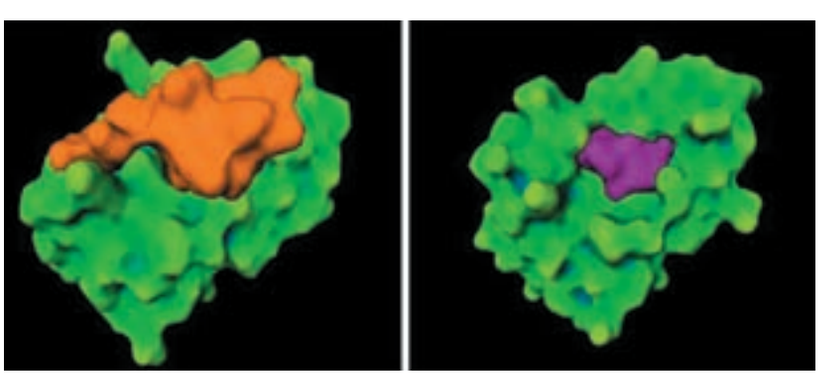


Figure 12
X-ray structures of hdm2 protein (green) complexed with the binding fragment of p53 (left) or a small molecule lead (right)

Summary

High throughput co-structure analysis coupled to focused library generation is already proving a powerful general tool in lead optimisation, and should increase the quality and probability of success of small molecule compounds entering development. Larger scale high throughput crystallisation of putative target proteins and crystal structure analysis will increasingly push this technology upstream into lead generation, especially when combined with a robust HT affinity binding screen for novel targets.

Acknowledgements

I wish to thank Drs John Spurlino and Roger Bone for their helpful contributions to this article. DDW

Formerly President, Research and Development, SmithKline Beecham Pharmaceuticals and International Research Director of Zeneca, ***Dr David U'Prichard*** *has overseen the entry of 10 compounds into clinical development and progressed the clinical development to NDA filing of Avandia, the blockbuster diabetes drug. Dr U'Prichard was instrumental in the launch of Nova Pharmaceuticals in 1983 and is the current Chief Executive Officer of 3D Pharmaceuticals, Inc.*