



**Trip Report:
Advances in Structure-Based Drug Design
Philadelphia Convention Center
Philadelphia, PA
October 16 – 19, 2005**

Helene Y. Decornez, Ph.D., Sean R. Dinn and Kristen N. Fleming, Ph.D.

Medicinal Chemistry Department
Albany Molecular Research, Inc.
21 Corporate Circle
Albany, NY 12212

***Abstract:** This ACS ProSpectives Conference was created to bring together medicinal chemists, computational chemists and structural biologists with the common goal of advancing the area of structure-based drug design. Approximately 180 scientists were in attendance. A summary of selected talks from the conference are presented.*

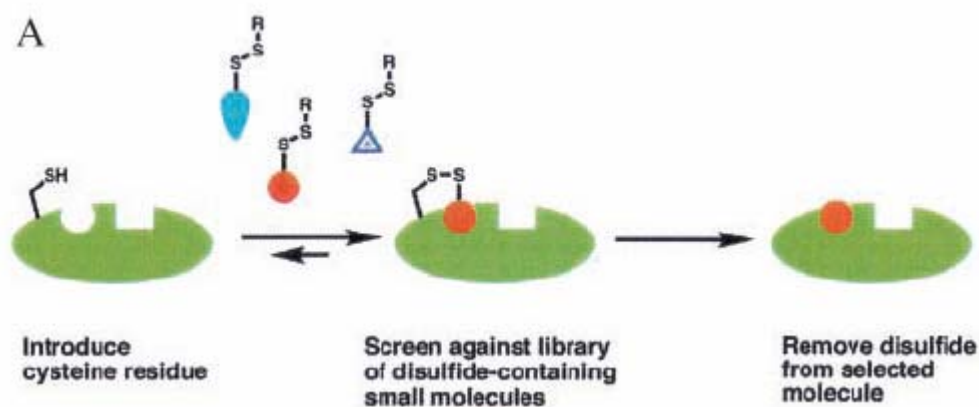
“Site-directed Small Molecule Discovery for Allosteric Sites,”

James A. Wells (University of California), San Francisco, CA.

Professor Wells described a new fragment-based drug discovery method called Tethering® which has aided in the identification of allosteric sites on proteins and helped to speed up the process of finding compounds that bind to these sites. This approach enables detection of weakly-bound small drug-like fragments at normal screening concentrations which are subsequently used as starting points toward the computer-aided design of potential leads. In addition, the tethering technology can be used to focus on specific sites of interest on the proteins that may otherwise be difficult to target selectively.

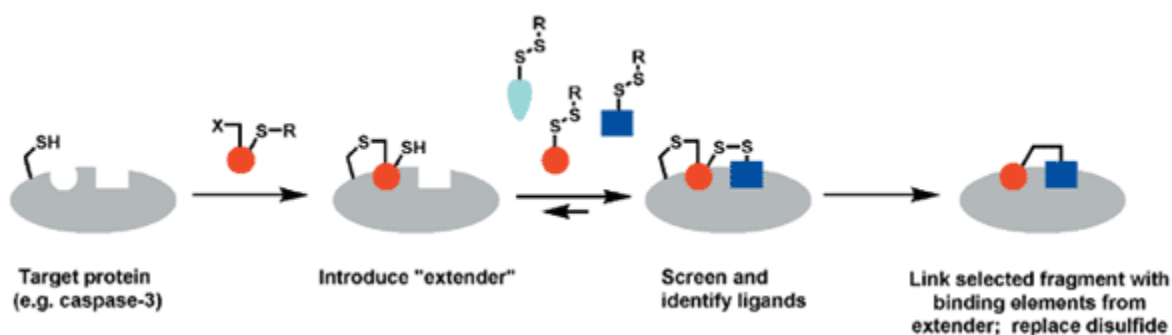
A small library of disulfide-containing compounds is allowed to react reversibly under thiol exchange conditions with specific native or synthetic thiols in a protein (Figure 1). The thiol-tethered ligands are then identified using mass spectroscopy and, many times, crystallization is possible at this stage. These modified proteins have proven to be quite stable while, in the absence of the tethers, many of the ligands would have bound to the sites only weakly. Ordinarily, moderate affinity compounds are difficult to identify in high-throughput screens, especially at typical screening concentrations. This approach allows for a higher success rate, even when a smaller set of compounds is screened and specific sites are targeted.

Figure 1



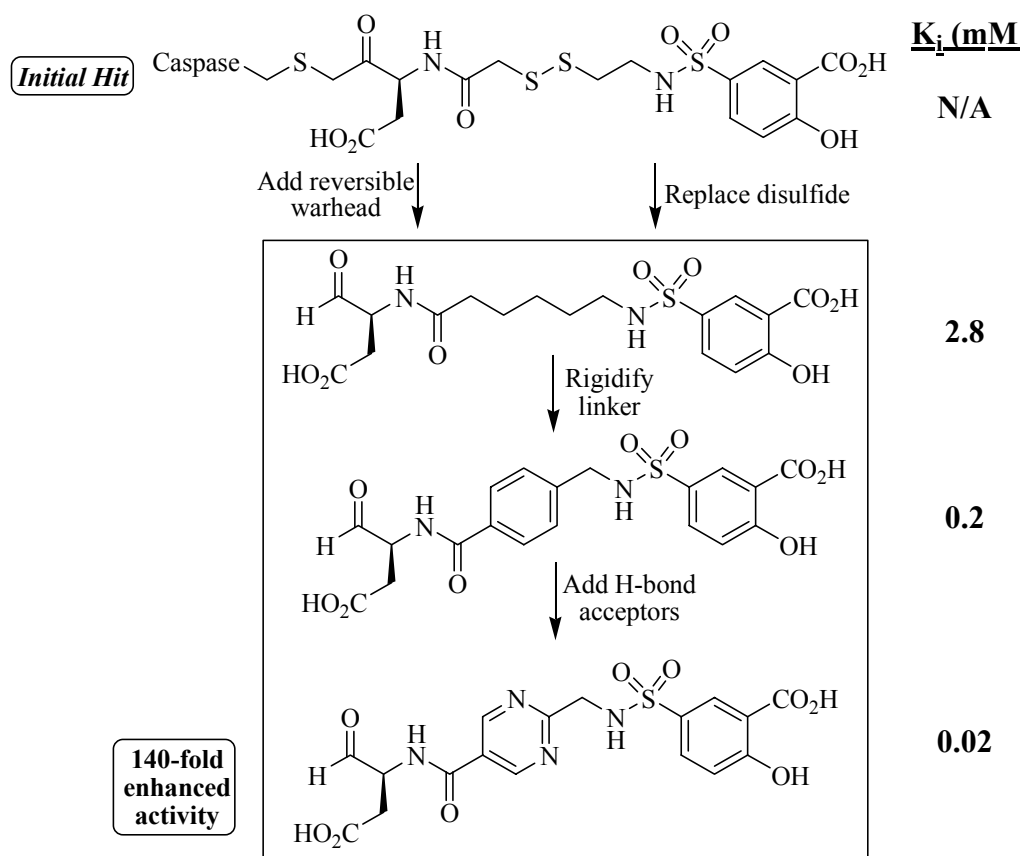
Additionally, this technology has been taken one step further to a method known as Extended Tethering® (Figure 2). A cysteine residue near the active site is modified covalently with an “extender”, a molecule possessing some affinity for the protein that bonds to the cysteine and also contains a protected thiol. After deprotection, the complex is screened against the library of small disulfide-containing fragments. Once again, the disulfide bonds that form help to stabilize bound fragments that otherwise may have bound to the proteins only weakly. Fragments identified by mass spectroscopy are then combined with elements from the extender and screened for inhibition of the protein. In some cases, this approach allows for quicker lead optimization than the original method of Tethering®.

Figure 2



This technology was subsequently used by Professor Wells and his co-workers to identify novel inhibitors of the proinflammatory caspase, Casp-3. One of their initial hits was subsequently converted to a non-covalent binding form and further optimization resulted in a compound with 140-fold greater activity as compared to the initial lead (Figure 3).

Figure 3



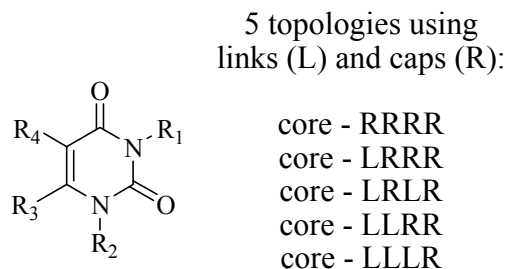
“Molecular Modeling Applied to Anti-HIV Drug Design,”

William L. Jorgensen (Yale University), USA.

Professor Jorgensen’s research group at Yale has utilized a new computer program called BOMB, in conjunction with other software, to design libraries of compounds based on known protein

structures and targeted cores. The BOMB program is used initially to “grow” analogs inside the protein’s binding site, then performs a conformational search and scores each of the analogs in terms of potential binding affinity. More specifically, BOMB will begin with a core structure, such as the one shown in Figure 1, and then generate either individual analogs or a combinatorial library. The example in Figure 1 allows for the possibility of varying four different attachments. As an alternative to growing the analogs, the program will link the structures to one another to provide either globular or elongated compounds. The example given below allows for five different topologies using links (L) and caps (R).

Figure 1



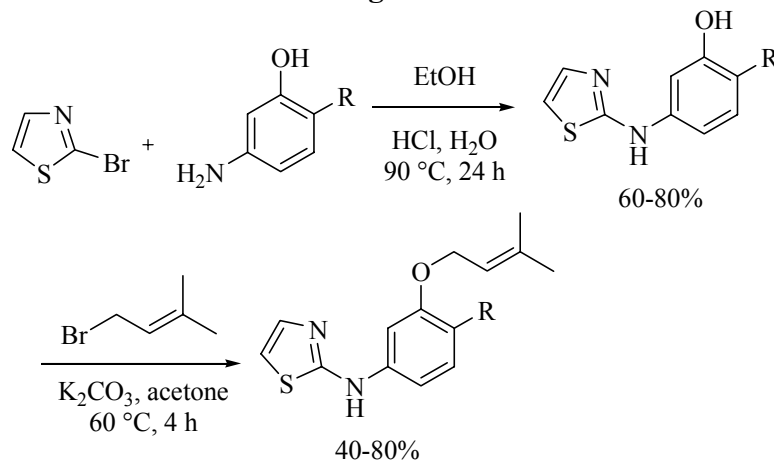
Since the computer software has >600 substituents in its database, a compound with four variable substituents and five possible topologies would result in a virtual library containing 10 trillion molecules! Fortunately, scoring functions are used to decrease this amount to a much more manageable number. Many times, very simple cores such as ammonia, methane, and benzene are used as the starting points, although the software contains over 100 cores.

Additionally, the program will build all possible conformers and then optimize them to select the best one. The first pass optimization does not take any protein flexibility into consideration and generally takes about one minute per compound. The second pass does in fact take protein flexibility into account and, for this reason, requires approximately five minutes per compound. Scoring functions are then used to predict the compounds’ potential activity.

The next stage of the process involves a program called QikProp which is used to determine which compounds possess favorable drug-like properties. Professor Jorgensen stressed that the most crucial properties to investigate at this point are solubility and cell permeability, but the program provides quite an extensive list of predictions. From there, MC/FEP simulations are used to refine the predictions for the best scoring leads.

This technology was subsequently used in the design of novel HIV reverse transcriptase inhibitors. Figure 2 shows the chosen core structure, along with the top-scoring substituent choices for two of the variable sites, Het and U, as predicted by BOMB.

Figure 4



Further analysis of this library was achieved using the programs QikSim and QikProp which filtered the library for similarities to known active compounds in this area. Two criteria used in the comparison were size and polarity. Once again, solubility was stressed as a key issue to consider as well. Work continues in this area in Professor Jorgensen's lab.

“Structure-Based Design of a Series of Potent and Selective Cell Permeable Inhibitors of Human β -Secretase,”

Joseph P. Vacca (Merck & Co, Inc.), USA.

Alzheimer's disease is on the rise globally and currently costs \$100,000,000,000 annually in the US alone. As it is the third most expensive disease to treat in the US and is always fatal, Alzheimer's most certainly represents an important target for drug discovery. Subsequently, Dr. Vacca and co-workers at Merck have been investigating new therapeutic agents for the treatment and prevention of this disease.

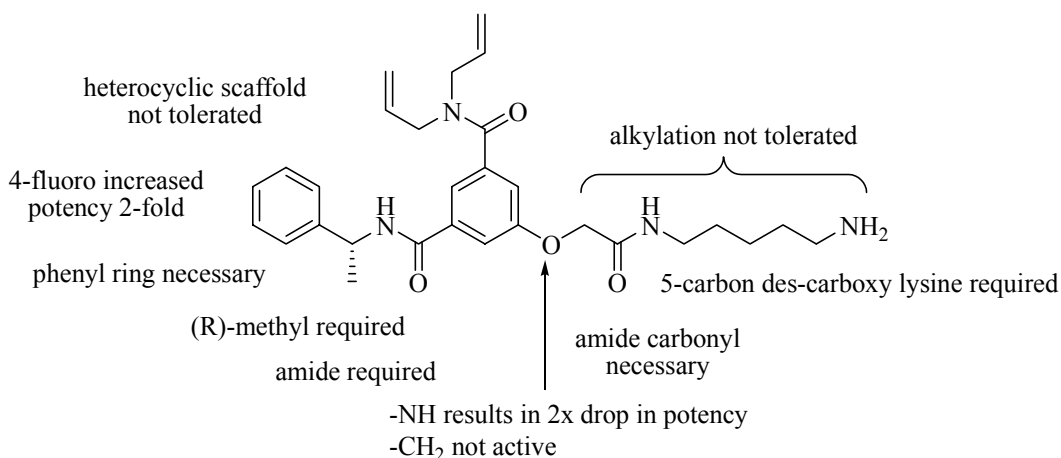
The Amyloid Cascade suggests that it may be possible to prevent Alzheimer's through the use of β -secretase or BACE-1 inhibitors. BACE-1 (β -site APP cleaving enzyme) is an aspartyl protease responsible for the accumulation of amyloid- β peptide in the brain, which leads to the formation of “plaques” and the onset of Alzheimer's. Therefore, researchers at Merck set out to perform a high-throughput screen in order to identify potent BACE-1 inhibitors.

Dr. Vacca noted a number of challenges associated with their BACE inhibitor research program. The BACE-1 active site differs substantially from sites for other aspartyl proteases in that it is more open and less hydrophobic. Additionally, they were challenged by the fact that their inhibitors must be selective for BACE-1 over BACE-2 as well as other aspartyl proteases. Another obvious obstacle is the fact that their compounds must be able to cross the blood-brain barrier, with effective concentrations reaching the target.

Researchers at Merck initially performed a screen using one of their existing sample collections plus a few compounds from their renin inhibitor program. Most of the compounds were large peptidomimetics and unfortunately were not orally bioavailable. They also generally did not penetrate the blood-brain barrier. Therefore, collaboration with Neogenesis was initiated and an ultra-high capacity screen was performed in order to identify non-peptide BACE-1 inhibitors. This

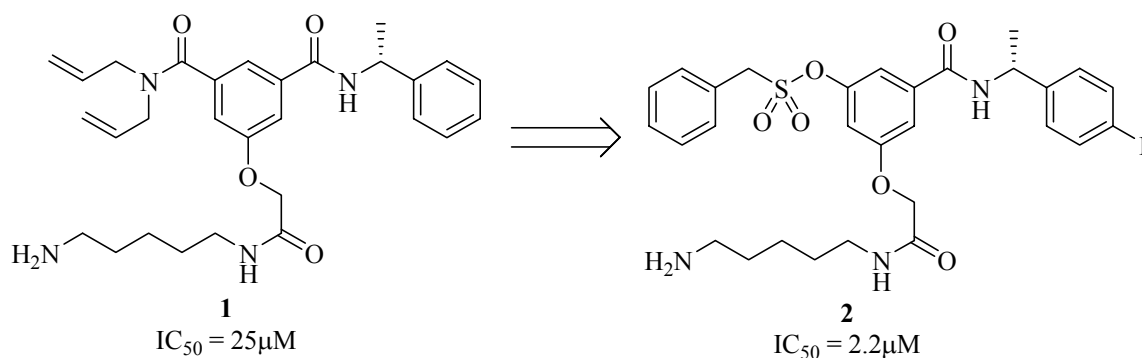
screen eventually provided the lead shown in Figure 1. Their subsequent lead optimization efforts are detailed as well.

Figure 1



Efforts toward optimizing these new BACE-1 inhibitors eventually resulted in formation of compound **2** (Figure 2), which exhibited an IC₅₀ of 2.2μM. Luckily, it was possible to obtain an x-ray crystal structure of this compound bound to BACE, providing valuable information regarding the inhibitor's mode of binding.

Figure 2



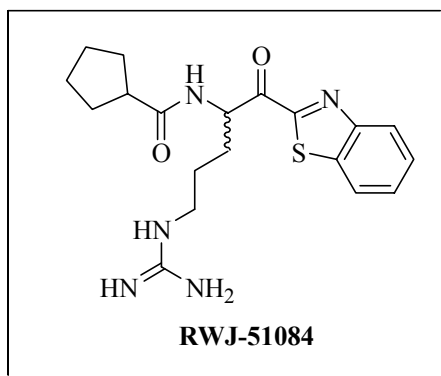
Unfortunately, these still did not prove to be brain penetrable compounds and so work continues in this area. Current efforts are focused on incorporating additional hydrophilic groups and attempting to “freeze” the compound in its bioactive conformation. It is anticipated macrocyclization will be able to accomplish the latter. They are also hoping to retain the compound size while increasing potency.

“Structure-Based Drug Design Applied to Serine Protease Inhibitors,”

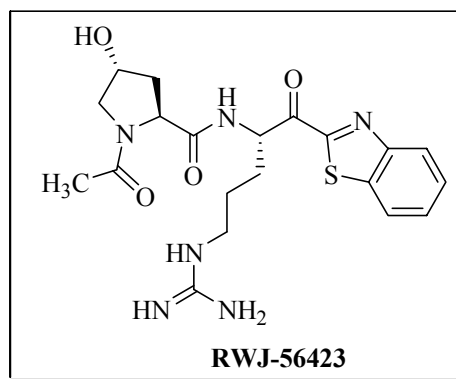
Bruce Maryanoff (Johnson & Johnson Pharmaceutical R&D).

Serine proteases are a very important class of enzymes that have been associated with a wide variety of disease states. Dr. Maryanoff has been extensively investigating several of these serine proteases and has developed inhibitors of them.

Tryptase is a proinflammatory mediator secreted by mast cells. It constitutes 20-25% of the total protein of human mast cells. Tryptase has been directly linked to the onset of asthma. Inhibitors of tryptase have the potential to treat a variety of inflammatory related conditions including asthma, atherosclerosis, inflammatory bowel disease and psoriasis. Tryptase is also an attractive target because it has no known endogenous inhibitors and it is only one of a few proteases that can activate protease-activated receptor 2 (PAR-2), which may be partly associated with the detrimental effects of tryptase. There have been no reported molecules that interrupt PAR-2 activity.

Figure 1

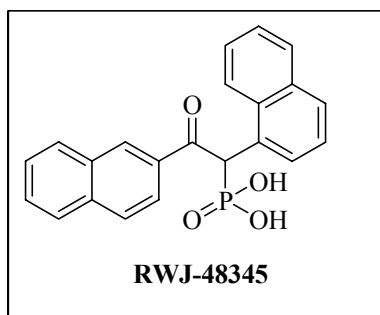
While studying truncated inhibitors of thrombin, another serine protease, RWJ-51084 was discovered to be a potent inhibitor of tryptase with a $K_i = 88$ nM (Figure 1). It also showed selectivity over other serine proteases including thrombin, factor Xa and chymotrypsin. RWJ-51084 is also a potent inhibitor of trypsin ($K_i = 30$ nM). Tryptase and trypsin have a high degree of homology in their active sites. A crystal structure of the (S) enantiomer of RWJ-51084 with trypsin was obtained. The inhibitor was occupying the active site with the guanidino group in the S1 specificity pocket interacting with Asp-189. The activated ketone formed a hemi-ketal with the O γ of Ser-195 and a hydrogen bond was observed between the benzothiazole nitrogen and N ϵ of His-57. Computer modeling was then performed using known information on the active site of human β -tryptase to obtain a more potent inhibitor. From this work, RWJ-56423 was identified as a promising inhibitor (Figure 2).

Figure 2

From this exercise it was observed that (S) stereochemistry at the arginine was necessary. The arginine moiety, acetyl substitution on the proline and benzothiazole group were needed for good potency. Selectivity was realized over many serine proteases including thrombin and factor Xa. Unfortunately this compound was not selective versus trypsin ($K_i = 8$ nM vs. 10 nM for trypsin). A co-crystal could not be obtained between β -trypsin and RWJ-56423, but a co-crystal with bovine pancreatic β -trypsin was obtained. The crystal structure indicated several features that led to affinity increases versus RWJ-51084. The acetyl group on the proline extends deeper into the S2 domain. The hydroxyl group picks up interactions with Ser-96 through hydrogen bonding network of water molecules. Unfortunately this modification doesn't lead to improved selectivity over trypsin. The racemic version of RWJ-56423 (RWJ-58643) has shown efficacy in a sheep model for asthma. Consequently, RWJ-58643 has entered clinical trials for the treatment of asthma. This compound is being administered as an aerosol to circumvent bioavailability issues.

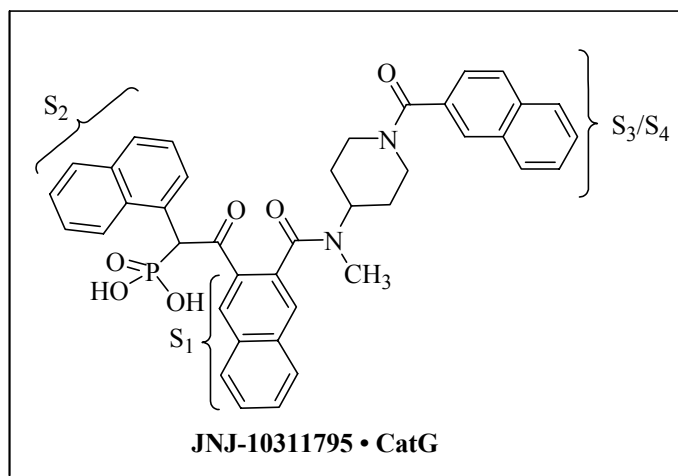
Another project involving dual inhibitors of cathepsin G (CatG) and chymase were described. CatG and chymase are both involved in initiating and propagating inflammatory responses associated with a variety of diseases including asthma and chronic obstructive pulmonary disease (COPD). CatG is present in neutrophils and chymase resides in mast cells. RWJ-48345 was identified via a HTS assay as an inhibitor of Cat G (Figure 3, $IC_{50} = 4.1$ μ m).

Figure 3



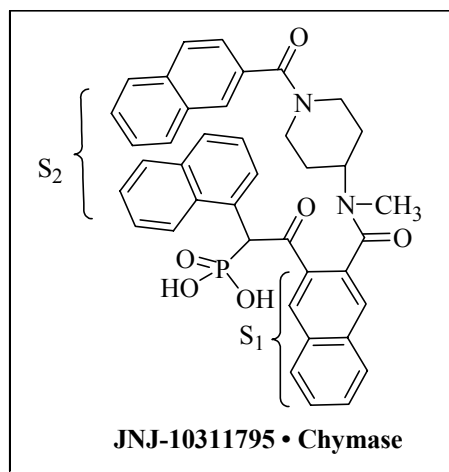
An x-ray crystal structure was obtained of CatG with RWJ-48345 with a resolution of 3.0 Å. The crystal structure revealed the enzyme is bound to the inhibitor with the (R) configuration and the hydrophobic S3 region was unoccupied. Molecular modeling analysis indicated the best improvements would be realized by substituting at the 3-position of the 2-naphthyl ring. Optimization was initiated and JNJ-10311795 was identified as a reversible competitive inhibitor of CatG with an IC_{50} improved to 82 nM. At this time JNJ-10311795 was assayed for enzyme selectivity and it was revealed to be a potent inhibitor of chymase ($IC_{50} = 4.5$ nM). The compound was selective over a variety of other serine proteases including trypsin, trypsin, factor Xa, leukocyte elastase and thrombin. Crystal structures were obtained for CatG and chymase with JNJ-10311795. The crystal structure with CatG (Figure 4) showed the 2-naphthyl ring occupying the S₁ pocket and the 1-naphthyl ring in the S₂ pocket involved in pi-stacking interactions with His-57. As expected the substituent in the 3 position was occupying the S₃/S₄ cavity defined by Tyr-215, Ile-99 and Phe-172.

Figure 4



Chymase shows >50% sequence homology and ~80% active site homology with CatG. It was expected to bind in a similar manner to CatG. The crystal structure of chymase and JNJ-10311795 (Figure 5) confirmed the binding in the S₁/S₂ domain was nearly identical to CatG. Surprisingly the crystal structure revealed there is no binding of this inhibitor in the S₃/S₄ pocket. Instead the acylpiperidine portion folds back on the molecule and the main interaction observed is π -stacking with the third naphthyl portion of the molecule.

Figure 5

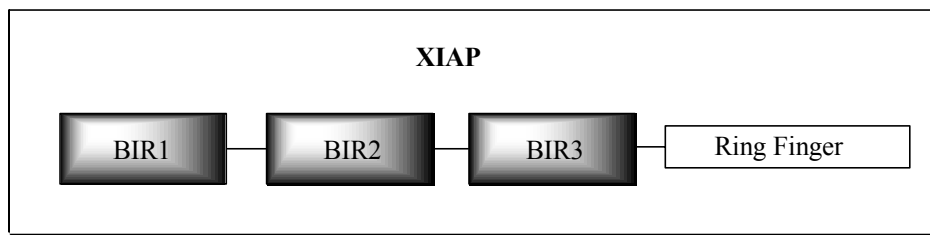


JNJ-10311795 has shown to be effective in rodent models for acute peritonitis and acute airway inflammation. Levels of neutrophils, cytokines IL-1 α , IL-1 β and TNF α and chemokine MCP-1 were reduced. As a result, JNJ-10311795 has entered human clinical trials for the treatment of asthma and COPD. Aerosol administration of this compound is necessary owing to its low oral bioavailability and short half-life.

“NMR-Based Drug Design: Discovery of Potent Antagonists of Anti-apoptotic IAP Proteins,”
Chaohong Son (Abbott Laboratories).

Inhibitor-of-apoptosis proteins (IAP's) are an important class of proteins involved in regulating programmed cell death in many organisms. IAP's function is to inhibit certain members of the caspase family of cysteine proteases. Caspases are involved in a cascade of events leading to cell death. Disruption of this delicate balance can lead to a variety of diseases including cancer. Analysis of human x-linked inhibitor-of-apoptosis protein (XIAP) was the focus of this presentation. XIAP is characterized by three repeating units called BIR domains, which are characterized by having 70 amino acids having a CX₂CX₁₆HX₆C key sequence (Figure 6). XIAP also has a ring finger portion which binds two zinc atoms.

Figure 6



The BIR2 domain is known to inhibit caspase-3. A truncated version of XIAP was created that encompassed the BIR2 domain and had a well-defined NMR derived structure. This version had similar inhibition compared to the wild-type protein. Site-directed mutants were prepared and were tested in inhibition studies for caspase-3. It was found that a group of conserved residues in the BIR1-BIR2 linker region were necessary for inhibition. Peptides containing only the linker region were not potent, so a combination of the linker region and the BIR2 domain were needed for potency. NMR- based binding studies were also performed using isotopically labeled XIAP mutant and caspase-3 mutants. Results from these experiments suggest that the linker region is binding to the active site of caspase-3 and the BIR2 domain is binding to an adjacent site on the protein, leading to a tightly bound complex.

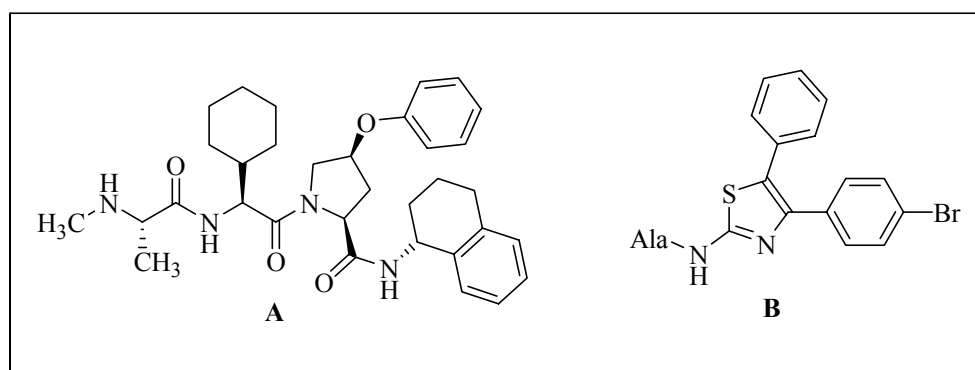
The BIR3 domain of XIAP is the region that inhibits caspase-9. The structure of the BIR3 domain was determined using NMR studies of a truncated analog. The three-dimensional structure is similar to BIR2 and resembles a classical zinc finger structure with a zinc atom chelated to 3 cysteines and 1 histidine. Again site-directed mutagenesis studies were used to determine the important residues needed for activity. Results of this study indicated that only the BIR3 domain is needed to potently inhibit caspase-9. Residues in the BIR2-BIR3 linker region were not needed for activity. Although the structure of the BIR3 domain is similar to the structure of the BIR2 domain, it seems XIAP is able to inhibit caspase-3 and caspase-9 using separate mechanisms.

Smac is another protein that was discovered that binds to XIAP. In this instance it serves to activate caspases. Peptides based on Smac were found to bind tighter to BIR3 than BIR2, so a NMR solution structure was obtained for BIR3 domain with a Smac 9-mer peptide that showed similar activity to mature Smac peptide. Overall, the structure of the complex was similar to free BIR3 domain. An electrostatic interaction was observed between Glu-314 and the N-terminus of the Smac peptide. Only the first four residues from the N-terminus of the Smac peptide showed interactions with BIR3. Structural changes to the Smac peptide and subsequent NMR based fluorescence binding assays corroborated these findings. A 5-mer peptide was synthesized and was found to retain all of the affinity of the 9-mer peptide. Mutants of BIR3 domain were also

synthesized and binding data was collected. In most cases, the ability of the BIR3 mutants to bind Smac peptide correlated with its ability to inhibit caspase-9. It was hypothesized that Smac and caspase-9 have overlapping binding sites within BIR3 and Smac exerts its effects by displacing caspase-9 from its complex with XIAP.

The first step in designing small molecule drugs that are Smac mimics was to make peptide libraries based on the Ala-Val-Phe-Ile tetrapeptide that interacts with the BIR3 domain. The parent tetrapeptide has a $K_D = 0.85 \mu\text{M}$ with the BIR3 domain. Results from this exercise showed converting Ala to N-Me-Ala and Ile to Phe increased affinity. To begin the transformation to a more druggable candidate, the tetrapeptide was converted to a capped tripeptide, where the Ile position was converted to a more stable moiety. The NMR solution structure of BIR3 with some of these capped tripeptides showed hydrophobic substituents could be added to the proline moiety to improve affinity by filling a vacant binding pocket. All this work culminated in **A**, which now has $K_d = 5 \text{ nM}$ (Figure 7). This compound was shown to provide relief of caspase inhibition by BIR3 and XIAP (caspase activity = 0.24 and 0.71 μM respectively). This and other similar compounds have shown activity against the MDA-MB-231 breast cancer cell line and have potent in vivo activity. Work has also begun making non-peptide small molecule inhibitors of XIAP. The best compound reported was **B** with a $K_i = 0.74 \mu\text{M}$.

Figure 7



“Structural Biology and Drug Discovery: Opportunities and Challenges,”

Tom Blundell (University of Cambridge and Astex Therapeutics).

The use of three dimensional structures of proteins to help guide drug discovery is an important way to advance compounds through lead optimization. Through advances in structural biology and genomic mining there are many more protein targets available for possible therapeutic applications. Unfortunately, the technology available and the time needed to crystallize complexes is often a roadblock for fully utilizing the vital information it provides to guide drug optimization.

In addition to being a professor, Tom Blundell is also a cofounder of Astex Therapeutics in England. Astex is a research and development company that has pioneered many new methods for using protein-ligand structure analysis to guide drug discovery. It is well documented that traditional high-throughput screening has not performed up to expectations for advancing more compounds into clinical trials. Astex is one of many companies employing fragment based drug

discovery. In this method small molecular fragments are designed to sample chemical space or target specific protein interactions and obey the Astex Rule of 3TM. These are drug-like fragments that have MW <300, H-bond donors ≤ 3 , H-bond acceptors ≤ 3 and ClogP ≤ 3 . Using PyramidTM, the drug discovery engine developed at Astex, the data obtained from analysis of these fragments can lead to the evolution of a drug lead and aid in optimization to candidate stage. A large part of their program is based on HTX[®] technology which uses high throughput x-ray crystallography techniques to probe ligand-protein interactions. These small fragments can show low affinity and might not be detectable in binding assays. X-ray crystallography is a useful way to explore and define these weak binding interactions (μM - mM). Through the use of Autosolve[®] software developed at Astex, this high-throughput method can reduce times to generate a crystal structure down to the order of minutes. To complement the crystallographic data, NMR spectroscopy is also used to probe these weak binding interactions. NMR analysis can reveal new interactions and probe protein function. In this process fragments are first chosen from computational analysis of virtual compound libraries. In silico docking experiments are run as a filter to create a subset of compounds. This subset is put into high-throughput crystallization assays. The compounds can be placed in the assay as single compounds or more frequently as cocktails of 4-10 compounds. There is no need to deconvolute the results from the cocktail because a positive result is the visual presentation of the fragment bound to the protein. From these results, hits are generated and can be built out to create an optimized compound.

Astex has used this technology to bring several compounds to advanced stages. The focus of their programs is to target proteins involved in the pathology of cancer. Their lead candidate entered phase 1 clinical trials in 2005. AT7519 is a CDK-2 cell cycle inhibitor that will be given intravenously to patients with solid tumors. AT9283 is being developed as an inhibitor of mitosis. AT9283 targets aurora kinases, which are involved in cell division and are over-expressed in a variety of tumors. Clinical trials could begin for this compound in 2006. An orally active cell cycle inhibitor (AT9311) is in preclinical development and could also begin clinical trials next year.

“Fragment-Based Drug Design in Large Pharma,”

Deborah A. Loughney (Bristol Myers Squibb Co.), USA.

Fragment-based drug design (FBDD) has emerged as an alternative approach to traditional lead identification and is increasingly being used in large pharmaceutical settings to improve the drug discovery cycle. Dr. Loughney discussed the rationale for considering a FBDD approach in drug discovery and its application to two targets at various stages of development. The results from these projects highlighted the strength and shortcomings of FBDD in assisting drug discovery.

Recently, early drug discovery has relied heavily on High Throughput Screening (HTS) to identify hit compounds for targets of interest. Because HTS is usually performed at lower concentrations (typically 10-30 μM), the resulting hits frequently have properties more reminiscent of drugs than lead compounds. HTS hits are frequently large and complex molecules with poor solubility and ADME profiles. These hits are thus poor candidates for the initiation of a chemistry program because these programs usually result in molecules with increased molecular weight and logP. Furthermore, the potency of ligands is not linearly related to increased molecular weight and in

fact tapers off for compounds consisting of more than 15 nonhydrogen atoms,¹ making the design of analogs to increase potency from an already large molecule more challenging.

From their utilization in combinatorial chemistry to computational chemistry, the use of fragments or small molecules in drug design is not a new concept, but has only recently been considered at the earliest stages of drug discovery. The idea is that small molecules with an inherently lower binding affinity would make better starting points for a chemistry program. Starting from a low molecular weight compound with optimum binding to the protein, chemical modifications that increase molecular weight can more easily be applied to optimize the drug profile of a compound.

Dr. Loughney discussed the tools necessary to embark on a FBDD approach and their use on two projects which served as proof of principles. The first target had already progressed to a stage where a large amount of information was already available. This included a good protein supply with crystallographic structures and a chemistry program that was already advanced. The second target had not yielded crystal structures and no good leads had been identified.

The team at Bristol Myers Squibb had access to all the components needed to embark on a FBDD approach. The computational chemistry team performed the fragment selection for initial screening and the further development of candidate structures. NMR techniques were developed to enable measurements at the higher concentrations that are necessary for fragment based screening. The subsequent analysis of bound fragments was investigated using either protein NMR or protein crystallography techniques. Lastly, the medicinal chemistry and combinatorial chemistry groups synthesized analogs based on the bound fragments using the available structural information for lead generation.

For the first target, they started with a fragment based virtual screen performed using a docking program with structural information from over 20 in house crystal structures. Compounds which possessed less than 30 atoms and less than five rotatable bonds were selected to serve as fragments. The results from the virtual screen were filtered and a small number of compounds were selected for NMR screening. Along with a visual inspection of the docked fragments, the selection process took into consideration the availability of fragments and their potential for facile analog synthesis. The compounds were screened by 1D NMR techniques and confirmed to bind in the active site by chemical shift perturbation mapping. Of the initial 62 fragments selected for NMR screening, 25 were confirmed to bind in the active site with 12 compounds showing >20% inhibition at 100 μ M and three compounds with >95% inhibition at 100 μ M in a cell based assay.

The structural information was collected to allow for elaboration of the chemical structures. Crystal and NMR structures of selected fragments showed an empty hydrophobic pocket and indicated an interaction with a water molecule that could be optimized. CADD designed compounds to address these issues and this resulted in core structures which were good starting points for library synthesis.

The second target had proven difficult and the team needed a new method in an effort to identify a lead. High throughput screening had failed to identify hits and the lead series they had been pursuing possessed too many liabilities. The FBDD approach proceeded similarly to the approach taken for the first target but started from a much larger set of potential fragments. Ten thousand

¹ Kuntz et al., *Proc. Natl. Acad. Sci. U. S. A.* **96** (1999) pp 9997-10002.

fragments were identified by virtual screening and run through a high concentration HTS assay. The top five hundred fragments from the virtual screen were also screened by NMR. For both the HTS screen and the NMR screen, binding fragments were confirmed to bind in the active site by NMR chase experiments.

The results from the screening campaigns for the second target showed that the fragments identified by HTS also contained the fragments identified by the NMR screen. The resulting structures were used by modeling to suggest modifications to the hit fragments and complex molecules that contained these fragments were found by standard HTS screening methods. Analoging synthesis around these initial hits resulted in a sub-micromolar lead compound. This initial lead compound was a desirable starting point for the medicinal chemistry team and the current leads which they have developed are potent inhibitors (20 nM in enzyme assays and 25 nM in cell based assays).

The results for the two targets presented by Dr. Loughney showed that FBDD is a powerful alternative approach for lead finding. Use of this approach is not always appropriate however. In the case of the first target, the FBDD results, while successful, showed that the iteration cycle was longer than desired to impact a program that was already well underway. The second target however, showed the strength of a fragment based approach even in the absence of structural information. These elegant studies presented by Dr. Loughney showed the potential of using small weak binders, an approach they term “Small is Beautiful”, to facilitate the drug discovery cycle. Dr. Loughney also identified the need for a commitment to screen at high concentrations and the development of NMR screening libraries as two areas of development for successful fragment based drug design with a shorter iteration cycle.

“Modeling Ligand Binding to the hERG K⁺ Channel: Application in Drug Discovery,”

Brett A. Tounge (J&J Pharmaceutical Research and Development, L.L.C.), USA.

Blockage of the human ether-a-go-go related gene (hERG) K⁺ channel by certain drugs has been linked to potentially fatal heart arrhythmias. Medications that prolong the repolarization of the cardiac action potential increase the QT interval of the EKG and result in a so called drug-induced long QT syndrome (LQTS). To date, 31 drugs are known to induce LQTS with a further 35 on a warning list.² Prediction of hERG K⁺ channel inhibition remains difficult due to a lack of structural information. Dr. Tounge discussed the current understanding of the hERG K⁺ channel and detailed computational efforts at J&J to develop computational models for prediction of hERG K⁺ channel binding.

The hERG K⁺ channel is a voltage gated channel formed by the assembly of four subunits each consisting of six transmembrane helices. The central portion of the channel is formed by a portion of hERG which consists of a so-called P-loop, a selectivity filter and the terminal helix (S6). This portion of the channel bears a strong resemblance to other potassium channels for which crystal structures are available. The homology model of the hERG K⁺ channel was built using the structure of the bacterial KcsA K⁺ channel. The protein alignments were based on the overlap of the selectivity filters which are involved in the transport of potassium to the extracellular side and are highly conserved features in K⁺ channels. The resulting structure indicated the orientation of

² <http://www.qtdrugs.org>

Y652 and F656 towards the inside of a cavity below the selectivity filter. Available mutagenesis data for these two key aromatic residues confirmed their involvement in ligand recognition.

The initially homology model was evaluated using 32 compounds from the literature with measured hERG binding affinities. These compounds were docked into the pore region of the hERG channel model. A training set of 15 compounds was defined to develop a linear interaction energy (LIE) model for this system. The training set consisted of: astemizole, cisapride, droperidol, thioridazine, mizolastine, bepridil, azimilide, mibefradil, chlorpromazine, imipramine, fexofenadine, diltiazem, sparfloxacin, grepafloxacin and gatifloxacin. These compounds had a pIC₅₀ range of 3.89 to 9.04. A plot of the experimental pIC₅₀ values versus the pIC₅₀ values resulting from the derivation of the LIE model gave poor results with an R² value of 0.28.

The initial KcsA K⁺ channel structure used to develop the original binding model was for a potassium channel in the closed state. For the channel to conduct ions, it must create an opening through which the ions can flow. This flexibility involves the movement of one end of the S6 helix in each subunit over several Ångströms and was captured in the crystal structure of the MthK K⁺ channel in its open state. Electrophysiology data of hERG inhibitors indicates that most compounds bind to an open state of the hERG K⁺ channel and a few bind to the closed state. Furthermore, Dr. Tounge described how the channel may close to varying degrees depending on the bound ligand. They wanted to capture this flexibility in a model.

The current homology models of the hERG K⁺ channel based on the closed KcsA K⁺ channel and the open MthK K⁺ channel were investigated and it was determined that these models gave unreliable pictures of the open and closed states of the hERG K⁺ channel. The homology model based on the closed KcsA K⁺ channel resulted in a hERG K⁺ channel where the F656 residues were packed too closely. The homology model based on the open MthK K⁺ channel on the other hand, gave the picture of a hERG channel that was too open for most inhibitors to form interactions with the F656 residues. Dr. Tounge concluded that intermediate descriptions of the hERG channel were necessary.

The generation of intermediate hERG K⁺ channel structures was performed using overlapped MthK and KcsA structures to define an axis of rotation for the S6 helix, since it is rotation about a Glycine residue in this helix which is responsible for the opening and closing of the channels. Rotational increments of 1° were investigated. To derive a model which could account for the binding of structures to either the closed or open hERG K⁺ channel states, two model channels were selected: one partially open and with a tilt in the S6 helix of 10° and one fully open with a tilt in the S6 helix of 19°.

The new hERG models were evaluated using the same procedure as the original KcsA based homology model. To develop LIE models for each of the states, protein-ligand complexes were evaluated and only one of these complexes was selected based on the sum of the electrostatic and van der Waals terms. This resulted in a set of 11 structures to develop a partially open state model and 21 structures for the open state model. These LIE models showed better correlation between the predicted and experimental IC₅₀ values than correlation derived from the LIE model based on the KcsA structure. In the new LIE model, it was a combined dual-state model for 27 ligands that gave the best results with an R² of 0.82 (clozapine, haloperidol, norastemizole, pimozide, and sertindole were outliers and were removed from the final regression).

Docked poses were investigated in both the partially open and open models for the inhibitors astemizole and cisapride. Astemizole showed a marked preference for the open model with much lower interaction energy between the ligand and the protein. Using the open state model, the predicted pIC_{50} for astemizole was 8.83 versus an experimental value of 9.04. Cisapride showed preference for the partially open model and resulted in a predicted pIC_{50} value of 7.31 versus an experimental value of 8.19. In both cases, the Y652 and F656 residues of the hERG K^+ channel formed strong interactions with the inhibitors as expected from mutagenesis data.

The homology models provide structural details that can be used when designing analogs with reduced hERG binding. Dr. Tounge indicated the possible need to look at different intermediate models in an effort to identify the interactions between the hERG K^+ channel and the inhibitors which were identified as outliers during the study. The present study indicated that homology models should be developed with care and that significant structural changes in the hERG K^+ channel may accompany inhibitor binding. This was evident by both the need to develop structural models which deviated from the current homology models based on KcsA and MthK, and also on the observation that a multiple state model gave superior results for hERG K^+ channel binding prediction than did the single state models.

“Structure-Based Design: The Long View,”

Irwin D. Kuntz (University of California at San Francisco), USA.

Dr. Kuntz gave an overview of the use and impact of structure-based drug design (SBDD) techniques in drug discovery. Computational tools are a valuable asset to drug discovery programs because they can be used to provide a visual image of a molecule's binding mode. Information from SBDD techniques can be brought together with information from experimental techniques to provide an in depth understanding of the interactions of the ligands with their targets. Currently, SBDD techniques are powerful but still difficult to use and do not always provide reliable results thus requiring careful analysis. Dr. Kuntz discussed SBDD applications and areas of difficulty where improvements are being pursued.

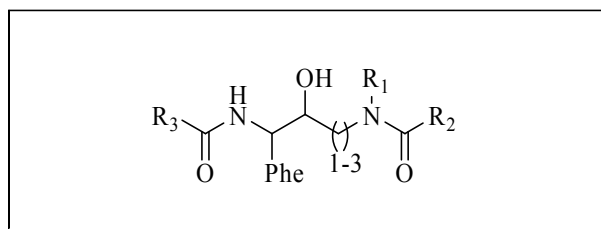
A wealth of structural information is now available which can be harnessed to guide drug discovery programs. The structural information may come from crystallographic and NMR data or may even be generated by computational techniques such as homology modeling. It is worth noting that even with access to structural information, it is still much easier to discover drugs (most frequently using screening techniques) than it is to design them. Furthermore, de-novo design techniques which build molecules up using fragments from an initial core are more successful than techniques which try to design molecules atom-by-atom. Designing an entire molecule using computational tools remains very difficult.

Docking strategies are one of the more successful modeling tools available to today's computational chemist when structural information of the target protein is available. Integration of molecular docking at various stages such as early screening, library design or compound optimization highlight the success of this technique in the drug discovery setting. Dr. Kuntz described several case studies where computational tools were used and were found to have a strong impact on the drug discovery project.

The first case study was aimed at mining a database to find compounds that could interact with TAR complexes of HIV. The structure from an NMR TAR-RNA complex could be harnessed to perform virtual screening using high throughput molecular docking. Obtaining good enrichment factors from virtual screening still depends on the purity of the compounds used for mining and on the quality of the library being searched. The development of further compounds is also sometimes complicated as the hit compounds might not be appropriate leads even if they are potent ligand for the selected target. Successful HTS screening or virtual screening do not always make a good start for a chemistry campaign.

The second case study concerned the design of libraries for screening of two targets: one for aspartyl protease inhibitors and the other for imidazole cyp450 inhibitors. In this case, candidate structures were more frequently good chemical starting points as many are selected with input from synthetic chemists. For the first target, given a starting scaffold and the structure of the target, the goal was to design a library that explores various sidechain orientations around the cathepsin D family of aspartyl proteases (see Figure 1). Issues that required particular care concerned stereochemical considerations as

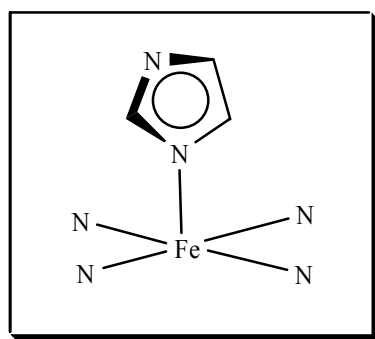
Figure 1



well as the conformational space available to these peptides. Comparison of the screening results for the library that was designed and a library containing diverse compounds showed that the resultant library yielded a larger number of inhibitors with increased potency.

The design of P450 inhibitors for the second target required a changing in the docking function to obtain an appropriate representation of imidazole interactions with the heme group of the enzyme (see Figure 2). Once this was performed, selection of compounds to

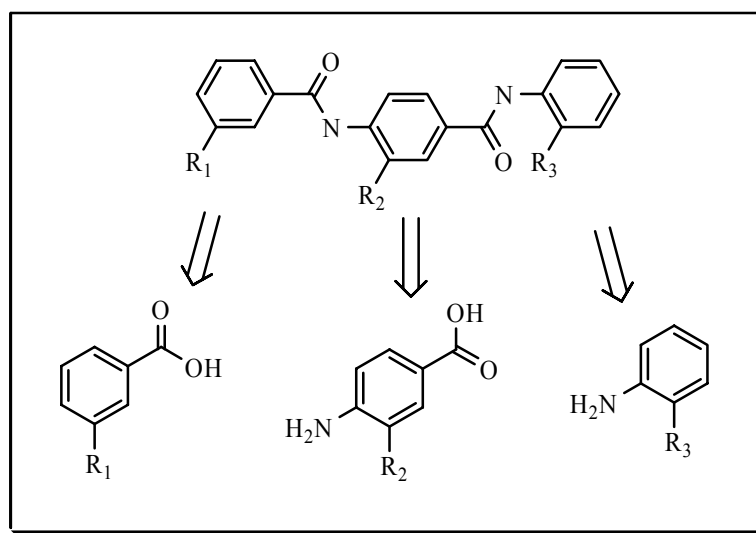
Figure 2



form a library was complicated by the flexible nature of the active site. The presence of water molecules which further stabilize the interactions between the putative ligands and the active site proved a further challenge. Lastly, the selection of compounds to serve as scaffolds for a library is difficult as small fragments usually poses information but are frequently weak binders. Nevertheless, this can be a powerful SBDD approach for drug discovery.

For the third case study, the aim was to design new inhibitors that mimicked a helical peptide to inhibit the formation of p53-MDM2 complexes. The design started with

Figure 3



replacements of the peptide backbone with semi-rigid organic scaffolds that could mimic p53 bound to the large hydrophobic groove in MDM2. Using a starting structure of p53 interacting with MDM2, evaluation of the key interactions yielded distance patterns that were used to design a scaffold using the tool CAVEAT.³ Once an appropriate scaffold had been chosen, the combinatorial library was built using a number of side chains which were selected to build from three attachment points (see Figure 3).

Currently, various improvements are still being targeted at molecular docking approaches for SBDD. Sampling of ligand and receptor conformations, including consideration of tautomeric and ionization states, continues to be an important area. The scoring functions that are used to rank order ligands following molecular docking still need a lot of improvement. These scoring functions are frequently based on force fields and errors are frequently linked to the choice of van der Waals radii. Also, it is still not clear whether an absolute prediction of binding affinity should be the goal of scoring functions. In many cases, relative ranking was successfully achieved as long as the series of ligands being investigated are within the scale that scoring functions were developed for. Also challenging is the calculation of contributions to ligand binding from solvation and entropic effects which remain time intensive and difficult to include. It is therefore

³ Lauri, G.; Bartlett, P.A., "CAVEAT: A Program to Facilitate the Design of Organic Molecules" *J. Comp. Aided Mol. Design* **1994**, 8, 51-66

still difficult to truly predict whether a compound that “fits” in the active site is a good compound for a target protein.

Dr. Kuntz highlighted the use of SBDD techniques to help find hits, design parts of molecules and even to design molecules in their entirety. From each study, it was evident that it is an interdisciplinary approach that is the most successful with different teams bringing key tools and information to the drug discovery process. Nevertheless, it was also evident that careful evaluation of the computational tools and their modification were also crucial for successful SBDD. Computational tools are now widely used and they are capable of strongly impacting drug discovery, but the optimal application of computational techniques is at the moment still far from being automated or even easy to implement. The goal remains to produce accurate results in a timely fashion so as to impact a drug discovery program.