

# Computational Techniques in Fragment Based Drug Discovery

Hugo O. Villar\* and Mark R. Hansen

ALTORIS, Inc. 7660-H. Fay Ave #347, La Jolla, CA 92037

**Abstract:** Fragment based drug discovery is gaining acceptance as a complement to other more established techniques to identify leads and optimize drug candidates. In this review we illustrate areas where fragment based drug discovery has had an impact and point to some examples that show how fragment based analysis is being applied to new arenas.

The traditional uses of computational methods in fragment based for lead discovery and optimization and for risk assessment are briefly summarized. The application of fragment analysis for the definition of bioisosteric replacements are discussed together with techniques to characterize the diversity of chemical libraries based on fragment distribution.

## INTRODUCTION

Implicit in the lead discovery and optimization stages in drug design is the idea that molecules are modular. Each molecular fragment contributes its own properties to the drug candidates. Researchers have relied on the modular nature of molecules since the early days of drug discovery. Functional groups are swapped for others in order to improve potency or conversely, substitutions are carried out to preserve the overall affinity but alter some other property of the molecule to increase its drug like characteristics. Molecular fragments have a long history in medicinal chemistry work.

Concurrently, many *in-silico* methods for prediction of molecular properties have been based on the decomposition of molecules into their parts to develop predictive models [1]. Fragment based techniques were pursued for a long time in the prediction of some physicochemical properties [2], such as logP, and several biological characteristics including toxicological studies [3]. In the recent years, fragment analysis contributions to discover and optimize lead molecules or to predict properties had generated great interest as an alternative paradigm for drug discovery and had a leap forward due to numerous new approaches [4,5].

The *de novo* construction of ligands based on insights gained by the study of the intermolecular interactions of low molecular weight molecules has become a promising approach for lead discovery and optimization [6]. The initial challenge for lead discovery based on fragment assembly is to identify very low molecular weight ligands, likely to also be low affinity. The second part of the challenge is to determine the way in which those fragments should be linked, including the nature of the linker moiety. Other significant challenges may exist, as chemistry can be difficult to perform in the regions where modifications are needed.

The low affinity of the very low molecular weight compounds were a trial to the implementation of fragment based discovery. Recently, it has been overcome with a new battery of biophysical methods to detect weak interactions. Indeed, fragment based drug discovery has become successful only with the tight integration of computational techniques, innovative Nuclear Magnetic Resonance (NMR) experiments and high throughput X-Ray crystallography [7].

The application of computational techniques in fragment based drug discovery is presented. The used of fragment based methods for lead discovery and optimization, including toxicological profile prediction have been previously reviewed [1-6] therefore we focus here on some particular developments that merit some discussion.

New applications of substructural analysis can provide alternative analysis tools for traditional chemoinformatics problems.

Among those reviewed, uses of substructural analysis in the definition and identification of bioisosteric replacements and the assessment of the diversity of chemical libraries are summarized.

## APPROACHES FOR FRAGMENT PLACEMENT AND LINKING

While the last decade has seen an acceleration of the development of techniques for fragment assembly, those techniques were already delineated some years back. Multiple solvent crystal structures introduced a few years earlier provided a way to characterize protein binding sites [8,9]. The use of different solvent molecules aided in the characterization of binding pockets in the protein, which in turn could be exploited to design ligands. The solvents indicated preferred positions for functional groups, which could potentially be assembled.

In parallel, computational techniques were developed to search for optimal positions and orientations of a set of solvent type molecules, by using docking methods. The multiple copy simultaneous search (MCSS) method [10] was a technique in this class. Even at this early stage the use of experimental information went hand in hand with computational tools. Other earlier fragment based computational techniques included LUDI [11] and can be used to identify the anchoring points in proteins and suggest replacement fragments or places to add given substructures. More recently, large fragment spaces were assembled following a sequential growth strategy, with several thousands of chemical fragments and a set of rules to specify fragments connectivity [12].

For some time, further development of fragment based methods was hindered due to the lack of techniques amenable to the detection of the very low affinity displayed by the small compounds required by the approach. While the study of the solvent binding had been done with the intent of linking binding spots in the protein, the procedure could not be done in a high throughput mode as it would be required of a modern technique for drug discovery. In subsequent years, biophysical methods expanded the repertoire of techniques available to the researchers and allowed determinations of even marginal target-ligand interactions. The introduction of SAR by NMR provided the first systematic use of biophysical techniques [13-14], which allowed the rapid generation of maps of the relative orientation of ligands in the binding sites. Once the evaluation of low affinity was possible in a high throughput mode, interest greatly increased in the approach as a complement to the dominant paradigms of library screening. Indeed, NMR guided techniques have been one of the preferred tools for fragment assembly. NMR provided a means for the screening of very low affinity ligands and to determine their relative orientation - for that reason it has been applied to various systems, as extensively reviewed in the literature in recent years [15-17].

\*Address correspondence to this author at the Altoris, Inc. 7660-H. Fay Ave #347, La Jolla, CA 92037; E-mail: hugo@altoris.com

X-ray crystallography for assembling molecules from fragments has also been used for quite sometime [18]. X-ray crystallography [19] can arguably have some advantages over NMR spectroscopy because of the detailed information it can provide about the ligand binding site and its orientation. Despite advances in automation for X-Ray crystallography [20] NMR is easier to set up for high throughput screening and provides an even faster turn around. Both techniques have a place, and ideally, should be used in conjunction.

A problem that has been studied less frequently and where additional effort should be expected in the computational techniques in the near future is on the linking of the groups placed by the different techniques. Methods such as HOOK [21] can be used to link the positioned fragments provide a complete molecule from the fragments. However, new automated techniques to design linkers with higher precision are still needed.

### FRAGMENT LIBRARY DESIGN

An important component of fragment based drug design methods is the initial collection of very low molecular weight ligands to be explored [22]. Libraries of molecules have been assembled with diverse uses in mind. For examples, fragment libraries can be broadly applicable or targeted to a specific protein class. An example of a broadly applicable library is SHAPES [23], derived largely from molecular frameworks most commonly found in known therapeutic agents. In this library, compounds from a medicinal chemistry database were classified by shape descriptors that included their ring structures and linkers, without regard for atom types or bond order. At this basic level, 32 frameworks encompassed half of all known drugs. When atom type and bond order were included, 41 frameworks represented nearly a quarter of known therapeutics.

The same trend towards specialized or focused libraries observed for conventional screening is also seen in the design of fragment based libraries. A survey of fragment-based lead discovery shows that approximately half of the hit fragments were discovered by screening generic libraries. In the other cases some knowledge about an initial ligands or the protein binding site was used. For example, protease inhibitors and substructures present in factor Xa inhibitors were used to uncover new leads for factor Xa [24], or a similar approach for kinases [25]. Systematic virtual screening of fragment databases has been only rarely reported [22].

Perhaps more interesting is the development of fragment based libraries to aid in chemoproteomics work. A new type of compound library was described that focuses on aiding in the functional annotation of novel proteins that have been identified from various ongoing genomics efforts. The functional chemical library comprised of small molecules with known biological activity such as, co-factors, inhibitors, metabolites and substrates, was assembled for NMR. This functional library was developed by mining several databases based on known ligand interactions with protein systems. [26]

Fragment-based approaches require their own rules, more stringent than those for drug-like characteristics, because the compounds identified in the screening of fragments are meant to be a substructure in the larger, final drug-like molecule. Simple rules have been put forward for that purpose [27]; the rule-of-three for fragment-based drug discovery parallels the ubiquitous rule-of-five for drug like characteristics. In this rule, the useful criteria for fragment library selection include molecular weight of less than 300 Da, the number of hydrogen bond donors or acceptors less than or equal to 3, calculated ClogP less than 3, number of rotatable bonds less than or equal to 3, and polar surface area of less than  $60 \text{ \AA}^2$  [27]. These observations are based on results from an analysis of a diverse set of fragment hits that were identified against a range of targets.

The issue of affinity needed for the fragment - target interaction was studied in a retrospective analysis of 18 highly optimized inhibitors is described in which the compounds were systematically deconstructed until the minimal binding elements could be identified. The potency changes showed approximately a linear relationship between molecular weight and binding affinity over the entire range of sizes and potencies represented in the dataset [28]. However, the deconstructive analysis suggests that some of those simple relationships may be deceiving, as the fragments may not adopt similar spatial interactions with the target than those observed when embedded in the final molecules [29].

An observation made some time ago has significant implication for the assembly of libraries for fragment based drug design. Simply stated, as a molecule becomes increasingly complex, that is more features are present in the molecule [30], the chance for interactions good and bad is also increased. Therefore, libraries of high complexity molecules have a low chance of individual molecules binding because of the increased potential for bad interactions. Fragment libraries, conversely, have less chance of a bad interaction, because they are feature poor, but for the same reason they are typically of modest activity. When working with complex molecules, the bad interactions are removed in the lead optimization process. In the case of fragment based methods, few or no bad interactions can exist and it is a question for the medicinal chemist, not to introduce them in the process of linking fragments to increase affinity and selectivity.

### FRAGMENTS AND LIGAND BINDING SITES

An important issue is the progress in experimental and computational methods for identifying and characterizing druggable ligand binding sites on protein targets by nuclear magnetic resonance, X-ray crystallography and tethering technologies. Among the computational techniques to identify valuable sites are those based on geometric features or energy-based computational methods. Solvent mapping and grand canonical Monte Carlo simulations can be used to reliably identify druggable sites on proteins and to facilitate the design of novel, low-nanomolar-affinity ligands [31,32].

The design of libraries to aid in chemoproteomics work is in line with observations made over time regarding common features in binding sites. Small organic molecules bind to well-defined, localized regions of proteins [33]. These sites define hot spots that can provide the main recognition points for the small molecules. The linking of fragments using different of the main recognition points, can provide the need selectivity and affinity [34].

The increasing evidence for similarities in binding sites suggests that an effective way to create new ligands for proteins would be to anchor a promiscuous compound in the binding site and use flanking regions to gain specificity [35]. The similarity in binding sites provides the basis for modular approaches to drug design, such as the well known SAR by NMR technique or the SHAPES procedures. The concept has been exploited in a systems-based approach for gene families that share a common cofactor. The basic premise is to identify a mimic for the cofactor that has drug-like properties and can be used as the central scaffold for a parallel synthesis effort. The resulting libraries would contain chemicals able to bind several different members of that gene family. Many challenges have to be overcome. First, as discussed earlier, the cofactor or other common ligands do not share the same conformation for all members of the gene family. Once a protein family is subdivided into classes that bind the cofactor in a similar way and a mimic for the cofactor has been identified, the next challenge is to decide how to extend the scaffold into a pocket in the protein that would confer selectivity. Spectroscopic techniques such as NMR are ideally suited for this purpose [36]. The result is a ligand that spans more than one pocket in the binding site, with thermodynamic advantages for binding. An example using

oxidoreductases illustrates the approach very clearly. The existence of similarities in binding sites supports this and other modular approaches to drug discovery.

### FRAGMENT SWAPING: BIOISOSTERISMS

The evolution of a fragment to generate a new chemical class has followed different paths. For example, the three dimensional locations of fragments from existing inhibitor structures obtained from X-Ray data can be combined to create novel compounds. The superimposition of experimental structures can be compared to identify appropriate reassembly of starting ligands to arrive at new chemical entities [37]. This procedure suggests which fragments could be replaced by which others, based on the experimental information. Thus fragment based approaches generalize based on structural information the concept of bioisosteres. Such information that can be compiled with high throughput structure determination methods can provide a very valuable alternative to methods that are based on databases of biologically equivalent molecules. The structural information can clearly indicate which groups are equivalent without ambiguity.

Fragment based design is not new in chemistry and the problem of bioisosterism can be viewed as one aspect of it. The identification of atoms or groups of atoms (fragments) that can replace for each other and elicit the same biological activity has been studied for decades in chemistry. A recent article presents a new approach based on the automated extraction of bioisosteric fragments. The automation leads to the identification of a significantly large number of bioisosteric fragments for the purposes of validation. The automation results in an unbiased method that extracted more than 2200 true bioisosteric R-groups, linkers, and cores. The method allows the identification of find potential bioisosteric replacements on the basis of a topological pharmacophore fingerprint and a database of more than 700 000 structural fragments.

Some alternative techniques aim to modify the basic scaffold to create and predict new molecules [38]. Site-directed ligand discovery, in which a covalent bond is used to stabilize the interaction between low-affinity fragments and a target protein, is an alternative approach [39,40]. A fragment based technique for discovery of new ligands that is based on swapping fragments, using retrosynthetic analysis and can be applied with as little information as the structure of a single ligand [41]

### FRAGMENT BASED TOXICOLOGICAL DESIGN

Some of the pioneer techniques for chemical risk assessment were based on substructural or fragment based analysis. Methods such as CASE [42,43], to predict carcinogenicity, mutagenicity, teratogenicity, cellular toxicity, and allergic contact dermatitis rely on the enumeration of linear fragments that are culled from toxic or non-toxic compounds, and counts of fragments in each set are used to determine the fragment contribution that correlates to the toxicity of the compound. Identification of toxicophores continues to be of interest, and there is a clear preference to use substructural analysis in methods to predict toxicological profiles [44].

In part due to the availability of data, the development of fragment based studies that aim to predict and analyze mutagenicity are plentiful. New methods continue to revisit this important toxicological property [44]. The availability of data provides a fertile means to test new algorithms and methods for prediction and analysis.

Variants of those techniques continue to be applied and added to an arsenal of fragment based methods for toxicological prediction. Most recently, hybrid methods where a novel pattern-recognition method to characterize substructures within a molecule has been developed and applied to the prediction of torsadogenicity [45]

These methods, which are based on easily recognizable chemical features, allow chemists to immediately use biological information when designing new libraries. Descriptor-based techniques can have better predictions, but because the molecular descriptors used in such approaches are abstract or theoretical constructs that are difficult to understand and visualize, incorporation of the knowledge is hindered. Conversely, the immediate application of the knowledge gained from fragment or substructure-based methods is central to their appeal as predictive techniques [46-48].

Analysis of biases in the distribution of molecular fragments has been carried out for a large variety of biological properties, ranging from prediction of CNS activity [49] to identification of substructures that confer a bitter taste to foods [50]. Substructure-based analytical methods are also starting to be used for predicting compound permeability and plasma protein binding [51]. In the case of permeability, a method based on the use of partial least squares has been described and appears to be a useful complement to 3D-based methods, which have difficulty in dealing with the problem of passive diffusion. In some cases, correlating substructural analysis with biological information can lead to an insight into the mechanism of action of the compounds. A characterization of the chemical substructure fragments in asthmagens versus control compounds led to the observation that fragments with bifunctional moieties are strongly associated with occupational asthma hazard across a range of chemical substructures [52].

### FRAGMENT BASED DIVERSITY ASSESSMENT

The current preferred paradigm in drug discovery involves the characterization of large chemical collections. The main goal is to assess the diversity or focus of libraries based on their structural, physicochemical or even biological properties, encoded into fingerprints or hash codes, with a variety of different metrics and classification techniques. Although these are all techniques that are computationally efficient, they are not intuitive to medicinal chemists, who prefer to organize them according to homologous chemical series.

The computational techniques used in fragment based drug design can be applied to organize chemicals according to scaffolds. An issue that has been considered is the definition of metrics that can be used to assess the diversity of chemical libraries, based on the scaffolds they contain.

Different algorithms are used to define scaffolds from the maximum common substructure, to largest rigid fragment or based on retrosynthetic analysis [1,53]. Depending on the type of definition adopted and the algorithms used, the scaffold or scaffolds used to represent a molecule may vary. Establishing relations among scaffolds is another important aspect of classification based on scaffolds.

Some simple indices have been put forward to get an idea of the library diversity based on the number of fragments that are present. The studies that are emerging in the use of fragment based methodology to assess chemical diversity go beyond simple counts of residues or study of frequencies to the generation of new diversity indices. For example, a dispersion index that is the simple ratio of total number of scaffolds computed to the total number of molecules a library contains was used to compare a number of libraries [54]. The simple metric gives an assessment of the diversity of the library. An alternative is the percentage of classes accounting for 50% of the classified compounds [55]. Since those metrics are independent of the size of a library, they can be used to compare collections of different sizes. Moreover, union and intersection of scaffold collections can be found and it can provide significant information to carry out comparison of chemical collections.

A similar set of tools can be applied to study if the occurrence of one fragment is related to the occurrence of another. Indeed the

co-occurrence of certain fragments and combinations is quite frequent and has been termed a chemical cliché [56]. Studies of these clichés can help in the identification of novel patentable substructures and the identification of areas with freedom to operate.

### FUTURE DIRECTIONS AND SUMMARY

Fragment based ligand discovery and optimization is growing in significance to complement other techniques available. Its appeal lies in the intuitive nature of the process. Numerous examples are available in the literature that illustrates how these techniques are being employed, largely in a combination of structural techniques and computational methods. Fragment based ligand discovery has successfully been applied to many targets in different classes including kinase inhibitors [57]; protease inhibitors [58]; dehydrogenases [59, 60], among others [61].

Computational techniques that complement fragment based experimental methods are still evolving and significant work is still needed. The techniques used for study of fragment distributions and for fragment docking are available. However, techniques for mining fragment based approaches still need to be further enhanced and in some cases novel statistical techniques need to be brought into the field.

Predictive techniques based on fragment distribution are likely to continue to attract attention, given their intuitive nature. There is still room for improvement in the prediction of both, physico-chemical and biological properties.

Fragment based analysis of chemical collections, as for example in diversity analysis or hit to lead process, is just being explored. As the lead discovery and optimization processes rely more and more in the generation of vast amounts of data, tools to mine it effectively are useful to the bench scientist will continue to be pursued. Fragment based analysis plays a very important role in this front and it is likely to increase its significance with new tools and analysis methods.

### REFERENCES

- Villar, H.O.; Hansen, M.R.; Kho, R. Substructural Analysis in Drug Discovery. *Curr. Comput. Aided Drug Des.* **2007**, *3*, 59-67.
- Clark, M. Generalized fragment-substructure based property prediction method. *J. Chem. Inf. Model.* **2005**, *45*, 30-8.
- Dearden, J.C. In silico prediction of drug toxicity. *J. Comput. Aided Mol. Des.* **2003**, *17*, 119-27.
- Erlanson, D.A. Fragment-based lead discovery: a chemical update. *Curr. Opin. Biotechnol.* **2006**, *17*, 643-52.
- Carr, R.A.; Congreve, M.; Murray, C.W.; Rees, D.C. Fragment-based lead discovery: leads by design. *Drug Discov. Today* **2005**, *10*, 987-92.
- Keseru, G.M.; Makara, G.M. Hit discovery and hit-to-lead approaches. *Drug Discov. Today* **2006**, *11*, 741-748.
- Leach, A.R.; Hann, M.M.; Burrows, J.N.; Griffen, E.J. Fragment screening: an introduction. *Mol. Biosyst.* **2006**, *2*, 430-446.
- Mattos, C.; Ringe, D. Locating and characterizing binding sites on proteins. *Nat. Biotechnol.* **1996**, *14*, 595-599.
- Ringe, D.; What makes a binding site a binding site? *Curr. Opin. Struct. Biol.* **1995**, *5*, 825-829.
- Cafilisch, A.; Miranker, A.; Karplus, M. *J. Med. Chem.* **1993**, *36*, 2142-2167.
- Bohm, H.J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput. Aided Mol. Des.* **1992**, *6*, 593-606.
- Degen, J.; Rarey, M. FlexNovo: structure-based searching in large fragment spaces. *Chem. Med. Chem.* **2006**, *1*, 854-68.
- Shuker, S.B.; Hajduk, P.J.; Meadows, R.P.; Fesik, S.W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531-1534.
- Hajduk, P.J. SAR by NMR: putting the pieces together. *Mol. Interv.* **2006**, *6*, 266-272.
- Zartler, E.R.; Shapiro, M.J. Fragonomics: fragment-based drug discovery. *Curr. Opin. Chem. Biol.* **2005**, *9*, 366-370.
- Schade, M. NMR fragment screening: Advantages and applications. *IDrugs* **2006**, *9*, 110-3.
- Villar, H.O.; Yan, J.; Hansen, M.R. Using NMR for ligand discovery and optimization. *Curr. Opin. Chem. Biol.* **2004**, *8*, 387-391.
- Mattos, C.; Ringe, D. *Curr. Opin. Struct. Biol.* **2001**, *11*, 761-764.
- Blundell, T.L.; Patel, S. High-throughput X-ray crystallography for drug discovery. *Curr. Opin. Pharmacol.* **2004**, *4*, 490-496.
- Mooij, W.T.; Hartshorn, M.J.; Tickle, I.J.; Sharff, A.J.; Verdonk, M.L.; Jhoti, H. Automated protein-ligand crystallography for structure-based drug design. *Chem. Med. Chem.* **2006**, *1*, 827-838.
- Eisen, M.B.; Wiley, D.C.; Karplus, M.; Hubbard, R.E. HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a macromolecule binding site. *Proteins* **1994**, *9*, 199-221.
- Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A.L.; Jahnke, W.; Blommers, M.; Selzer, P.; Jacoby, E. Library design for fragment based screening. *Curr. Top Med. Chem.* **2005**, *5*(8), 751-62.
- Lepre, C.A.; Peng, J.; Fejzo, J.; Abdul-Manan, N.; Pocas, J.; Jacobs, M.; Xie, X.; Moore, J.M. Applications of SHAPES screening in drug discovery. *Comb. Chem. High Throughput Screen.* **2002**, *5*, 583-590.
- Fielding, L.; Fletcher, D.; Rutherford, S.; Kaur, J.; Mestres, J. Exploring the active site of human factor Xa protein by NMR screening of small molecule probes. *Org. Biomol. Chem.* **2003**, *1*, 4235-4241.
- Baurin, N.; Aboul-Ela, F.; Barril, X.; Davis, B.; Drysdale, M.; Dymock, B.; Finch, H.; Fromont, C.; Richardson, C.; Simmonite, H.; Hubbard, R.E. Design and characterization of libraries of molecular fragments for use in NMR screening against protein targets. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2157-2166.
- Mercier, K.A.; Germer, K.; Powers, R. Design and characterization of a functional library for NMR screening against novel protein targets. *Comb. Chem. High Throughput Screen.* **2006**, *9*, 515-534.
- Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. Today Fragment-based lead discovery: leads by design. *Drug Discov.* **2003**, *8*, 876-877.
- Hajduk, P.J. Fragment-based drug design: how big is too big? *J. Med. Chem.* **2006**, *49*, 6972-6976.
- Babaoglu, K.; Shoichet, B.K. Deconstructing fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2006**, *2*, 720-723.
- 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.
- Vajda, S.; Guarnieri, F. Characterization of protein-ligand interaction sites using experimental and computational methods. *Curr. Opin. Drug Discov. Devel.* **2006**, *9*, 354-62.
- Clark, M.; Guarnieri, F.; Shkurko, I.; Wiseman, J. Grand canonical Monte Carlo simulation of ligand-protein binding. *J. Chem. Inf. Model.* **2006**, *46*, 231-42.
- DeLano, W.L. Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol.* **2002**, *12*, 14-20.
- Arkin, M.R.; Wells, J.A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **2004**, *3*, 301-317.
- Ciulli, A.; Williams, G.; Smith, A.G.; Blundell, T.L.; Abell, C. Probing hot spots at protein-ligand binding sites: a fragment-based approach using biophysical methods. *J. Med. Chem.* **2006**, *49*, 4992-5000.
- Sem, D.S.; Bertolaet, B.; Baker, B.; Chang, E.; Costache, A.D.; Coutts, S.; Dong, Q.; Hansen, M.; Hong, V.; Huang, X.; Jack, R.M.; Kho, R.; Lang, H.; Ma, C.T.; Meininger, D.; Pellicchia, M.; Pierre, F.; Villar, H.; Yu, L. Systems-based design of bi-ligand inhibitors of oxidoreductases: filling the chemical proteomic toolbox. *Chem. Biol.* **2004**, *11*, 185-94.
- Pierce, A.C.; Rao, G.; Bemis, G.W. BREED: Generating novel inhibitors through hybridization of known ligands. Application to CDK2, p38, and HIV protease. *J. Med. Chem.* **2004**, *47*, 2768-2775.
- Aronov, A.M.; Bemis, G.W. A minimalist approach to fragment-based ligand design using common rings and linkers: application to kinase inhibitors. *Proteins* **2004**, *57*, 36-50.
- Erlanson, D.A.; Hansen, S.K. Making drugs on proteins: site-directed ligand discovery for fragment-based lead assembly. *Curr. Opin. Chem. Biol.* **2004**, *8*, 399-406.
- Erlanson, D.A.; McDowell, R.S.; O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* **2004**, *47*, 3463-3482.
- Fechner, U.; Schneider, G. Flux (1): a virtual synthesis scheme for fragment-based de novo design. *J. Chem. Inf. Model.* **2006**, *46*, 699-707.
- Klopman, G.; Zhu, H.; Fuller, M.A.; Saikhov, R.D. Searching for an enhanced predictive tool for mutagenicity. *SAR QSAR Environ. Res.* **2004**, *15*, 251-263.
- Rothfuss, A.; Steger-Hartmann, T.; Heinrich, N.; Wichard, J. Computational prediction of the chromosome-damaging potential of chemicals. *Chem. Res. Toxicol.* **2006**, *19*, 1313-1319.
- Kazius, J.; McGuire, R.; Bursi, R. Derivation and validation of toxicophores for mutagenicity prediction. *J. Med. Chem.* **2005**, *48*, 312-320.
- Bhavani, S.; Nagargadde, A.; Thawani, A.; Sridhar, V.; Chandra, N. Substructure-based support vector machine classifiers for prediction of adverse effects in diverse classes of drugs. *J. Chem. Inf. Model.* **2006**, *46*, 2478-86.
- Kho, R.; Hodges, J.A.; Hansen, M.R.; Villor, H.O. Ring systems in mutagenic database. *J. Med. Chem.* **2005**, *48*, 6671-6678.
- Johnson, D.E.; Rodgers, A.D. Computational toxicology: heading toward more relevance in drug discovery and development. *Curr. Opin. Drug Discov. Devel.* **2006**, *9*, 29-37.
- von Korff, M.; Sander, T. Toxicity-indicating structural patterns. *J. Chem. Inf. Model.* **2006**, *46*, 536-544.

- [49] Engkvist, O.; Wrede, P.; Rester, U. Prediction of CNS activity of compound libraries using substructure analysis. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 155-160.
- [50] Rodgers, S.; Busch, J.; Peters, H.; Christ-Hazelhof, E. Building a tree of knowledge: analysis of bitter molecules. *Chem. Senses.* **2005**, *30*, 547-57.
- [51] Clark, D.E.; Grootenhuys, P.D. Predicting passive transport in silico--history, hype, hope. *Curr. Opin. Drug Discov. Dev.* **2002**, *5*, 382-390.
- [52] Jarvis, J.; Seed, M.J.; Elton, R.; Sawyer, L.; Agius, R. Prediction of asthma hazard of glutaraldehyde substitutes. *Occup. Environ. Med.* **2005**, *62*, 243-250.
- [53] Kazius, J.; Nijssen, S.; Kok, J.; Back, T.; Ijzerman, A.P. Substructure mining using elaborate chemical representation. *J. Chem. Inf. Model.* **2006**, *46*, 597-605.
- [54] Hansen, M.R.; Hodges, J.A. Identification of Unique and Redundant Scaffolds in Chemical databases. The 232nd ACS National Meeting, San Francisco, CA, September 10-14, 2006. CINF 54; <http://www.chemapps.com/resources.html>
- [55] Krier, M.; Bret, G.; Rognan, D. Assessing the scaffold diversity of screening libraries. *J. Chem. Inf. Model.* **2006**, *46*, 512-524.
- [56] Lameijer, E.W.; Kok, J.N.; Back, T.; Ijzerman, A.P. Mining a chemical database for fragment co-occurrence: discovery of "chemical cliches". *J. Chem. Inf. Model.* **2006**, *46*, 553-562.
- [57] Warner, S.L.; Bashyam, S.; Vankayalapati, H.; Bearss, D.J.; Han, H.; Von Hoff, D.D.; Hurley, L.H. Identification of a lead small-molecule inhibitor of the Aurora kinases using a structure-assisted, fragment-based approach. *Mol. Cancer Ther.* **2006**, *5*(7), 1764-73
- [58] Rummey, C.; Nordhoff, S.; Thiemann, M.; Metz, G. In silico fragment-based discovery of DPP-IV S1 pocket binders. *Bioorg. Med. Chem. Lett.* **2006**, *16*(5), 1405-9.
- [59] Johnson, S.L.; Pellicchia, M. Structure- and fragment-based approaches to protease inhibition. *Curr. Top Med. Chem.* **2006**, *6*, 317-29.
- [60] Merckle, L.; de Andres-Gomez, A.; Dick, B.; Cox, R.J.; Godfrey, C.R. A fragment-based approach to understanding inhibition of 1-deoxy-D-xylulose-5-phosphate reductoisomerase. *ChemBiochem* **2005**, *6*, 1866-1874.
- [61] Erlanson, D.A. Fragment-based lead discovery: a chemical update. *Curr. Opin. Biotechnol.* **2006**, *17*, 643-652.

---

Received: November 30, 2006

Accepted: January 12, 2006

Not For Distribution