BREED: Generating Novel Inhibitors through Hybridization of Known Ligands. Application to CDK2, P38, and HIV Protease

Albert C. Pierce,* Govinda Rao, and Guy W. Bemis

Vertex Pharmaceuticals, 130 Waverly Street, Cambridge, Massachusetts 02139

Received October 24, 2003

In this work we describe BREED, a method for the generation of novel inhibitors from structures of known ligands bound to a common target. The method is essentially an automation of the common medicinal chemistry practice of joining fragments of two known ligands to generate a new inhibitor. The ligand-bound target structures are overlaid, all overlapping bonds in all pairs of ligands are found, and the fragments on each side of each matching bond are swapped to generate the new molecules. Since the method is automated, it can be applied recursively to generate all possible combinations of known ligands. In an application of this method to HIV protease inhibitors and protein kinase inhibitors, hundreds of new molecular structures were generated. These included known inhibitor scaffolds not included in the initial set, entirely novel scaffolds, and novel substituents on known scaffolds. The method is fast, and since all of the ligand functional groups are known to bind the target in the precise position and orientation present in the novel ligand, the success rate of this method should be superior to more traditional de novo design techniques. In an era of increasingly high-throughput structural biology, such methods for high-throughput utilization of structural information will become increasingly valuable.

Introduction

As the determination of protein−inhibitor complex structures has become more routine, the amount of structural information useful in inhibitor design has grown tremendously. Unfortunately, as the supply of this information increases, it becomes increasingly difficult to take full advantage of it. It is simple enough to draw the important lessons from a single complex structure. Extracting all such information from a pair of complex structures is substantially more difficult, and simultaneous comparison of more than three structures is virtually impossible. Yet it is not uncommon to have dozens of protein−ligand complex structures available for comparison. Also, as projects have moved toward gene family-based ligand design, this number has reached the hundreds and continues to grow. Clearly there is a need for methods to efficiently handle this tremendous quantity of information, distilling it into a form that can be readily used in the design of novel ligands.

One simple method for taking advantage of structural information is to use the known positions of two ligands to recombine fragments from each to generate a novel ligand, as in Figure 1. This new molecule will generally be a hybrid of the two scaffolds or a transfer of a substituent from one scaffold to the other. The latter practice is a common design strategy in medicinal chemistry projects, where it is hoped that a potency-enhancing group from one scaffold should have the same effect in a new scaffold if positioned correctly. The difficulty with this method is that it requires manual inspection and recombination of each fragment in each pair of ligands, a prospect that becomes tedious and error-prone with more than a few structures.

BREED was developed to automate this process, using a bond-matching and fragment-swapping algorithm similar to that developed by Ho and Marshall[1] to ensure that all structurally reasonable pairings of ligand fragments would be generated. Also, the automation of this process has allowed it to be taken a step further. The 'offspring' of the original ligands can be added into the pool of initial compounds for recombination into a new set of ligands, as in Figure 2. In this manner, a small number of initial structures can be used to create a very large set of novel inhibitors. These inhibitors do not simply combine the scaffold of one known inhibitor with the side chain of another. Nor are they limited to hybrids of two different scaffolds. With only two iterations of BREED crossing, many of the compounds generated bear very little resemblance to any of the initial ligand structures, combining scaffold and side chain elements from as many as four of the lead compounds. In this capacity, the method begins to resemble de novo design approaches, though obviously it can only be applied in cases where several ligands are already known and structurally characterized. Though this is a disadvantage, in the not infrequent circumstances where initial leads have proven unsuitable (due to insolubility, poor cell permeability, metabolic instability, toxicity, etc.) such late stage de novo design would still be valuable. Another situation in which inhibitor design would be valuable even after large amounts of structural data are available is in the case of gene family-based inhibitor design, where structures from prior projects could be used to design new inhibitors. It has even been suggested recently that this sort of inhibitor design by ligand combination might be applied to ligands of unrelated targets with topologically similar binding pockets.[2]

Two target systems were used in the development and validation of BREED. The first system was HIV-1

* To whom correspondence should be addressed. Phone: 617-444-6825. Fax: 617-444-6566. E-mail: al_pierce@vrtx.com.

10.1021/jm030543u CCC: $27.50
Published on Web 04/20/2004

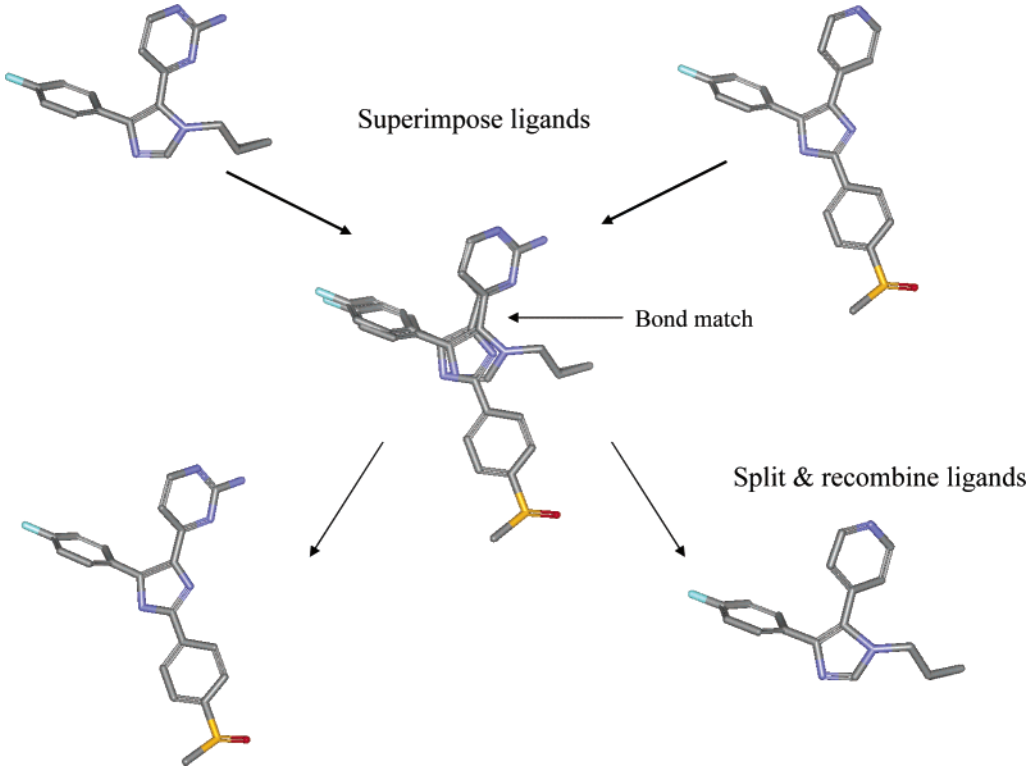


Figure 1. Illustration of the steps carried out by BREED to generate new ligands.

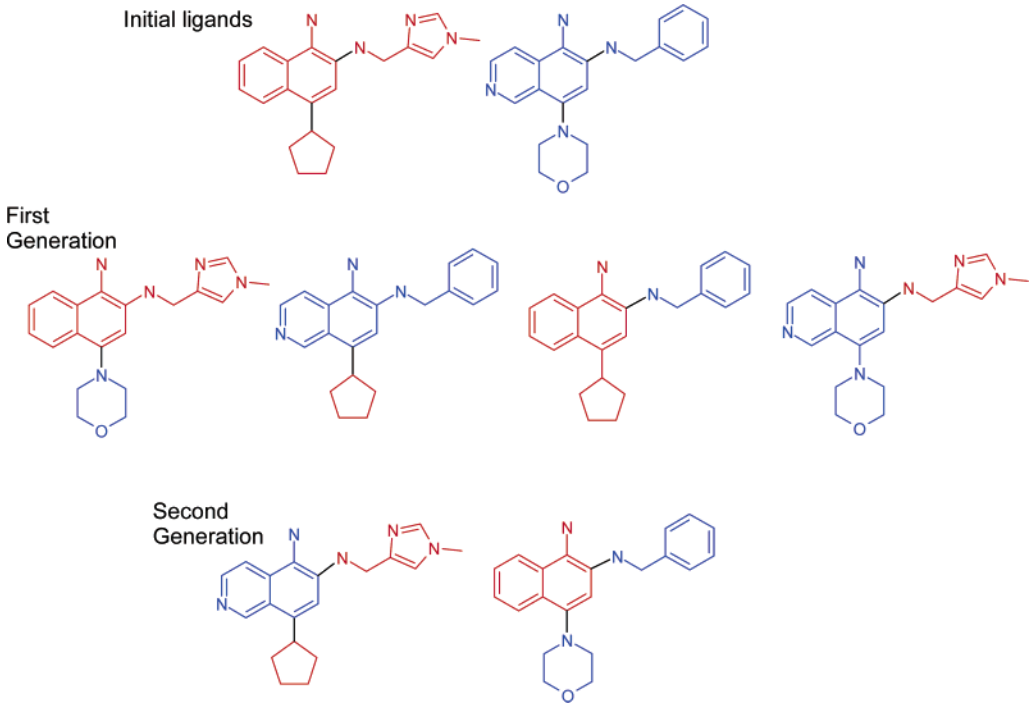


Figure 2. Illustration of the recursive application of BREED processing. In the initial ligands, overlapping bonds are colored black. In the first and second generation molecules, black bonds indicate bonds about which fragments have been swapped to create the new compounds.

aspartyl protease, an enzyme vital to the replication of the AIDS virus. This target was chosen as a single protein system due to its large number and wide variety of potent inhibitors and its numerous publicly available crystal structures. The family of protein kinase enzymes was also studied to establish the applicability of the method to a number of closely related targets. Numerous protein kinase crystal structures have been published and many inhibitors are known, due to these enzymes' roles in cancer, inflammation, diabetes, and other conditions.

Methods

The first step necessary for the application of BREED is the overlay of the target structures. In the case of HIV-1 protease, all backbone atoms were used to overlay the protein structures on to PDB entry 1HSG.[3] For the kinase structures, the backbone atoms of the hinge

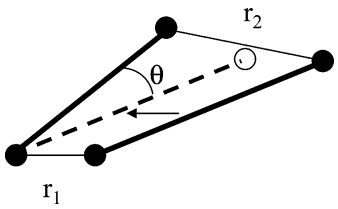


Figure 3. Illustration of the bond matching criteria. Dark circles represent atoms, thick lines represent bonds. Thin lines show r_1 and r_2, the distances between the atoms at each end of the bond. The dashed line is the right-hand bond translated left to show the angle between the two bonds, θ.

region were overlaid with the corresponding atoms of a structure (Vertex, unpublished) of c-Jun N-terminal kinase 3 (JNK3). The hinge residues are 145−149 in JNK3, and the corresponding residues in other kinase structures were determined from amino acid sequence alignments with the kinase domain of JNK3. These hinge residues were chosen for alignment based on their structural similarity among protein kinases and the fact that alignment of these residues gives excellent alignment of inhibitors in the ATP binding pocket. Where compounds are shown overlaid in Figures 5, 8, and 9, the proteins were overlaid based on the same atom sets used in this initial superposition.

Once the set of structures is aligned and the ligands are in a common reference frame, all of the ligands are saved to an SDfile[4] without explicit hydrogen atoms. The SDfile is then processed in the following manner. Each pair of ligands is considered in turn to find all matching bonds between the two molecules. A pair of bonds is considered to be matching if four conditions are met. First, the two bonds must be of the same order (i.e. a single bond and a double bond cannot be considered a match.) This requirement is necessary to maintain the hybridization/geometry of the bonded atoms in the new molecule. Second, the atoms at each end of the bond must be within 1 Å of each other. Third, the angle between the bond vectors of the two bonds must be no greater than 15°. The final requirement is that neither of the bonds is in a ring system. While ring fusion is an additional reasonable strategy for ligand hybridization, it is a considerably more complex problem, requiring consideration of additional geometrical requirements as well as issues of atom type, valence, and aromaticity. As such, ring-handling will be addressed in future work.

The geometric requirements described above are depicted in Figure 3, and although the exact distance and angle are somewhat arbitrary, we have found these values to yield a reasonable number of bond matches without leading to excessively distorted geometries in the recombined molecules. Although these margins of error might seem large relative to tolerances in van der Waals contacts, they are reasonable relative to limits of crystal structure resolution, protein flexibility, and the accuracy of protein alignments. It might actually be preferable, particularly in cases where the target proteins used in alignment are different, to use looser tolerances for bond matching.

For each pair of matching bonds among two molecules, two new molecules are generated, as depicted in Figure 1. If the initial molecules are considered to be split into two 'halves' at the matching bond, one new molecule consists of the first half of molecule one and the second half of molecule two. The other new molecule is made up of the second half of molecule one and the first half of molecule two. All of the atoms in the new molecules have exactly the same atom types, positions, and bonds as the corresponding atoms in the parent compounds, except for the two atoms that made up the matching bond. These two atoms are identical except for their Cartesian coordinates, which are the average of the Cartesian coordinates of the corresponding atom in both parent molecules. This averaging gives the new bond between the two halves of the molecule a more appropriate bond length and a bond vector closer to the bond vectors of the parent compounds. As each new molecule is generated, its canonical SMILES[5] string is compared to a list of the canonical SMILES strings of all the input and previously generated molecules. If it is not a duplicate, the molecule is output into an SD file of results.

The method described above represents the core of BREED. The molecules in the new SD file can be viewed in many packages of molecular visualization software, minimized, scored, or otherwise processed as potential lead compounds. However, this SD file is also appropriate for reprocessing by BREED. If the SD file of initial compounds is concatenated with the file of output compounds, this new file can be processed to generate further molecules combining input with offspring, and offspring with other offspring.

For both the HIV-1 protease inhibitors and protein kinase inhibitors, four compounds were initially selected for BREED processing. These sets were chosen to ensure that each compound shared at least one matching bond with another molecule in the set. The small sets allowed manual inspection of the results to verify that they are both correct (that no errors were made in the splitting or recombining of molecules) and complete (that recombination was executed at all matching bonds). With the method refined, six additional inhibitors were added to both the HIV-1 protease and protein kinase sets. These new compounds were chosen to give a diverse set of 10 inhibitors for each system, to verify that BREED can generate new inhibitors from a larger set of arbitrarily selected ligands. The initial four ligands for the HIV-1 protease system come from the PDB crystal structures 1HPV,[6] 1HSG,[3] 1HPX,[7] and 1HXB[8] and are shown in Figure 4. The remaining six ligands come from the PDB crystal structures 1B6J,[9] 1B6K,[9] 1HII,[10] 1IIQ,[11] 1OHR,[12] and 4PHV.[13] The first four kinase ligands (compounds **14**−**17**, Figure 7) come from the p38 MAP kinase (p38) and cyclin-dependent kinase 2 (CDK2) crystal structures with PDB codes 1A9U,[14] 1BMK,[14] 1DI9,[15] and 1JSV.[16] The remainder of the kinase ligands came from Vertex-determined crystal structures of inhibitors bound to glycogen synthase kinase 3 (GSK3), Src kinase, and Aurora 2 kinase.

Results and Discussion

HIV-1 Protease. A first pass of the four original HIV-1 protease inhibitors (compounds **1**−**4**, Figure 4) through BREED generated 20 novel compounds, and a second round of processing led to an additional 81 compounds, for a total of 101 novel potential inhibitors. Elimination of structures with undesirable characteristics (no key hydrogen bonding hydroxyl group or chemically unstable functionality) yields a list of 62

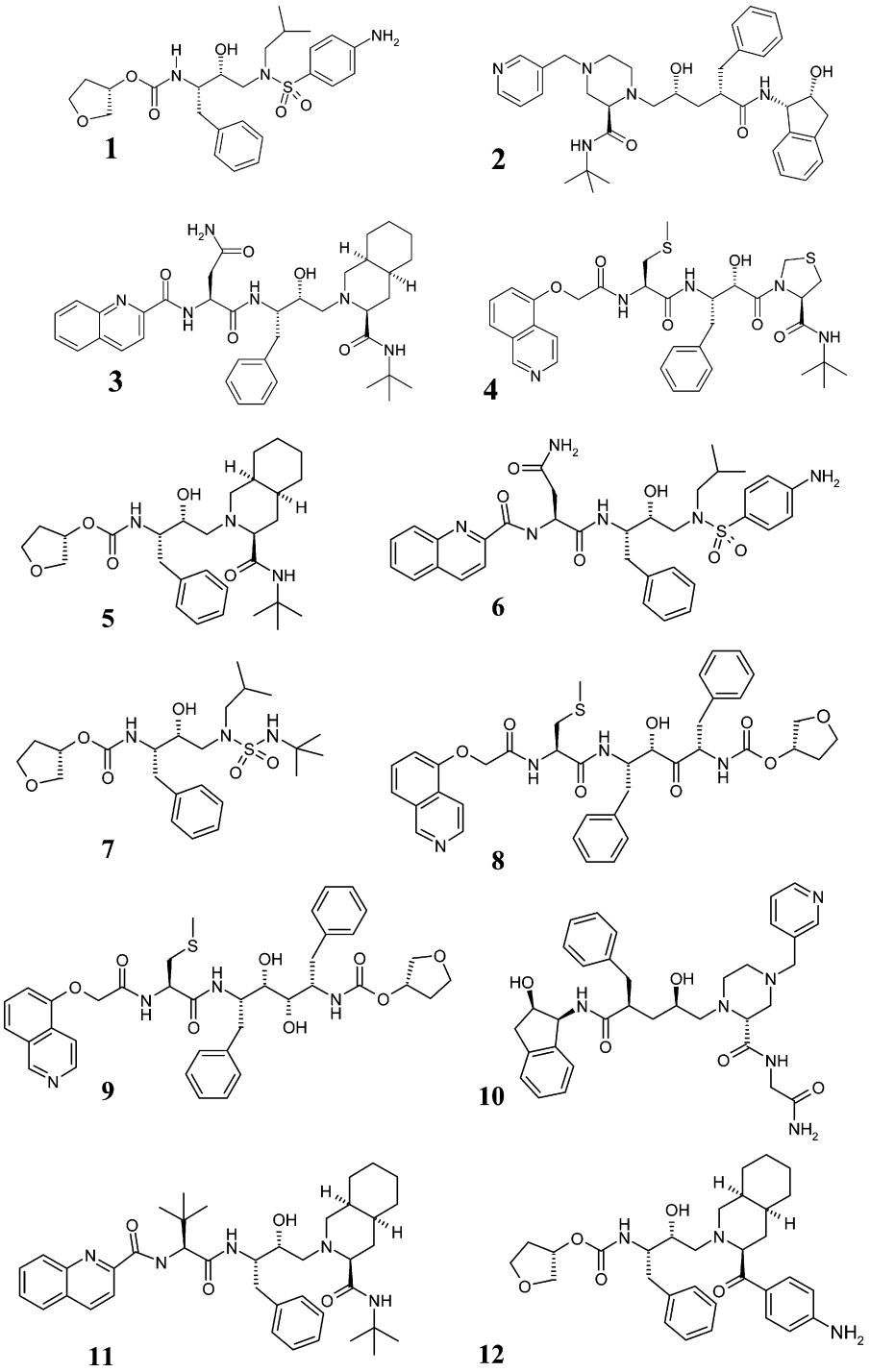


Figure 4. Four HIV-1 protease inhibitors used as input to BREED processing (molecules **1**−**4**) and a sampling of eight molecules generated by BREED processing (molecules **5**−**12**).

novel, chemically viable compounds. Eight of these structures (compounds **5**−**12**) are shown in Figure 4. Compounds **5**−**11** were generated in the first round of BREED processing, while compound **12** is from the second round. Compounds **5** and **6** are hybrids of structures **1** and **3** from Figure 4a. Compound **5** has been synthesized as an HIV-1 protease inhibitor and is reported to have an enzymatic IC50 of 160 nM.[17] A very close analogue of **6**, with the aniline group replaced by benzoxadiazole, was synthesized, and a crystal structure has been determined (K_i = 0.1 nM, unpublished data). Because these compounds are designed by combining target-bound inhibitors, the new structures are created in the appropriate conformation and position for binding to their target. In essence, they are pre-docked. Simply superimposing the protein of the crystal structure of the benzoxadiazole version of compound **6** onto the protein in 1HSG shows (Figure 5) that BREED has closely reproduced the binding mode of inhibitor **6**, giving an rms deviation from the X-ray structure of only 0.8 Å among the 46 atoms shared between the two structures.

Output structure **7** is a variation of another compound designed and synthesized in our efforts to develop a second-generation HIV protease inhibitor. This compound had *N*,*N*-dimethyl rather *tert*-butyl sulfamide and inhibited the enzyme with a K_i of 42 nM, validating the

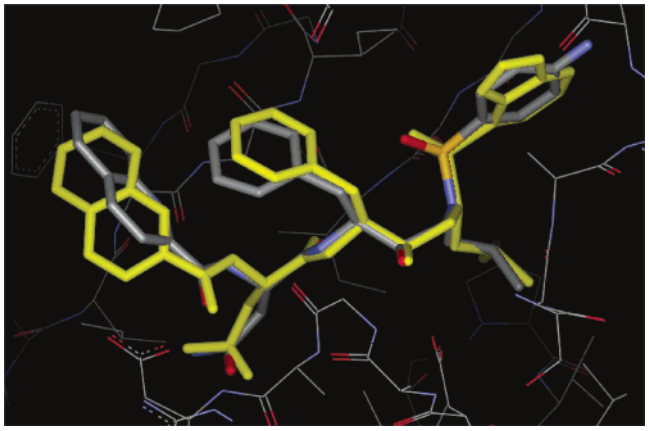
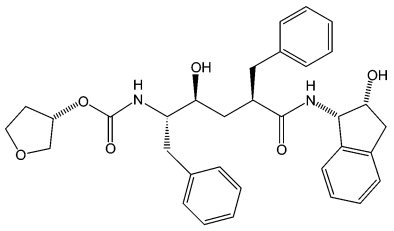


Figure 5. Compound **6** generated by BREED processing and a close analogue (in yellow) for which a crystal structure is available. The molecules are overlaid by overlaying the protein structures, yet the RMS deviation among shared ligand atoms is only 0.8 Å.



13

Figure 6. A hybrid of HIV-1 ligands **1** and **2** that was not generated by BREED due to relatively tight tolerances for bond matching.

sulfonamide to sulfamide transformation suggested by BREED. Another interesting result among these hybrid compounds is the generation of new backbone functionality for the critical central hydroxyl group. Structures **8** and **9** show the generation of two backbone hydroxyl motifs not present in the input structures, the hydroxy-ketone and dihydroxy moieties. The dihydroxy motif is well-known among HIV-1 protease inhibitors,[18] and while the hydroxy-ketone is a novel functionality for inhibitors of this protease, it has been successfully incorporated into inhibitors of the aspartyl protease renin.[19] Compounds **10** and **11** have the backbones of parent structures **2** and **3**, respectively, but have exchanged an Asp side chain and a *tert*-butyl group between the two backbones. This exchange of side chains suggests that substituents and SAR may be transferable between these two sites, a valuable finding for a pair of scaffolds. A final class of novel results is exemplified by compound **12**. In this case, the novel structure combines both ends of input structure **1** with a central element of structure **3**. Such results are only possible in the second and subsequent generations of BREED results. They are formed by an initial combination of hypothetical structures A and B, followed by a combination of this first hybrid with either A or B, as shown in Figure 2. This result is reminiscent of the program CAVEAT[20] which was developed to find such linkers, though BREED provides only linkers known to bind the target of interest.

It should be noted that not all possible known hybrids of the input molecules are generated by BREED. Compound **13** in Figure 6 is a potential hybrid of input structures **1** and **2** and is known to be a potent inhibitor of HIV-1 protease.[21] This molecule is not generated due to the tight geometric tolerances used in determining a bond match and the fact that **1** and **2** do not appear to be particularly appropriate for hybridization based on visual inspection. The two molecules have opposite stereochemistry at the central hydroxyl, which causes the angle between their "matching" bond to be roughly 50°. Loosening the tolerances sufficiently to generate **13** would generate hundreds of compounds unlikely to bind the target of interest. While it is possible that an energy minimization and scoring protocol could adequately filter these results, this begins to undermine the advantages of BREED, and more difficult hybridizations such as this may simply be best left to manual modeling.

When BREED was applied to a set of 10 HIV-1 protease inhibitors, 75 new structures were generated in the first round of crossing, and 716 compounds were generated in the second round. Among these 791 compounds, 767 contain the Asp-bridging hydroxyl. These results suggest that BREED is broadly applicable for the design of HIV-1 protease inhibitors.

Protein Kinases. An obvious drawback of this method is that it requires a significant amount of structural information to be implemented. In the early phases of a drug discovery program, when BREED results would be most valuable, such information is often in short supply. However, in cases where structures are available for related targets, nonselective inhibitors from these targets become excellent starting points for BREED processing in the new program. Of course, the ideal situation for such cross-target inhibitor hybridization is within a gene family-based drug design program. In such an environment, large amounts of structural information would be available even before the program for a particular target is initiated, and BREED could have a tremendous impact.

Protein kinases, the enzymes responsible for transfer of the gamma phosphate of ATP to the hydroxyl side chains of their substrate proteins, have been the subject of considerable interest in gene family-based design and were chosen as a second test case for the BREED method. The first four inhibitors processed from this class (**14**−**17**) are shown in Figure 7, along with the eight compounds produced in the first pass of BREED processing (**18**−**25**). A second pass of these compounds through BREED gave four more output structures. One particularly interesting output molecule is **24**, a combination of **15** and **16** with an IC_{50} of 160 nM against p38.[14] As there is even a crystal structure available for this compound (PDB code 1BL6[14]), a comparison can be made between the BREED "docked" structure and the experimentally determined structure. These two structures are shown in Figure 8, where the superposition has been performed by overlaying the protein of 1BL6 with the JNK3 structure used for BREED processing of all kinase structures. The rms deviation in atom position between the two molecules is 1.07 Å, remarkably small given the method of superposition. It seems extremely likely that **25** is also a potent inhibitor of p38, given that **15**, **16**, and **24** are all p38 inhibitors, and visual inspection suggests that all of the same key interactions will be made by **25**. Although there is no

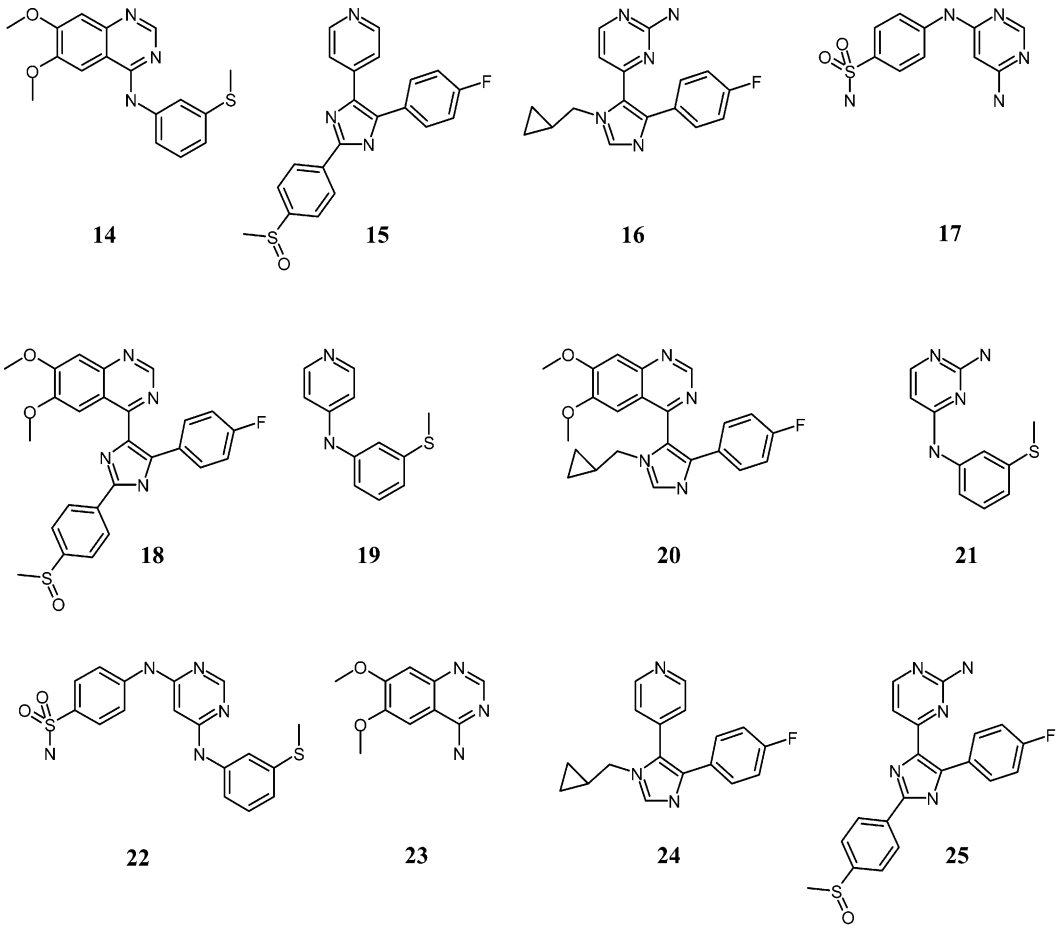


Figure 7. Four kinase inhibitors used as input to BREED (**14**−**17**) and eight compounds (**18**−**25**) generated by BREED processing.

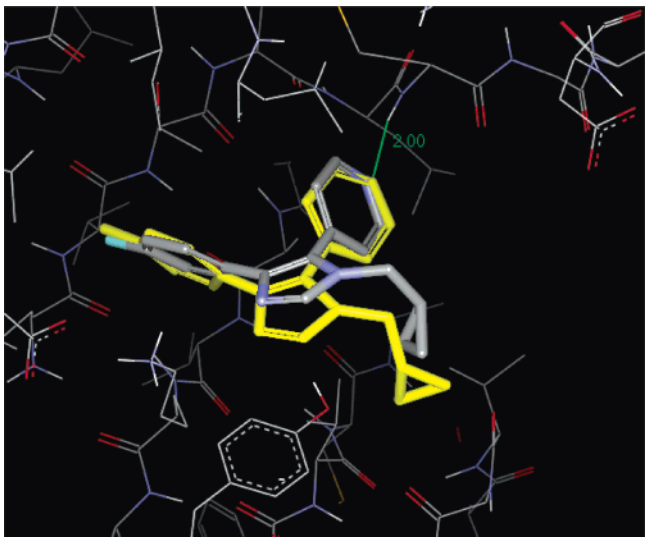


Figure 8. Compound **24** "docked" by BREED (in yellow) overlaid with the published crystal structure. Though only protein atoms were used in superimposing the ligands, the RMS deviation in atom positions is 1.07 Å.

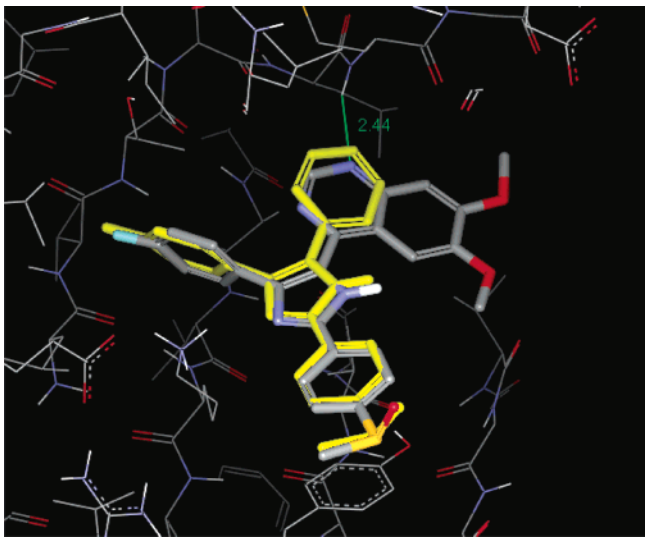


Figure 9. Compound **18** overlaid with crystal structure of input structure **15** (in yellow).

published data for **25** itself as a kinase inhibitor, it does fall within the generic claim of a patent of p38 inhibitors.[22] Several of the other compounds in Figure 7 have similarly compelling evidence for their status as kinase inhibitors. All of them are composites of known kinase inhibitors, with hydrogen bonding functionality at the ATP site hinge, reasonable binding conformations, and hydrophobic functionality that fits well within the ATP site. Additionally, there are numerous patents covering 2-amino-4-anilinopyrimidines similar to **21** as inhibitors of the kinases FGFR-1, CDK's, GSK3β, FAK, and kinases in general.[23−26] Also, there are patents covering 4-amino-6-anilinopyrimidines such as **22** as inhibitors of the CDK's, p38, and tyrosine kinases.[27−29] 4-Aminoquinazolines such as **23** are also known to be inhibitors of kinases as exemplified by **14**, so this compound would also be expected to inhibit kinases, though its small size makes it unlikely to be very potent. The most exciting

results, however, are compounds such as **18**, **19**, and **20**. These molecules represent entirely novel scaffolds with potential for development as kinase inhibitors.

To establish that such results are typical of BREED processing of kinase inhibitors in general, and not just a product of the particular four compounds **14**−**17**, processing was repeated on all 10 kinase inhibitors described in the Methods section. Two rounds of processing yielded 119 compounds, a significant number, but far fewer than the 791 produced from the application of BREED to the 10 HIV-1 protease compounds. This is because fragment swapping to make new molecules only takes place at acyclic bonds, and the ring-rich kinase compounds have significantly fewer such bonds than the protease inhibitors. On average, kinase inhibitors have six acyclic bonds per molecule while the HIV-1 protease inhibitors have 20. This, of course, implies that some classes of inhibitors will be much more prolific than others in BREED processing. Macrocycles and steroids, for example, have few acyclic bonds and would therefore generate few new ideas without BREED modifications to handle these special cases.

Regardless, the relatively small number of molecules generated from the kinase leads does not imply that BREED could not generate a large number of kinase inhibitor ideas. Several factors allow for a potential abundance of new ideas for this system. Due to the large number of publicly (and perhaps privately) available ligand-bound kinase structures, the recombination of these inhibitors will generate a very large number of novel compounds. This is an excellent example of using chemogenomics to leverage knowledge from past programs to facilitate research against new targets. Another option for increasing the number of starting structures for BREED processing is to generate them by docking known inhibitors into the appropriate kinase. Combination of docked structures with the experimentally determined structures would greatly expand the number of ideas generated by this method, though at some risk of generating "false positives" based on improperly docked structures. A final possibility for increasing the output of BREED would be the addition of explicit hydrogen atoms to the initial structures before processing. This would significantly increase the number of acyclic bonds available for matching and consequently increase the yield of novel hybrids.

The rapid generation of such large numbers of novel compounds for potential synthesis and testing, of course, presents its own problems. Namely, how does one decide which compounds to make given limited time and resources. Fortunately, this is a topic that has been covered extensively elsewhere. Scoring functions and other binding affinity prediction tools[30] can be used to filter the compounds based on likely in vitro activity versus the target of interest. Alternately, if novelty is a priority, the new compounds with the lowest molecular similarity to any known inhibitors might be given the highest priority. Other tools have been developed to predict the druglikeness of compounds,[31] their solubility,[32] intestinal absorption,[33] metabolic stability, toxicity, etc.[34] Any of these tools could also be used to rank compounds based on a given set of properties appropriate for the ultimate goal of the design program. Assuming that the methods described above narrow the list to a manageable length, the remaining compounds can be visually inspected for synthetic accessibility before final selections are made.

At this point it is worth considering how BREED compares to other ligand design programs. BREED holds an advantage in speed and (presumably) success rate. Both of these advantages come from using experimentally determined structures of known ligands bound to the target of interest. A typical structure-based de novo ligand design program spends the majority of its effort sampling conformational space to place ligands or fragments into the protein active site and then scoring their "fitness". BREED, by using experimentally determined structures, does not require conformational searches or fitness evaluation. As a result, while other programs can take a day or more to generate a set of tens to hundreds of ideas,[35,36] BREED takes approximately 50 min to generate the 789 potential inhibitors of HIV-1 protease on a 900 MHz Pentium 3. One reviewer accurately notes that setup and user visualization of results can be the most time-consuming aspects of computational ligand design. To this point, BREED's use of precise ligand-placement information from crystal structures should also allow a speedup in visualization due to the generation of fewer false positives. Since traditional de novo design depends on imperfect scoring functions for ligand placement and fitness calculation, a considerable number of inactive compounds are generated, adding to the burden of user visualization. The main disadvantage of this method is that it requires a reasonable number of target−ligand complex structures to implement, though with the rise of high-throughput structural biology and gene family-based drug discovery, this requirement will be more routinely met.

The other aspect in which BREED might be compared to other methods is as a means of determining positions on different scaffolds at which substituents may be interchanged. For this process, the state-of-the-art is fairly primitive. The only technique available to our knowledge is manual inspection by a chemist or molecular modeler. Although this is not particularly difficult or time-consuming for a few structures, the exponential increase in comparisons with the number of starting structures renders this method slow and tedious with a handful of structures and infeasible with more. BREED makes all of the comparisons and recombinations quickly and flawlessly.

Conclusions

The application of structure-based drug design has become increasingly successful in recent years as structures become more common and the techniques more refined. With the advent of high-throughput structural biology, it is expected that this trend will continue and perhaps accelerate. However, the tools of structure-based drug design have generally not evolved to take advantage of the larger quantities of available structural information. BREED is a first computational implementation of the sort of design methods that have traditionally been performed manually when fewer structures were available. This automation of the common practice of swapping fragments between different ligands of a given target has been shown to reproduce known inhibitors of HIV-1 protease and protein kinases. The

hybridizations and backbone modifications generated, as originally reported in the literature, were the result of considerable design efforts by highly skilled medicinal and computational chemists. We believe it is a significant breakthrough to be able to generate such ideas quickly and routinely through application of the BREED method. The compounds produced by these ligand recombinations include examples of both novel scaffolds and substituent exchanges between scaffolds. In the former examples, the method represents a de novo-like method for ligand design, with significant enhancements in speed and incorporation of known structural information. The latter examples are representative of the results of manual side chain swapping, with improvements in speed and thoroughness. The automation of the method also allows it to be performed recursively, such that the molecular structures produced by BREED can be passed through the process again, recombining with each other to generate still more novel compounds. In this way it is even possible to generate CAVEAT-like results in which a novel linker is introduced between two important binding elements. BREED's efficient use of large amounts of structural information will make it particularly useful in the current era of high-throughput structural biology and gene family-based drug design.

References

(1) Ho, C. M. W.; Marshall, G. R. SPLICE: A program to assemble partial query solutions from three-dimensional database searches into novel ligands. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 623−647.

(2) Schmitt, S.; Kuhn, D.; Klebe, G. A New Method to Detect Related Function Among Proteins Independent of Sequence and Fold Homology. *J. Mol. Biol.* **2002**, *323*, 387−406.

(3) Chen, Z.; Li, Y.; Chen, E.; Hall, D. L.; Darke, P. L. et al. Crystal structure at 1.9-A resolution of human immunodeficiency virus (HIV) II protease complexed with L-735,524, an orally bioavailable inhibitor of the HIV proteases. *J. Biol. Chem.* **1994**, *269*, 26344−26348.

(4) Dalby, A.; Nourse, J. G.; Hounshell, W. D.; Gushurst, A. K. I.; Grier, D. L. et al. Description of Several Chemical Structure File Formats Used by Computer Programs Developed at Molecular Design Limited. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 244−255.

(5) Weininger, D. SMILES 1. Introduction and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31−36.

(6) Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G. et al. Crystal-strucure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. *J. Am. Chem. Soc.* **1995**, *117*, 1181−1182.

(7) Vazquez, M. L.; Bryant, M. L.; Clare, M.; DeCrescenzo, G. A.; Doherty, E. M. et al. Inhibitors of HIV-1 Protease Containing the Novel and Potent (*R*)-Hydroxyethyl)sulfonamide Isostere. *J. Med. Chem.* **1995**, *38*, 581−584.

(8) Krohn, A.; Redshaw, S.; Ritchie, J. C.; Graves, B. J.; Hatada, M. H. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (*R*)-hydroxyethylamine isostere. *J. Med. Chem.* **1991**, *34*, 3340−3342.

(9) Martin, J. L.; Begun, J.; Schindeler, A.; Wickramasinghe, W. A.; Alewood, D. et al. Molecular recognition of macrocyclic peptidomimetic inhibitors by HIV-1 protease. *Biochemistry* **1999**, *38*, 7978−7988.

(10) Priestle, J. P.; Fassler, A.; Rosel, J.; Tintelnot-Blomley, M.; Strop, P. et al. Comparative analysis of the X-ray structures of HIV-1 and HIV-2 proteases in complex with CGP 53820, a novel pseudosymmetric inhibitor. *Structure* **1995**, *3*, 381−389.

(11) Dohnalek, J.; Hasek, J.; Duskova, J.; Petrokova, H.; Hradilek, M. et al. Hydroxyethylamine isostere of an HIV-1 protease inhibitor prefers its amine to the hydroxy group in binding to catalytic aspartates. A synchrotron study of HIV-1 protease in complex with a peptidomimetic inhibitor. *J. Med. Chem.* **2002**, *45*, 1432−1438.

(12) Kaldor, S. W.; Kalish, V. J.; 2nd, J. F. D.; Shetty, B. V.; Fritz, J. E. et al. Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. *J. Med. Chem.* **1997**, *40*, 3979−3985.

(13) Bone, R.; Vacca, J. P.; Anderson, P. S.; Holloway, M. K. X-ray crystal structure of the HIV protease complex with L-700,417, an inhibitor with pseudo C2 symmetry. *J. Am. Chem. Soc.* **1991**, *113*, 9382−9384.

(14) Wang, Z.; Canagarajah, B. J.; Boehm, J. C.; Kassisa, S.; Cobb, M. H. et al. Structural basis of inhibitor selectivity in MAP kinases. *Structure* **1998**, *6*, 1117−1128.

(15) Shewchuk, L.; Hassell, A.; Wisely, B.; Rocque, W.; Holmes, W. et al. Binding mode of the 4-anilinoquinazoline class of protein kinase inhibitor: X-ray crystallographic studies of 4-anilinoquinazolines bound to cyclin-dependent kinase 2 and p38 kinase. *J. Med. Chem.* **2000**, *43*, 133−138.

(16) Clare, P. M.; Poorman, R. A.; Kelley, L. C.; Watenpaugh, K. D.; Bannow, C. A. et al. The cyclin-dependent kinases cdk2 and cdk5 act by a random, anticooperative kinetic mechanism. *J. Biol. Chem.* **2001**, *276*, 48292−48299.

(17) Ghosh, A. K.; Thompson, W. J.; Lee, H. Y.; McKee, S. P.; Munson, P. M. et al. Cyclic Sulfolanes as Novel and High Affinity P2 Ligands for HIV-1 Protease Inhibitors. *J. Med. Chem.* **1993**, *36*, 924−927.

(18) Kempf, D. J.; Norbeck, D. W.; Codacovi, L.; Wang, X. C.; Kohlbrenner, W. E. et al. Structure-Based, C2 Symmetric Inhibitors of HIV Protease. *J. Med. Chem.* **1990**, *33*, 2687−2689.

(19) Patel, D. V.; Rielly-Gauvin, K.; Ryono, D. E.; Free, C. A.; Smith, S. A. et al. Activated Ketone Based Inhibitors of Human Renin. *J. Med. Chem.* **1993**, *36*, 2431−2447.

(20) Lauri, G.; Bartlett, P. A. CAVEAT: a program to facilitate the design of organic molecules. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 51−66.

(21) Ghosh, A. K.; Thompson, W. J.; McKee, S. P.; Duong, T. T.; Lyle, T. A. et al. 3-Tetrahydrofuran and Pyran Urethanes as High-Affinity P2−Ligands for HIV-1 Protease Inhibitors. *J. Med. Chem.* **1993**, *36*, 292−294.

(22) Adams, J. L.; Lee, D.; Long, S. A. Novel aryloxy substituted pyrimidine imidazole compounds. WO9857966, 1998.

(23) Pease, E. J.; Breault, G. A.; Morris, J. J. Preparation of pyrimidin-2-amines as cyclin-dependent serine/threonine kinase (CDK) inhibitors. WO0164654, 2001.

(24) Pease, E. J.; Breault, G. A.; Williams, E. J.; Bradbury, R. H.; Morris, J. J. Preparation of 2,4-di(hetero)arylamino(oxy)-5-substituted pyrimidines as antineoplastic agents. WO0164655, 2001.

(25) Nuss, J. M.; Harrison, S. D.; Ring, D. B.; Boyce, R. S.; Johnson, K. et al. Preparation of aminopyrimidines and -pyridines as glycogen synthase kinase 3 inhibitors. WO0220495, 2002.

(26) Armistead, D. M.; Bemis, J. E.; Pietro, L. V. D.; Geuns-Meyer, S. D.; Habgood, G. J. et al. Preparation of arylaminopyrimidines as Kinase inhibitors. WO0160816, 2001.

(27) Thomas, A. P. Preparation of 4,6-dianilinopyrimidines derivatives as tyrosine kinase inhibitors. WO9515952, 1995.

(28) Breault, G. A.; Jewsbury, P. J.; Pease, J. E. Preparation of bis-(anilino)pyrimidines as CDK inhibitors. WO0012486, 2000.

(29) Cumming, J. G. Preparation of 4-anilinopyrimidines as p38 kinase inhibitors. WO0127089, 2001.

(30) Stahl, M.; Rarey, M. Detailed Analysis of Scoring Functions for Virtual Screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.

(31) Walters, W. P.; Murcko, M. A. Prediction of 'drug-likeness'. *Adv. Drug Deliv. Rev.* **2002**, *54*, 255−271.

(32) Jorgensen, W. L.; Duffy, E. M. Prediction of drug solubility from structure. *Adv. Drug Deliv. Rev.* **2002**, *54*, 355−366.

(33) Egan, W. J.; Lauri, G. Prediction of intestinal permeability. *Adv. Drug Deliv. Rev.* **2002**, *54*, 273−289.

(34) Ekins, S.; Boulanger, B.; Swaan, P. W.; Hupcey, M. A. Towards a new age of virtual ADME/TOX and multidimensional drug discovery. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 381−401.

(35) Pegg, S. C.-H.; Haresco, J. J.; Kuntz, I. D. A genetic algorithm for structure-based de novo design. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 911−933.

(36) Pearlman, D. A.; Murcko, M. A. CONCERTS: Dynamic Connection of Fragments as an Approach to *de Novo* Ligand Design. *J. Med. Chem.* **1996**, *39*, 1651−1663.

JM030543U