Scaffold Replacement in MOE

Simon Grimshaw

Chemical Computing Group Inc.
1010 Sherbrooke Street West, Suite 910 Montreal, Canada H3A 2R7

Abstract. A new application for performing scaffold replacements, fragment linking, and R-group optimization is presented. The application is applied to the discovery of novel p38 MAP kinase inhibitors, using the structures of a screening hit and the mature BIRB-796 inhibitor as starting points.

INTRODUCTION

The progression of an initial hit to a lead candidate is accompanied by a variety of molecular alterations. For example, a portion of the molecule could be replaced, or a group might be added to achieve a particular polar or steric interaction. In fragment-based drug discovery, fragments need to be extended or linked with a scaffold. Computational methods to accelerate these processes are highly desirable.

The application presented in this article builds on MOE's Linker pharmacophore features, providing a simple-to-use interface that automatically creates complex scaffold replacement queries. Resulting molecules can be refined and scored against any available receptor atoms, and also filtered using molecular properties, pharmacophore, QSAR or fingerprint similarity models. Additionally the synthetic plausibility of the proposed molecules is predicted and can be included as a filter. Examples are shown of the use of the application to optimise an initial hit to the p38 kinase and to replace the scaffold of a known p38 ligand.

METHODS

Computational methods for scaffold replacement include CAVEAT, Recore and MOE's Linker pharmacophores; these have been described in a previous article [Deschênes 2007]. MOE's Linker pharmacophore features encode CAVEAT-like bond vectors as part of a pharmacophore query. A simple, "Add group to ligand" type of modification requires a single pair of such features to define the bond vector. Combining two (or more) pairs of features, each defining the bond from an R-group to the ligand's scaffold, will select fragments able to bond to all the specified R-groups, thus replacing the scaffold of the input molecule.

Similar molecular alterations are required for scaffold replacement, fragment linking and fragment growth. Once the bonds between the scaffold and the R-groups have been selected, the bonds are cut and the endpoints are used to define bond vectors. In the case of fragment growth, one bond vector is typically identified, along with other pharmacophore features in the volume to be filled. For fragment linking and scaffold replacement functions, two or more bond vectors are used. If at least four points are defined (whether from two pairs of bond vector endpoints, or from pharmacophore query features), there is sufficient information to superpose the feature points of candidate compounds onto compatible pharmacophore features. If a sufficiently close superposition can be achieved, the candidate compounds can be bonded to the initial molecule.

Bond vectors are defined using Link pharmacophore annotations at the position of heavy atoms and projected Link2 and Link3 annotations at the position of R-group atoms (Figure 1). Projected annotations are placed at a distance of 1.5Å, roughly equivalent to the length of a carbon-carbon single bond, and at an angle that is consistent with the hybridization of the heavy atom. Link2 projected features are sp^2-geometry, i.e. in the plane of the atoms attached to the scaffold heavy atom, while Link3 features are consistent with the scaffold atom being sp^3-hybridized. During the scaffold search, a constraint is imposed that ensures that both heavy and projected annotations are matched by the same atom, thus encoding the "exit vector" concept of Bartlett's CAVEAT method.

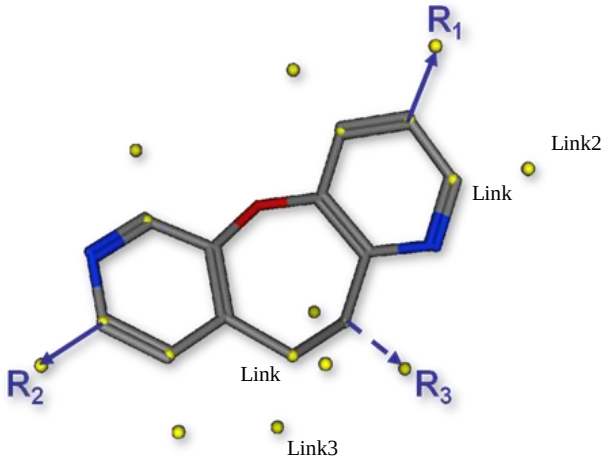


Figure 1. *The combination of a Link feature with Link2 and/or Link3 features defines the bond vectors between scaffold and R-groups.*

Link annotations are placed at the position of atoms with reliable local minima; for example, on alkane carbon atoms they are placed in staggered conformations. Conjugated atoms and terminal free rotors present problems in that there may be no reliable exit vector direction. For example, methyl groups on sp^2 systems are often free rotors; consequently, 16 projected Link features are created around such groups to reflect this uncertainty. Aniline nitrogen atoms are also a problem, in that their planarity depends upon the attached groups. Amines are annotated both for sp^3 geometry and sp^2 geometry to allow for amide bond formation.

Two fragments can be positioned such that a bond could be formed between them by superposing their Link-Link{2,3} features in opposing geometry. For example, given two fragments, the Link feature of the first fragment is overlaid with the Link2 of the second, and the Link feature of the second fragment is overlaid with the Link3 of the first. The Link features are typically given very small radii (e.g., 0.3Å), and so the overlay will produce nearly optimal bond lengths and angles. If there are no other features available to constrain the dihedral of the new bond, a dihedral angle search must be undertaken. Alternatively, if there are such extra features, the dihedral angle formed by the connection must be checked (π-bond dihedrals must be near to 0 or 180 degrees and sp3-sp3 dihedrals should not be eclipsed).

A clash test is performed to ensure that the connected fragments have no van der Waals clashes. In such a test, the locations of hydrogen atoms are taken into account. A similar clash test is performed between the receptor and the proposed scaffold, but not between the receptor and input R-groups, since they are assumed to be reasonable. The stringency of the clash tests is an important parameter. A permissive threshold will lead to many van der Waals clashes that must be resolved with coordinate refinement. A strict threshold will produce much fewer candidates, but refinement of the result molecules may not be required.

An excluded volume is created based on all ligand atoms with a topological distance of two or more bonds from a connection point. This will ensure that linkers will not form clashes with the retained R-groups. Additionally, if a receptor is present, an excluded volume is created from the pocket atoms, to provide a shape guide for candidate linkers.

The workflow followed during the process of computational ligand optimization is shown in Figure 2. The first stage is query generation, i.e. defining what molecular alteration will be performed. This can be run in four different modes:

1. Replace Scaffold (Select Scaffold)
2. Replace Scaffold (Select R Atoms)
3. Link Fragments
4. Add Group to Ligand

These modes automatically create one or more pharmacophore queries and encode a variety of chemical modifications (linking, growing, ring fusion and cyclization).

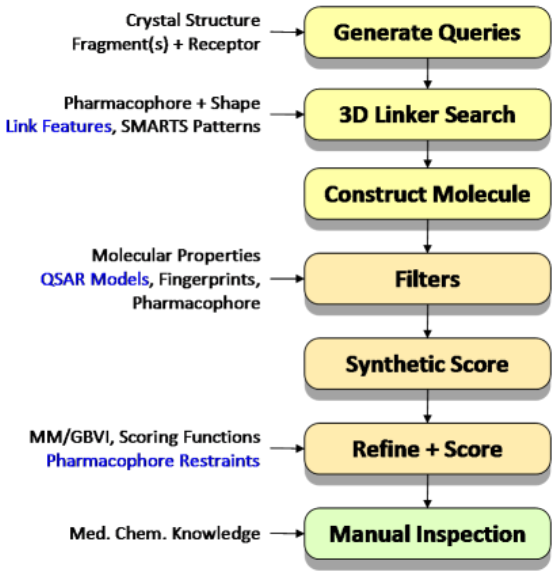


Figure 2. *The workflow used for scaffold replacement and ligand optimisation.*

Replace Scaffold. A portion of the input ligand can be selected for replacement using one of these modes:

1. Select Scaffold, where the selected atoms are replaced.

2. Select R-Group Atoms, where the connecting atom of each retained R-group is selected

Using the first mode, "Select Scaffold," and selecting the atoms indicated in red below (Figure 3) results in two bond vectors, shown as green arrows. The R-groups to be retained in the result molecules are indicated in black.

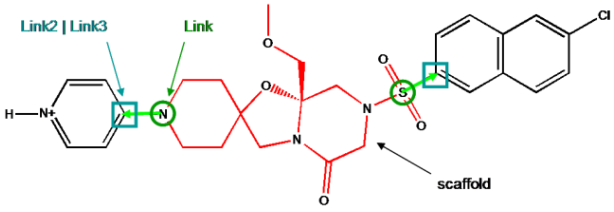


Figure 3*. A scaffold replacement query showing two exit vectors shown as green arrows, with the Link and "Link2 OR Link3" pharmacophore features highlighted by circles and squares respectively.*

A single pharmacophore query is generated that places "Link2 | Link3" (sp2 OR sp3) projected Linker features on the bridgehead R-group atoms (marked with blue squares). Link features are generated 1.5Å away at the correct position for the heavy atom (green circles). Each pair is constrained such that they are formed by the same heavy atom to ensure that a bond can be formed.

Connections can be either essential or optional. Multiple queries are automatically generated that include the essential connections as well as all combinations of the optional connections, thus allowing a broad range of possible molecular modifications to be investigated in a single search. Optional connections are defined by selecting hydrogen atoms, which serve to orient the bond vector. Selecting the connected heavy atom as well as the hydrogen atom marks the connection as

essential. For example, if an optional connection is added to the pyridine ring (Figure 4), two queries are generated: one involving the two essential connections, and one involving all three connections. Molecules that are returned by the third query will contain a ring fused to the pyridine R-group.

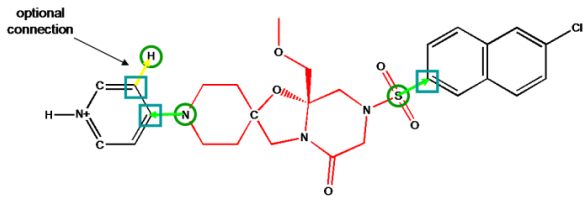


***Figure 4**. A scaffold replacement query showing two essential connections (green arrows) and a single optional connection (yellow arrow).*

Using the application in the second Replace Scaffold mode, "Select R Atoms," and selecting the atoms in the blue squares in Figure 3 to define the R-groups, results in the same calculation as above. In this mode a range of contiguous bonds can be selected in which the connection can occur, resulting in multiple substitution points being covered by a single search. For example, selecting the sulfur atom as well generates two queries, one involving the N-C and S-C bonds and one involving the N-C and N-S bonds (Figure 5).

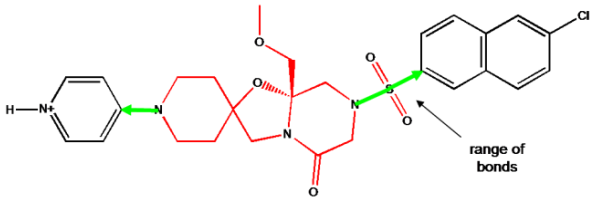


***Figure 5**. Using the "Select R Atoms" mode, multiple contiguous exit vectors can be selected in a single search, automatically creating multiple pharmacophore queries.*

Link Fragments. This mode allows two or more fragments to be joined with a linker. Connections are indicated by picking hydrogen atoms, with at least one connection required for each fragment.

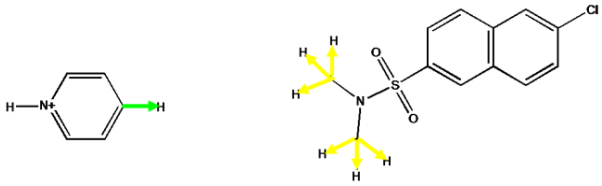


***Figure 6**. Linking multiple fragments with a combination of essential (green arrow) and optional (yellow arrows) exit vectors.*

In Figure 6, two fragments are shown with seven hydrogen atoms selected as connection points. The connection from the pyridine on the left is considered essential as it is the only connection specified for that fragment. The six on the right are optional, although at least one of them must be present in every generated query. This example generates 56 queries, all involving the essential left-hand connection plus a combination of up to four optional connections from the right-hand fragment. (The maximum number of optional connection points is four). If one of the optional connections is replaced by an essential connection, the number of queries is reduced to 8: one with no optional connections, 3 with one optional connection, 3 with two and 1 with three.

Add Group to Ligand. This mode allows the user to extend a ligand at one or more sites. As in the Link Fragments mode, connections are indicated by picking hydrogen atoms, and if the heavy atom is also selected, the connection is essential. In Figure 7 there is one essential connection (green arrow) and two optional connections (yellow arrows). Four queries are generated: one with just the essential connection, two queries with the essential connection and one of the optional connections and one with all three connections. If a single connection point is specified, a large number of hits can result. In this case, extra filters such as pharmacophore queries, QSAR or fingerprint models can be applied to ensure a reasonable number of hits.

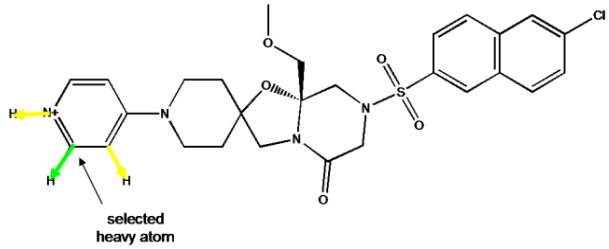


***Figure 7**. Add Group to Ligand mode, showing one essential connection (green arrow) and two optional connections (yellow arrows).*

Scaffold Database Preparation. Once the queries have been automatically generated, one or more databases of conformations of small molecular fragments are searched. There is no clear scientific definition of a molecular fragment, although there are some commonly used rules that help identify suitable molecules [Congreve 2003]. In addition to such rules, when using fragments in scaffold replacement it is necessary for the fragment to contain a site suitable for chemical substitution. Fragments can be found by searching vendor catalogues specifically designed to meet such fragment rules, although such catalogues often do not provide sufficient geometric diversity to serve as R-group linkers. Techniques based on the enumeration of molecular graphs can also be used to generate novel scaffolds [Blum 2009], however these methods can produce molecules that are difficult to synthesize. Consequently, CCG has focused on the partitioning of existing molecules, thus increasing the likelihood of synthetic feasibility whilst retaining geometric and pharmacophoric diversity.

When filtering a database of potential scaffolds it is important to remember that connection to the R-groups may change the chemical type of the newly bonded atoms, or even affect the properties of a distantly-bonded atom. For example, an aldehyde group, generally undesirable due to its reactive nature, may be joined to an amine to form an amide bond. Similarly, the new bond can stabilize minor tautomeric forms. In Figure 8, a keto-enol group that would prefer the aldehyde tautomer when free in solution would show the alternative bonding pattern when the oxygen forms part of an ether group. It is therefore important to

calculate a range of tautomers for the searched scaffolds, as their final chemical context will not be known, and result molecules should be filtered for undesirable properties, rather than the input scaffolds and fragments. Finally, protonation states should also only be assigned once the result molecule has been assembled.

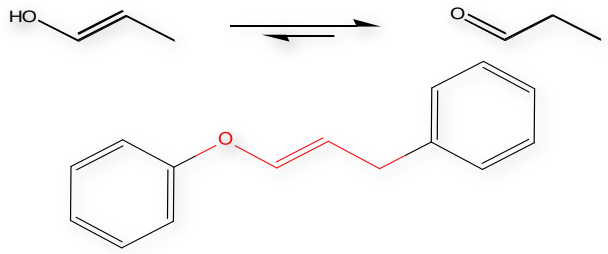


***Figure 8**. A minor tautomer of a scaffold (highlighted in red) may be stabilized when bonded to R-groups.*

The following method was used to create the MOE fragment conformational database suitable for fragment linking, growing and scaffold replacement. 4.3 million molecules were obtained from vendor catalogues and 10 years of medicinal chemistry literature. A rule-based tautomer and ionization state enumerator was applied, retaining most minor tautomers and protonation states. The resulting molecules (after filtering for lead-likeness) were subjected to four fragmentation methods. Two of these methods emphasize ring blocks, retaining either the ring block plus exocyclic double bonds, or ring blocks connected by a single rotatable bond (e.g., biphenyl). The other two methods are retrosynthetic in nature. The first was to fragment molecules by cutting bonds identified using RECAP rules [Lewell 1998]. The second was to subject each molecule to a Schuffenhauer decomposition [Schuffenhauer 2007] to produce a sequence of fragments, each corresponding to a step-wise reduction in the complexity of the molecule. The resulting unique fragments were subjected to a full conformational search to produce a set of conformations with low strain energy, as determined using the MMFF94 forcefield [Halgren 1996]. The resulting database contained 800,000 fragments and 16.5 million conformations.

Filtering the Result Molecules. Once the new fragment has been bonded to the target R-groups, user-defined filters are automatically applied. These filters can include other MOE pharmacophore query features, which gives the possibility of using 3D SMARTS matches in addition to the more standard feature types, along with QSAR and fingerprint models. The use of a 3D SMARTS match in a pharmacophore query enables specific chemical groups to be required at particular substitution points, or to ensure a particular group appears in the result molecules. A simple text interface for applying molecular property and 2D SMARTS-based filters has been developed to allow the user to easily define custom filter rules. By default this is set up to ensure reasonable, drug-like molecules using molecular weight, SlogP and TPSA [Wildman 1999, Ertl 2000]. Additionally for these three molecular properties, an estimate of their value for the result molecule is made during the search phase to avoid investigating linkers that would not be able to match the filter.

Estimation of Synthetic Plausibility. Linking molecular fragments purely on geometric considerations can lead to unstable or even unsynthesizable molecules. There are two solutions to this problem: a) guide the growth of molecules by following reaction pathways or RECAP-style connections; b) decide after the fact whether the proposed molecule is stable and synthesizable via some scoring scheme or disconnection protocol. In MOE Scaffold Replacement the latter option has been used, since fragment placement and geometric constraint filtering is relatively fast.

There is a speed-accuracy tradeoff when assessing synthetic feasibility. The more accurate retrosynthetic analysis systems, e.g. LHASA [Johnson 1992], are not fast enough to be applied to hundreds of thousands of proposed molecules. Methods based upon molecular graph complexity or other descriptors are very fast but highly inaccurate due to the difficulty of encoding sufficiently detailed chemical rules into such descriptors. Approximate techniques such as RECAP or BRICS [Degen 2008] focus on cutting chain connections rather than ring bonds, which can result in an overestimation of complexity. CCG's approach to calculating chemical plausibility is to perform extensive retrosynthetic disconnections and report the fraction of atoms in the resulting fragments found in a reference database of starting materials. Heterocycles are not necessarily preserved; for example, most 5-ring heterocycles can be formed with imine (or related) hydrolysis, a reaction that is absent from the RECAP and BRICS rule set. In Figure 9, the depicted oxazole can be opened by imine hydrolysis, resulting in two simpler fragments. By repeated application of approximately 20 retrosynthetic pathways, most starting materials can be decomposed to simple precursors that can be compared to the database of reference starting materials.

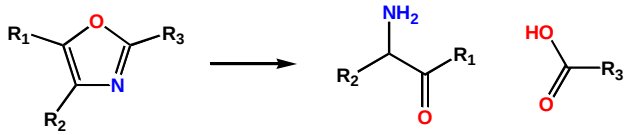


***Figure 9**. Imine hydrolysis can be used to open heterocycles.*

To illustrate the process on a complete molecule, BIRB 796, a ligand for p38 MAP kinase, is fragmented as shown in Figure 10. The pyrazole is cleaved with imine hydrolysis, the moropholine fragment with aryl ester chemistry and the urea with amide hydrolysis. All of the resulting fragments are found in the reference database, leading to a chemical plausibility score of 100%. The procedure is very fast allowing for a processing rate of hundreds of molecules per second. Most importantly, it is easy to add new fragments to the database of starting materials, allowing users to customize the scores to better account for particular chemistries available to them.

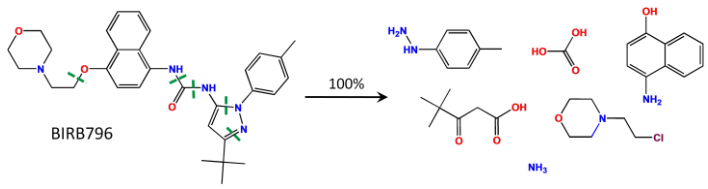


***Figure 10**. BIRB 796 is reduced to a set of starting materials using various retrosynthetic disconnections. All the starting materials can be identified in the reference database, so the chemical plausibility score is 100%.*

RESULTS AND DISCUSSION

The p38 Mitogen-Activated Protein Kinase (MAPK) pathway is a therapeutic target for inflammatory diseases such as psoriasis, rheumatoid arthritis and chronic obstructive pulmonary disease [Kumar 2003, Kaminska 2005]. Many p38α kinase inhibitors have been identified, most of which have been shown to be competitive with ATP, binding to the kinase active site [Kaminska 2005, Lee 1994, Adams 2001]. A few inhibitors have been found to bind to a neighbouring site, called the "DFG-out" site, formed by a conformational change in the conserved DFG sequence, one of which is a bi-aryl urea compound known as BIRB-796 [Pargellis 2002]. This compound has well-defined SAR available [Regan 2002, Regan 2003] and will serve as the basis for the following demonstration of MOE Scaffold Replacement.

The structures of p38 kinase in complex with BIRB-796 and the initial screening hit that lead to BIRB-796 have been deposited in the PDB under codes 1KV1 and 1KV2, respectively (Figure 11).

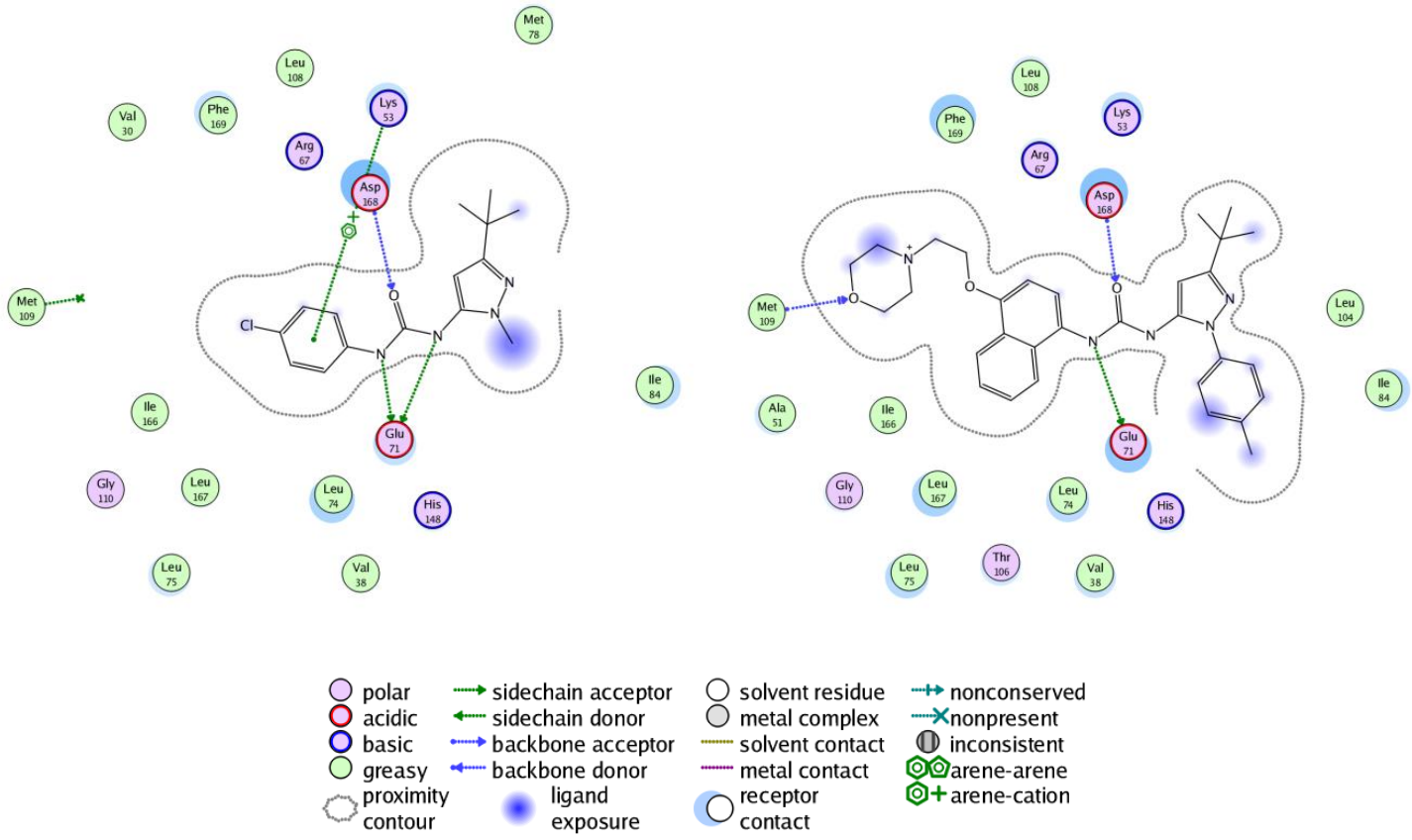
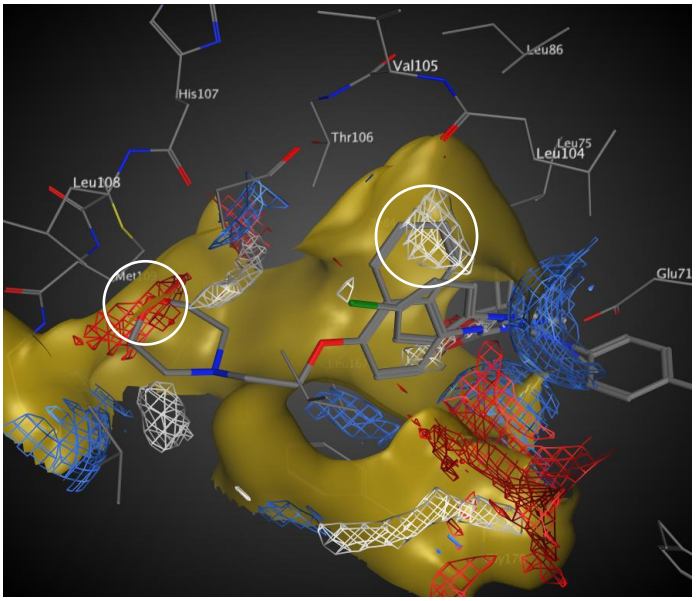


Figure 11*. The protein-ligand interaction diagrams calculated from PDB structures 1KV1 (left) and 1KV2 (right). Residues have been laid out in a consistent manner between the two diagrams to allow easy comparison.*

During the progression from initial hit to BIRB-796, three modifications occurred:

1. The replacement of a methyl group on the pyrazole with a tolyl group (resulting in a ~140-fold enhancement in affinity)
2. The replacement of a chlorophenyl with a naphthyl group (~15-fold enhancement)
3. The addition of an ethoxymorpholine group (~11-fold enhancement)

It was decided to test whether MOE Scaffold Replacement could have predicted the final compound based on the initial hit ligand. To begin with, changes two and three were investigated. The structure in 1KV1 was prepared using MOE's Protonate3D [Labute 2009] tool to add hydrogen atoms and assign ionisation states. The electrostatic field for the p38 kinase structure in 1KV1 (Figure 12) was then calculated and regions of favourable location for donor, acceptor and hydrophobic probes were plotted as isocontours at -2, -2 and -3 kcal/mol, respectively. It can be seen in the figure that the morpholine oxygen is in a region of acceptor preference, while the naphthyl group is situated in a region of hydrophobic preference, which agrees with the experimentally determined increases in affinity from the initial hit to BIRB 796. A three-point pharmacophore was developed, as shown in Figure 13. Two of the pharmacophore features encode the bond between the morpholine oxygen and the backbone of Met-109 (features F1 and F2 in the diagram), while the third feature requires an aromatic group in the pocket that BIRB-796 fills with the naphthyl group (F3).



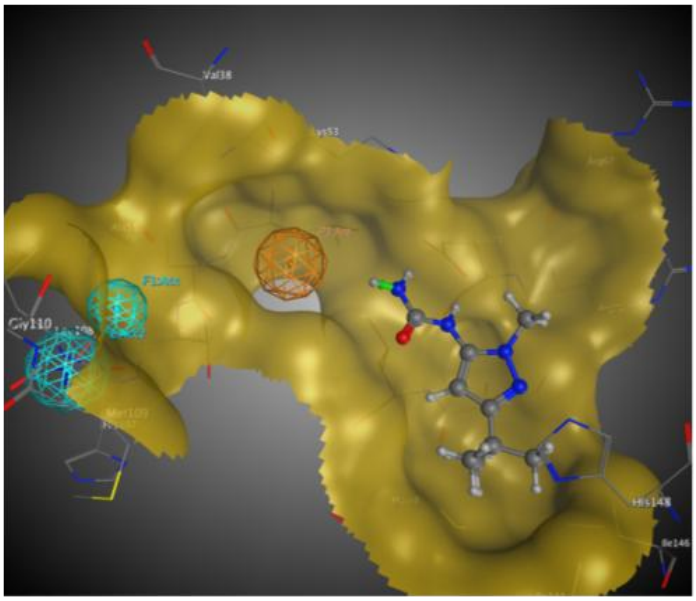


Figure 12. *The active site of 1KV1 with the two ligands overlaid. The electrostatic field was calculated using MOE's Surfaces and Maps tool, and wireframe volumes show regions of favourable interaction between the receptor atoms and acceptor (red), donor (blue) and hydrophobic (white) probe atoms, plotted as isocontours of interaction energy at -2 kcal/mol, -2 kcal/mol and -3 kcal/mol respectively. Two regions where BIRB-796 has complementary chemical groups in these regions are highlighted with white circles. The surface of the receptor, calculated using a grid-Connolly method around receptor atoms within 4.5 Å of the ligand, is shown in gold.*

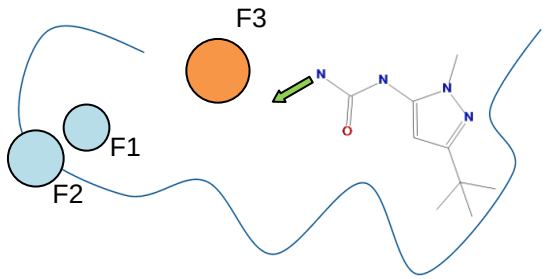


Figure 13*. The scaffold replacement and pharmacophore query used to generate novel p38 kinase inhibitors, shown both in 3D and schematically. The green arrow indicates the exit vector from the trimmed 1KV1 ligand, and the three pharmacophore features are shown as wireframe circles in the 3D view and solid circles in the 2D view.*

The ligand in 1KV1 was edited to remove the chlorophenyl group, and a single bond vector was specified in the Scaffold Replacement tool (shown as a green arrow in Figure 13). The search database was a set of 16.5 million conformations calculated for over 800,000 fragments extracted from a combination of vendor catalogues and medicinal chemistry literature, as described in the methods section. Result molecules were filtered to ensure a molecular weight less than 500 Da, SlogP in the range -4 to 8, TPSA in the range 40, 140, and rsynth (the synthetic plausibility score described above) greater than zero. Excluded volumes were automatically generated based on the pocket atoms, to ensure the results did not clash with the receptor, and on the retained portion of the ligand, to avoid intramolecular clashes. Molecules were ranked using the London dG scoring scheme.

This search resulted in 8958 hits, a number of which share features with the optimised BIRB 796 ligand. Naphthyl groups are frequently found in the hits, in common with BIRB 796. The acceptor features are commonly matched by alcohol groups, the presence of which can result in poor DMPK properties. To exclude these molecules, feature F1 in the query was modified to be "Acc!Don" (acceptor NOT donor). The result molecules were filtered using this new query, which reduced the number of hits to 5919. The top ten ranked molecules are shown in Table 1. In this reduced list, the first molecule containing a group similar to the morpholino of BIRB 796 is ranked seventh.

Figure 14 shows two results from the search with features similar to BIRB 796. The compound on the left of the figure, ranked first by London dG, has the same naphthyl group as BIRB-796, and forms the hydrogen bond with Met-109 using a carbonyl group. The compound on the right, ranked seventh by London dG, uses a morpholino group to form the hydrogen bond, just like BIRB-796. This shows that it would have been possible to proceed from the original compound to a near neighbour of the clinical candidate with a suitably constructed scaffold replacement query.

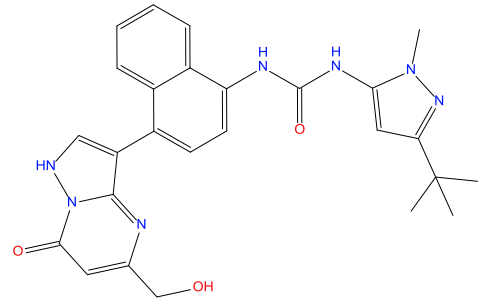
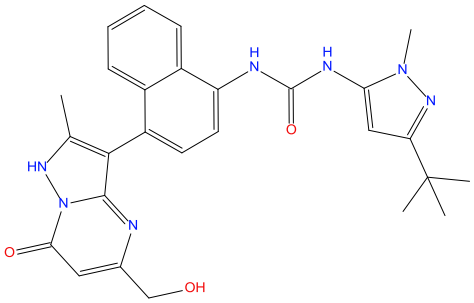
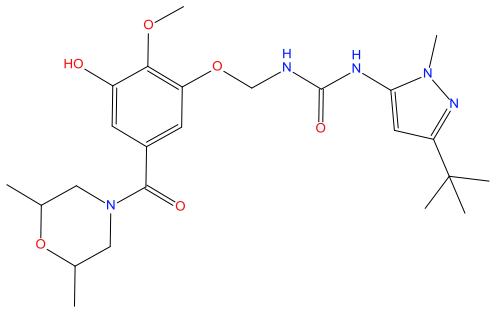
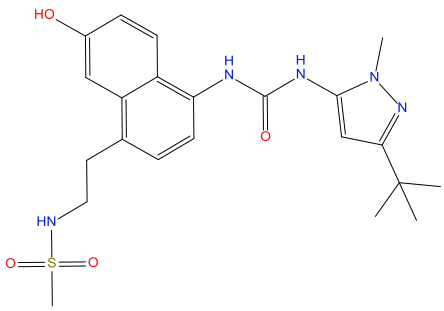
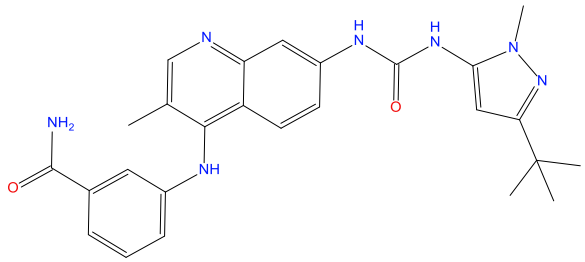
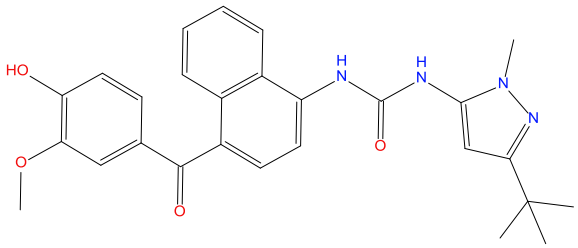
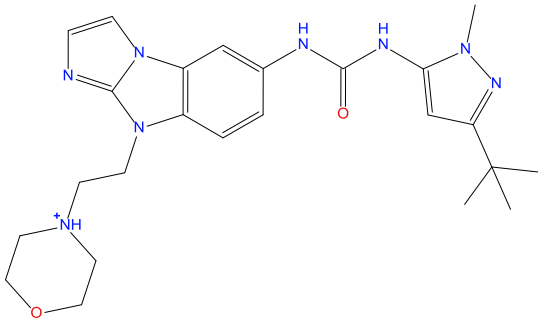
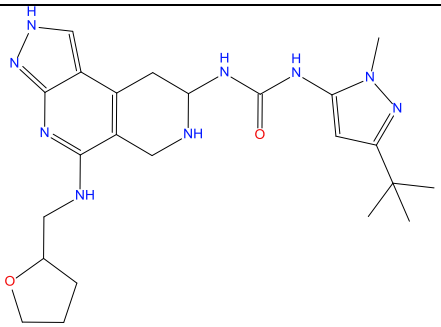
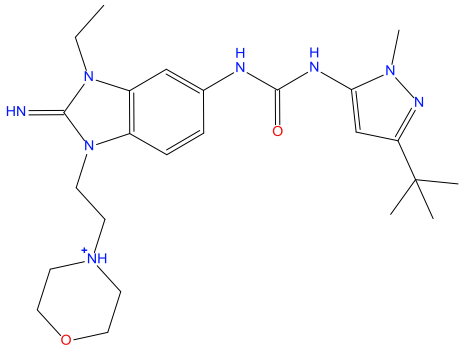
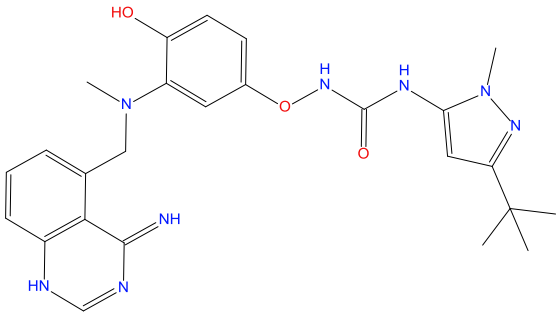
 London dG: -18.61	 London dG: -18.21
 London dG: -18.05	 London dG: -17.62
 London dG: -17.39	 London dG: -17.32
 London dG: -17.20	 London dG: -17.17
 London dG: -17.05	 London dG: -17.04

Table 1. The top ten proposed inhibitors of p38 kinase, as ranked by their London dG score

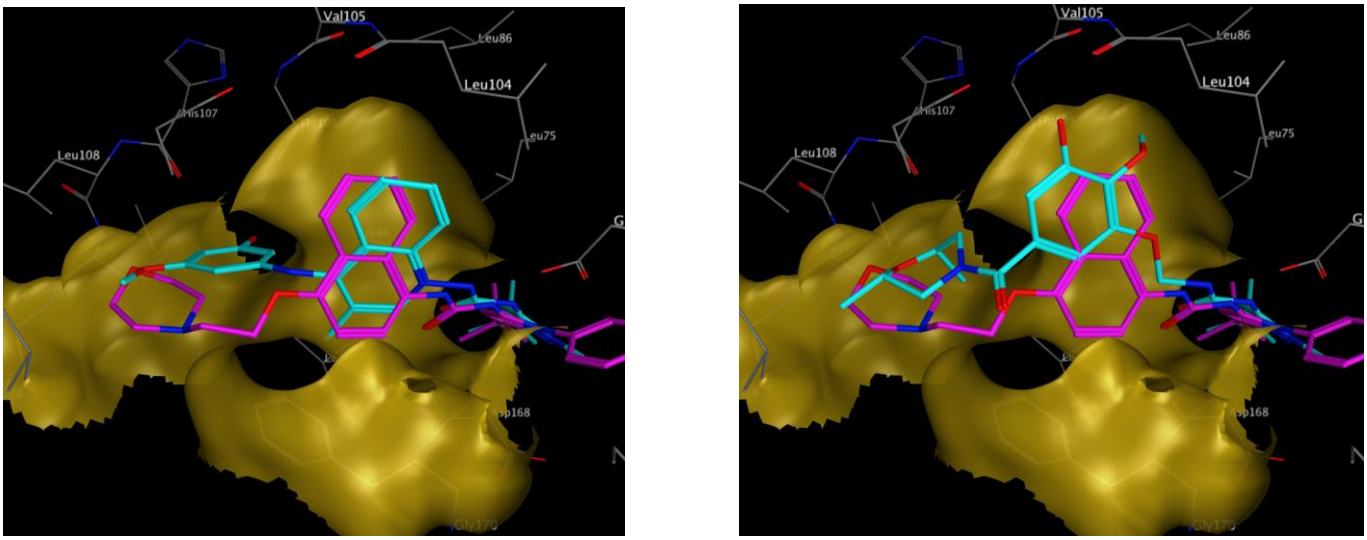


***Figure 14**. Sample results from the search shown in Figure 13. Proposed molecules are shown with carbon atoms coloured cyan, BIRB-796 is shown with carbons coloured magenta.*

A second experiment was performed to investigate the possibility of replacing the urea scaffold of BIRB-796, along with the pyrazole and *tert*-butyl groups, starting from the 1KV2 structure. The role of the pyrazole is structural, maintaining the position of the tolyl and *tert*-butyl groups, hence it can be targeted for replacement. The urea group forms important hydrogen bonds with the sidechain of Glu-71 and the backbone nitrogen of Asp-168. The region to be replaced is shown in Figure 15. The 1KV2 structure was prepared using Protonate3D. In this experiment the Scaffold Replacement tool was used in the Replace Scaffold (Select R-Atoms) mode, with one atom from the naphthyl and one atom from the tolyl group selected. The tolyl group was not replaced because this has a large contact surface with Glu-71, filling the pocket vacated by Phe-169 during the conformational change to the DFG-out structure [Pargellis 2002]. To ensure that the result molecules form the same interactions with the protein as BIRB-796, a three point pharmacophore was used (Figure 15):

1. A projected donor feature on the sidechain carboxylate of Glu-71.
2. An acceptor feature on the urea oxygen to maintain the interaction with Asp-168.
3. A hydrophobic feature placed at the centroid of the *tert*-butyl group near the hydrophobic pocket that is exposed in the DFG-out conformation.

Excluded volumes were automatically generated to prevent candidate linkers from clashing with the receptor and R-group atoms. Result molecules were minimised in the active site of the kinase as seen in the 1KV2 structure using the MMFF94x forcefield and Reaction Field solvation model with a final gradient of 0.01 kcal/mol/Å, and ranked using the London dG scoring scheme. Receptor atoms were held fixed during minimisation, and ligand atoms that matched the pharmacophore features included in the search were tethered with a flat-bottomed potential to the location of the pharmacophore feature.

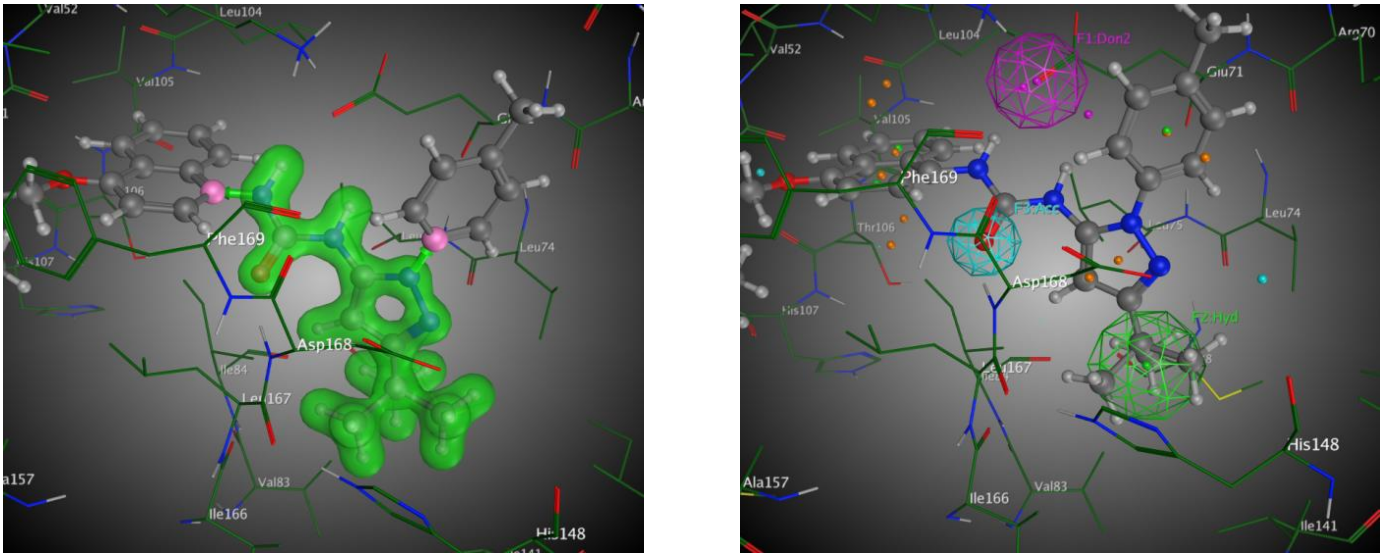


***Figure 15**. Replacing the urea, pyrazole and tert-butyl groups. Left: The R-group root atoms selected for the search are highlighted in pink, while the region to be replaced is highlighted in green. Right: The pharmacophore features used to focus the search are shown as wireframe spheres.*

In order to calculate a benchmark score to judge the new molecules, BIRB-796 was minimized in the active site using the same procedure that was used on the result molecules. The London dG score for the minimised structure was -13.95. Searching the database of conformations for the 800,000 fragments identified 16 hit molecules, with London dG scores ranging from -12.55 to -15.25. Five selected molecules from this set are shown in Table 2. Figure 16 shows two of the selected molecules in the active site; on the left, an amide group replaces the urea and the pyrazole is replaced by a six-membered ring, while on the right, the urea is replaced by a sulphonamide, which is linked to the tolyl group by an alkane chain.

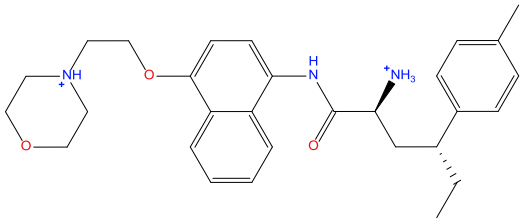
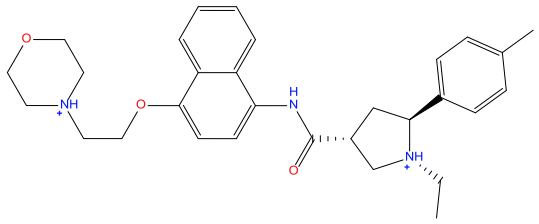
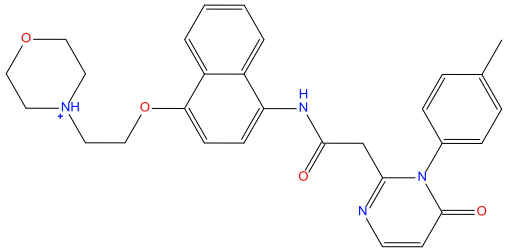
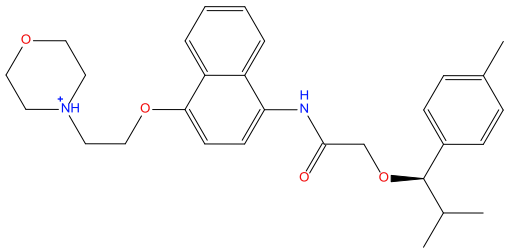
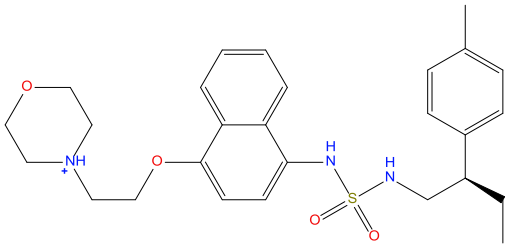
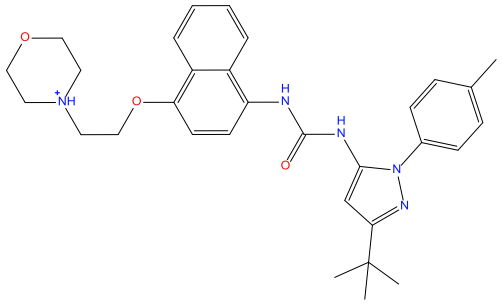
 London dG: -15.24	 London dG: -14.34
 London dG: -13.72	 London dG: -13.15
 London dG: -13.12	 **BIRB-796** London dG: -13.95

***Table 2**. The novel scaffolds for p38 kinase based on the BIRB-796 ligand (shown lower left).*

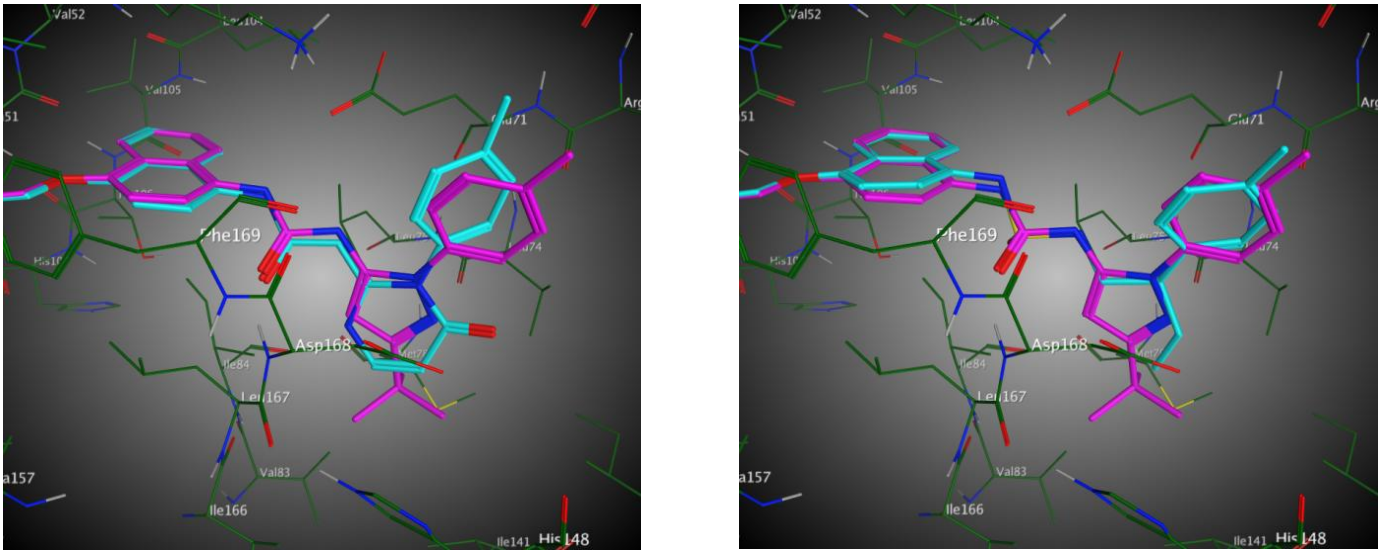


***Figure 16**. Sample results from the search shown in Figure 15. Proposed molecules are shown with carbon atoms coloured cyan, BIRB-796 is shown with carbons coloured magenta.*

CONCLUSIONS

MOE's scaffold replacement tool is an integrated tool for scaffold replacement and fragment linking and growing. Complex queries can easily be constructed, including multiple exit vectors from the template molecules and a mix of optional and essential connections. Result molecules can be filtered using both 2D and 3D constraints, such as QSAR and fingerprint models. Pharmacophore queries can also be used to ensure the proposed molecules make particular interactions with a receptor. If a receptor is available, excluded volumes are automatically generated to prevent clashes with receptor atoms, and optionally the result molecules can be refined in the binding pocket, which can be assigned varying degrees of flexibility, and the interactions with the receptor can be scored. Any database of molecular conformations can be searched using this method. CCG provides a database of 800 000 fragments derived from medicinal chemistry literature and chemical vendor catalogues; this can easily be augmented with in-house compounds to create a custom linker library. A method for calculating the synthetic feasibility of the new molecules has been developed, based upon the fragmentation of input molecules by known reactions, including those required for ring-opening, and the comparison of the result fragments to those from known compounds.

The results of two example searches based on ligands for the p38 kinase have been shown. In the first search, the progression from hit compound to clinical candidate BIRB-796 is reproduced *in silico*, and in the second search, replacement scaffolds for BIRB-796 are calculated, suggesting possibilities for novel p38 inhibitors.

REFERENCES

[Adams 2001] Adams, J.L., Badger, A.M., Kumar, S., Lee, J.C.; P38 MAP Kinase: Molecular Target for the Inhibition of Pro-Inflammatory Cytokines; *Prog. Med. Chem 38* **(2001)** 1-60.

[Blum 2009] Blum, L.C., Reymond, J.L.; 970 Million Druglike Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13; *J. Am. Chem. Soc. 131(25)* **(2009)** 8732-8733.

[Congreve 2003] Congreve, M., Carr, R., Murray, C., Jhoti, H.; A Rule of Three for Fragment-Based Lead Discovery; *Drug Discov. Today 8(1)* **(2003)** 876-877.

[Degen 2008] Degen, J., Wegscheid-Gerlach, C., Zaliani, A., Rarey, M.; On the Art of Compiling and Using 'Drug-Like' Chemical Fragment Spaces; *Chem. Med. Chem. 3(10)* **(2008)** 1503-1507.

[Deschênes 2007] Deschênes, A., Sourial, E.; Ligand Scaffold Replacement Using MOE Pharmacophore Tools; *J. Chem. Comp. Group* **(2007)**; see also http://www.chemcomp.com/journal/scaffold.htm

[Ertl 2000] Ertl, P., Rohde, B., Selzer, P.; Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties; *J. Med. Chem. 43(20)* **(2000)** 3714-3717.

[Halgren 1996] Halgren, T.A.; The Merck Force Field; *J. Comp. Chem. 17* **(1996)** 490-641.

[Johnson 1992] Peter Johnson, A., Marshall, C., Judson, P.N.; Starting Material Oriented Retrosynthetic Analysis In The LHASA Program; *J. Chem. Inf. Comput. Sci. 32(5)* **(1992)** 411-417.

[Kaminska 2005] Kaminska, B.; MAPK Signalling Pathways as Molecular Targets for Anti-Inflammatory Therapy – from Molecular Mechanisms to Therapeutic Benefits; *Biochim. Biophys. Acta. 1754(1-2)* **(2005)** 253-262

[Klaus 2002] Regan, J., Breitfelder, S., Cirillo, P., Gilmore, T., Graham, A.G., Hickey, E., Klaus, B., Madwed, J., Moriak, M., Moss, M., Pargellis, C., Pav, S., Proto, A., Swinamer, A., Tong, L., Torcellini, C.; Pyrazole Urea-Based Inhibitors of p38 MAP Kinase: from Lead Compound to Clinical Candidate; *J. Med. Chem. 45(14)* **(2002)** 2994-3008.

[Kumar 2003] Kumar, S., Boehm, J., Lee, J.C.; P38 MAP Kinases: Key Signalling Molecules as Therapeutic Targets for Inflammatory Diseases; *Nat. Rev. Drug Discovery 2* **(2003)** 717-726.

[Labute 2009] Labute, P.; Protonate3D: Assignment Of Ionization States And Hydrogen Coordinates To Macromolecular Structures; *Proteins 75(1)* **(2009)** 187-205.

[Lee 1994] Lee, J.C., Laydon, J.T., McDonnell, P.C., Gallagher, T.F., Kumar, S., Green, D., McNulty, D., Blumenthal, M.J., Heys, J.R., Landvatter, S.W., et al; A Protein Kinase Involved in the Regulation of Inflammatory Cytokine Biosynthesis; *Nature 372* **(1994)** 739-746.

[Lewell 1998] Lewell, X.Q., Judd, D.B., Watson, S.P., Hann, M.M.; RECAP — Retrosynthetic Combinatorial Analysis Procedure: A

Powerful New Technique for Identifying Privileged Molecular Fragments with Useful Applications in Combinatorial Chemistry; *J. Chem. Inf. Comput. Sci. 38(3)* **(1998)** 511-522.

[Pargellis 2002] Pargellis, C.A., Tong, L., Churchill, L., Cirillo, P., Gilmore, T., Graham, A.G., Grob, P.M., Hickey, E.R., Moss, N., Pav, S., Regan, J.; Inhibition of P38 Map Kinase by Utilizing a Novel Allosteric Binding Site; *Nat. Struct. Biol. 9* **(2002)** 268-272.

[Regan 2003] Regan, J., Capolino, A., Cirillo, P.F., Gilmore, T., Graham, A.G., Hickey, E., Kroe, R.R., Madwed, J., Moriak, M., Nelson, R., Pargellis, C.A., Swinamer, A., Torcellini, C., Tsang, M., Moss, N.; Structure-Activity Relationships of the P38alpha MAP Kinase Inhibitor 1-(5-tert-butyl-2-p-tolyl-2H-pyrazol-3-yl)-3-[4-(2-morpholin-4-yl-ethoxy)n aph-thalen-1-yl]urea (BIRB 796); *J. Med. Chem 46(22)* **(2003)** 4676-4686.

[Schuffenhauer 2007] Schuffenhauer, A.; Ertl, P., Roggo, S., Wetzel, S., Koch, M.A., Waldmann, H.; The Scaffold Tree - Visualization of the Scaffold Universe by Hierarchical Scaffold Classification; *J. Chem. Inf. Model. 47(1)* **(2007)** 47-58.

[Wildman 1999] Wildman, S.A., Crippen, G.M.; Prediction of Physiochemical Parameters by Atomic Contributions; *J. Chem. Inf. Comput. Sci. 39 (5)* **(1999)** 868-873.