Scaffold-hopping and fragment-linking using the Cambridge Structural Database

Aim

1) To use the Cambridge Structural Database to provide ideas for novel scaffolds giving an active inhibitor of known binding mode
2) To use the Cambridge Structural Database to find good linkers between two fragments binding in different parts of the active site

Introduction

Scaffold-hopping and fragment-linking are both drug-design methodologies that are widely used in drug discovery today. Scaffold-hopping requires an active molecule with either an experimentally determined, or a hypothetical, binding conformation. A central part of the active molecule is then replaced by a new scaffold which is able to retain the original binding groups of the molecule in their good binding orientation. Re-scaffolding allows new scaffolds to be found which might impart better physicochemical properties or which might avoid existing patent art[1].

Fragment-linking is an important methodology in fragment based design. Two fragments with known binding-modes may be identified by high-throughput X-Ray or NMR methods[2]. The task is to find good linkers between the fragments that can retain them in the same orientation without significant strain energy being introduced[3].

Clearly from the computational chemistry standpoint these two methodologies are closely related. Past approaches to the problem commonly involve first constructing a set of bond vectors representing the bonds that a) will be the connection points for the new scaffold in a scaffold-hopping approach or, b) that represent the bonds that will be made to each fragment. 3D databases or collections of conformers (either stored or generated on the fly) are then matched to the vectors and good matches identified[4,5].

A key criterion is that the scaffold/linker should have achievable low strain geometry. For this reason the Cambridge Structural Database (CSD) is recognised to be an important source of information to be used for either of these approaches[4] and libraries of CSD derived scaffolds are available for use with other re-scaffolding packages for CSDS license holders.[6]

Software has now been developed by CCDC that can allow both scaffold hopping and fragment linking to be carried out relatively simply utilising the information in the CSD. This use-case demonstrates some simple examples to demonstrate both methodologies.

Method

Scaffold Hopping: The Packing Feature search functionality in Materials Mercury is used. Scaffold hopping is illustrated using an inhibitor of Factor Xa. Factor Xa is an important element in the coagulation cascade and Factor Xa inhibitors have been designed as orally available antithrombotic agents. The ligand from PDB entry 2w26 is an antithrombotic candidate drug in clinical testing[7]. This ligand was imported into Mercury in its protein bound conformation. Five atoms were picked to represent the conserved vectors between which the new scaffold was to be placed (Fig. 1). There is available an option to vary atom type, valence and number of hydrogens for the selected atoms and this was used to allow the retrieval of matches where the sp2 nitrogen atoms could optionally also be sp2 carbon atoms. The second most stringent restriction on geometric similarity was applied (distances between vectors in the hits must be conserved within 20%; angles within 12%). The full Cambridge Structural Database was searched. 409 hits were obtained, of which 26 had RMSD values to the five query atoms of < 0.25Å.

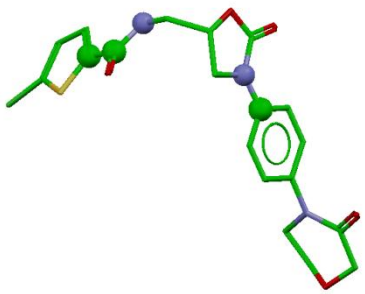

Figure 1. Factor Xa inhibitor from PDB structure 1w26 with atoms highlighted, which define the section for rescaffolding.

Fragment linking: The Packing Feature search functionality in Materials Mercury is used. In this example the fragments to be linked are taken from the crystal structures of two different agonists of PPAR-Gamma. The pdb codes for the appropriate protein structures are 2fvj and 2f4b. An initial required step is that these two proteins be superimposed in the binding site region. This was carried out with Relibase+ using a Similar Binding Sites search and superposition[8]. The fragments selected for linking together are situated in two different sub-pockets. Each ligand only fills one of the sub-pockets (Fig. 2). In this example a hybrid molecule that did potentially fill both pockets might be an interesting target and might show a different balance of antagonist versus agonist behaviour than that of the two source agonists. Alternatively the fragment placements might come from X-Ray or NMR based fragment screening experiments.

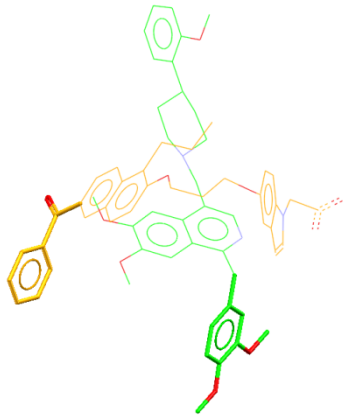

Figure 2. PPAR-G ligands from PDB structures 2fvj and 2f4b. The atoms highlighted indicate the fragments to be linked.

Both fragments are required to be merged into the same entry and an 'intramolecular' option is required to be set so that the packing Feature search retrieves structures which contain linkers to both fragments. Figure 3 shows the relative orientation of both fragments and indicates the atoms that were selected in order to carry out the Packing Feature search. The linker atom on the dimethoxy benzyl fragment was allowed to be oxygen as well as carbon.

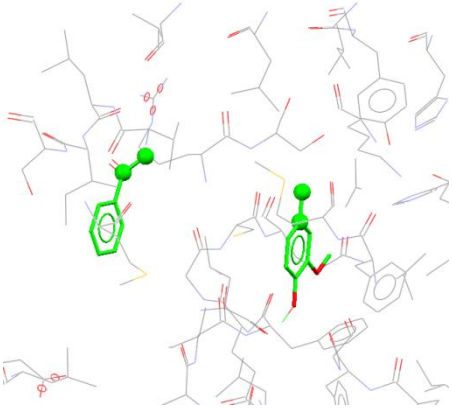

Figure 3. PPAR-G model with the two fragments bound. The atoms between which a linker will be sought, are highlighted.

Results

Scaffold Hopping: Figure 4 illustrates some of the scaffolds generated by the search that might be valid replacements for the scaffold in 2w26. The identification of such scaffolds then provides ideas and material for further design work. This will be necessary to demonstrate these are genuine design ideas that should be seriously considered. Additional steps might include checking the scaffold geometry with the geometry validation tool Mogul to confirm that it is indeed a low strain scaffold; and carrying out further modelling on the composite molecules, for instance by docking them using the docking package GOLD.

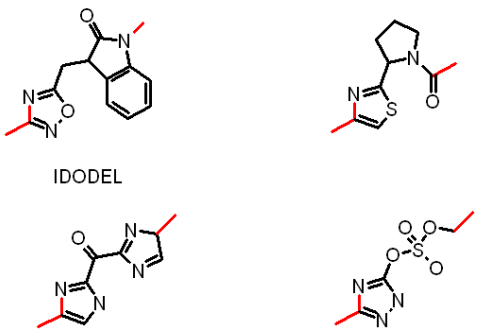

Figure 4. Four examples of scaffold replacements for the 2w26 factor Xa inhibitor. The link bonds are marked in red.

Figure 5 shows how one hit, CSD entry IDODEL, overlays in the active site of Factor Xa (relevant portions of the molecule are highlighted in yellow, the original ligand is in green). The central scaffold appears to be capable of making a good hydrogen bond with the Gly 218 and could easily be modified further (e.g. from the benzo group) to make better interactions with the binding site. This appears a particularly interesting overlay because IDODEL also contains a group that sits well in the S1 pocket, thus providing a new idea here too.

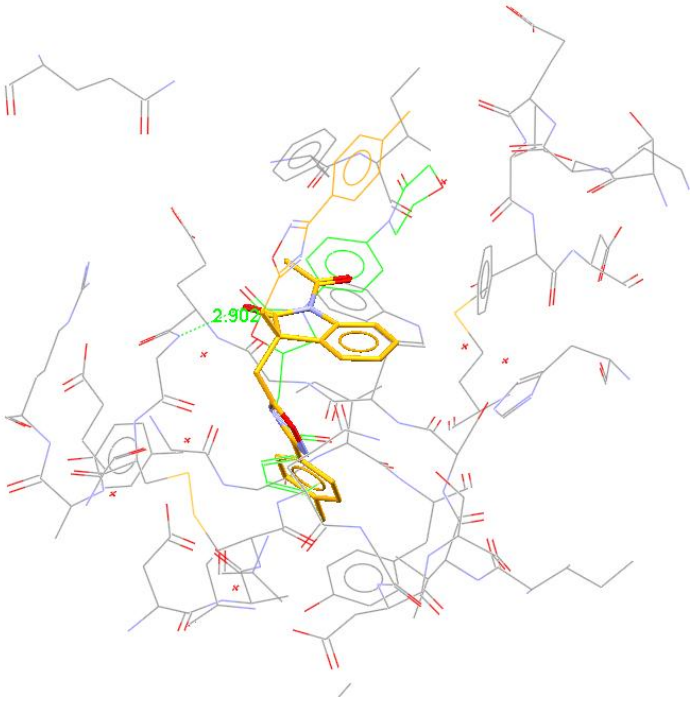

Figure 5. Superposition of the hit IDODEL over the 2w26 ligand. The hydrogen bond that this scaffold might make to Gly 218 is shown.

Fragment Linking: Figure 6 shows some of the linking groups that might be considered further for joining the two PPAR-G fragments. Figure 7 shows the closeness of fit of two of the linkers to the fragments to be bridged. These and others generated by this search, provide design ideas that, with additional modelling validation and consideration of synthetic chemistry, can lead to reasonable design candidates.

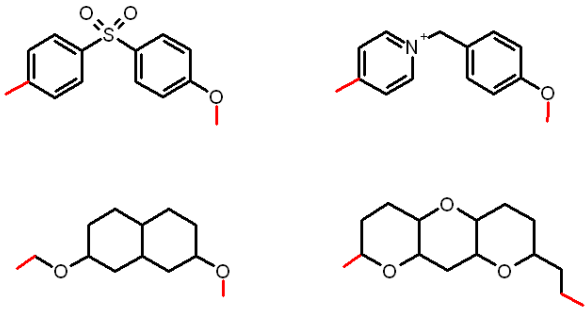

Figure 6. Linkers that are good geometric matches (as measured by RMSD at the linker attachment points of the two fragments). Link bonds are in red.

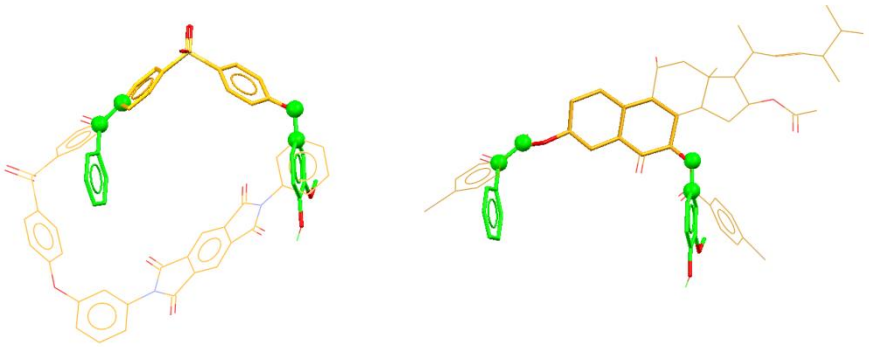

Figure 7. Superposition of the CSD entries BALVEQ and VAHKEV, hits from the Packing Feature search, with the two fragments to be linked, indicating what the composite designs look like.

Conclusions

This worked example illustrates how the Packing Feature search tool, part of the Materials module of Mercury, can be used in the structure-based redesign of ligands to a protein active site. The tool can very easily uncover re-scaffolding or fragment linking ideas that already exist in the Cambridge Structural Database. Consequently these ideas are already likely to have reasonable low strain geometry and therefore can be incorporated in to a composite design without adversely affecting the binding of the retained highly efficient fragments.

This module, now adapted for use by drug discovery researchers, is currently available as a Beta release add-on module to the Cambridge Structural Database System. We are currently looking for

beta testers for this module. If you are interested in evaluating this drug design software please contact support@ccdc.cam.ac.uk

References

1. H.-J. Boehm, A. Flohr, M. Stahl., *Drug. Disc. Today: Technologies,* **1**, (2004), 217-224.
2. M.Pellecchia, D. S. Sem, K. Wuthrich, *Nat. Rev. Drug. Discov.*, **1**, (2002) , 211, T.L. Blundell, H. Jhoti, C. Abell, *Nat. Rev. Drug. Discov.*, **1**, (2002) , 45,
3. M. Congreve, G. Chessari, D. Tisi, A. J. Woodhead, *J. Med. Chem.*, **51**, (2008), 3661-3680.
4. P. Maass , T. Schulz-Gasch, M. Stahl, M. J. Rarey, *J. Chem. Inf. Model.*, **47**, (2007), 390−399.
5. D. C. Thompson, R. A. Denny, R. Nilakantan, C. Humblet, D. Joseph-McCarthy, E. Feyant, *J. Comput. Aided Mol. Des.,* **22**, (2008), 761-772.
6. http://www.biosolveit.de/press/archive/2010-03-24_CCDC.html
7. S. Roehrig, A. Straub, J. Pohlmann, T. Lampe, J. Pernerstorfer, K. Schlemmer, P.Reinemer, E. Perzborn, *J. Med. Chem.,* **48**, (2005), 5900-5908.
8. http://www.ccdc.cam.ac.uk/products/life_sciences/relibase/

Products

CSD – the world's only comprehensive, fully curated database of crystal structures, containing over 500,000 entries.

Materials module of Mercury – a powerful exploration and comparison tool for solid state structures.

Mercury – a versatile and feature-rich visualisation tool for molecular structures.

Relibase+ - an essential tool for searching, exploring and comparing all protein-ligand data from public and in-house data sources. Relibase+ enables users to get the maximum benefit out of the structure data available.

Mogul – a knowledge base if CSD-derived molecular geometries which provides a quick route to validated and trustworthy molecular models.

GOLD – an accurate and reliable protein-ligand docking program.

For further information please contact Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK. Tel: +44 1223 336408, Fax: +44 1223 336033, Email: admin@ccdc.cam.ac.uk