# The Identification of Bioisosteres as Drug Development Candidates

By Tim Cheeseright at Cresset BioMolecular Discovery

Molecular fields give a richer, more intuitive view of active molecules. Application of this technology to the search for bioisosteres results in relevant, non-obvious suggestions that make a significant impact on the drug development process.

**Bioisosterism has played a central role in the development of drug molecules almost from the outset of the pharmaceutical industry. The promise of bioisosterism is that the properties of a compound can be fine-tuned without affecting its underlying biological activity. This promise is not however without its challenges. Successfully applying bioisosterism to achieve the desired molecular outcome is difficult because of the fundamental problem that chemical structure is an unreliable indicator of biological activity. Small changes in a molecule can have a profound impact on a compound's activity, specificity and toxicity, whilst completely different chemotypes may have near identical biological activity profiles. More rigorous and reproducible methods for suggesting relevant, non-obvious and yet synthetically intuitive bioisosteres would have wide applicability.**

## BIOISOSTERES IN DRUG DISCOVERY

Bioisosteres are used by researchers throughout the pharmaceutical industry to find new hits and leads by modifying known actives or substrates, to develop leads by modifying physicochemical properties and protecting their knowledge using patents.

Traditional drug discovery had bioisosterism at the core of its strategy for finding hits and leads. Having identified an interesting target, researchers often had little choice in finding an active inhibitor or antagonist, except through bioisosteric modification of the natural

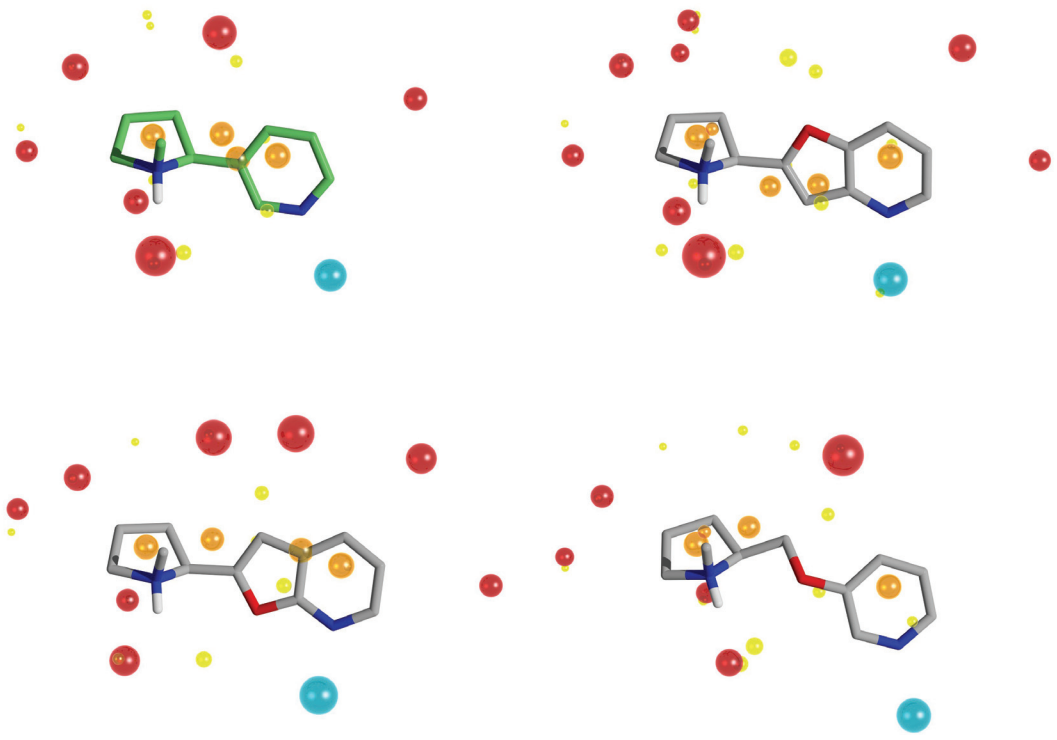


**Figure 1:** Nicotine (shown in green) bioisosteres (3) shown using 'field points'. The field point patterns for these single digit nanomolar actives at α4β2 neuronal nicotinic receptor are clearly highly similar

ligand in a systematic and thoughtful manner. The modern HTS era has provided plenty of potential leads, but the need for bioisosteres remains. Actives found through HTS can have undesirable properties (either physical or biological) and often lack novelty.

Once hits are found, the principle of bioisosterism is central to converting the hit into a more useful lead or drug. Medicinal chemists make biologically relevant changes to the active structure that result in better physicochemical properties or that move the hits into desirable intellectual property (IP) space, facilitating their move from hits to leads. During lead optimisation, the lead molecule may be 'tweaked' to introduce the correct activity/selectivity profile combined with good physicochemical properties. More radically, replacement of core groups in the lead series with new scaffolds that introduce better selectivity or physical properties can rescue what would otherwise be a failed project.

Having successfully negotiated the maze of drug discovery, companies often need to find a backup series that uses all the knowledge gained in a successful project but has unique IP and/or better off-target effects. The challenge is to introduce significant structural diversity whilst at the same time retaining as much of the information gained as possible. Bioisostere substitution is an obvious route to achieving this.

The requirement to protect research positions through patent applications is critical for the development of new medicines. In this respect, IP protection is probably the most important use of bioisosteres in the modern drug discovery project. Chemical structures derived from a common scaffold are often used as surrogates for true bioisosteres to define a range of chemical space that is deemed to have a common biological activity. However, due in large part to the inadequacy of structure to describe biological activity, there are numerous examples where the definitions in a patent are not sufficient to prevent competitors from finding closely related drug development candidates. This has led on many occasions to significant reductions in the value of the original application. Clearly, more automated methods for identifying biologically relevant bioisosteres would be of significant value in the definition of IP space.

## EXISTING APPROACHES TO FINDING NOVEL BIOISOSTERES

The application of bioisosterism in the current drug development process is widespread but also somewhat *ad hoc*, relying on the experience of medicinal chemists to invent new molecules. It is consequently firmly grounded in the chemical structure descriptions of active molecules, resulting in only small evolutionary jumps in the underlying chemistry. This limits the tested chemical space to closely related structures, potentially synthesising and testing inactive structures and, more importantly, missing a range of interesting candidates whose structures are not closely related to the lead.

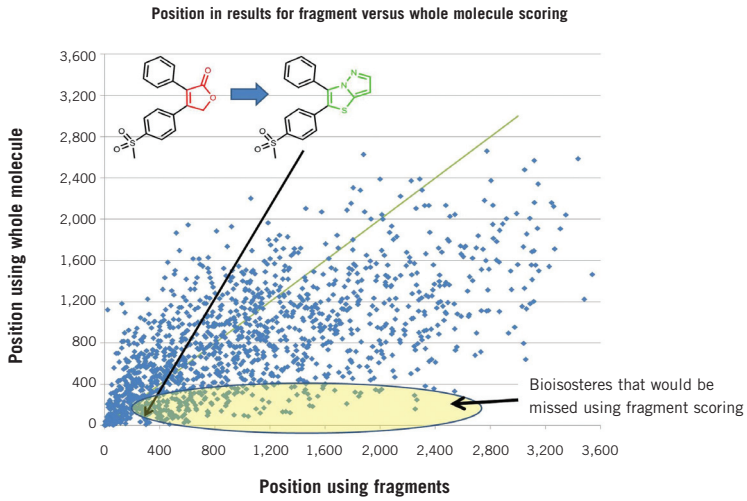
A method for describing molecules in a manner more related to their biological activity would have the potential to explore much broader sets of chemical space, suggest a wider range of modifications and enable research projects to progress faster. There has been increasing interest in methods that can suggest relevant, non-obvious, accurate bioisosteres to project teams. These methods broadly fall into two categories: knowledge-based approaches and computational techniques.

Knowledge-based approaches tend to use data mining techniques to find moieties that have previously been substituted for each other without a significant change in the activity under study. Thus the replacement of arginine by benzamidine in the Thrombin and Factor Xa projects of the 1990s would register benzamidine and arginine as bioisosteric. This approach has broad appeal; it can highlight changes that are known to be successful together with detailed examples. However, many replacements are specific to a particular protein and take no account of which parts of the moiety to be replaced are most important for activity. Equally, if the moiety to be replaced is not present in the literature then no suggestions are possible.

The variety and availability of computer algorithms to suggest bioisosteric replacements has increased significantly in recent years. Most methods attempt to excise a chosen moiety from a molecule and replace it with a fragment from a fragment database. These fragments are typically scored against the moiety to be replaced using shape or electrostatic measures of similarity, or by using the presence or absence of pharmacophore points.

All of these methods seem to suffer the same problem: they rarely suggest truly novel scaffolds to the experienced medicinal chemist. The reason for this is not clear but the one commonality is that these computational methods, like the literature methods, rely on fragment-to-fragment comparison. In this method, the replacement fragment is scored as an isolated molecule in itself, and not in the context of the final molecule. This is a subtle but critical problem, which means that there is no possibility for the properties of the

**Figure 2:** Comparison of fragment and whole molecule scoring. Each axis shows the position in the respective results file for 3,600 possible bioisosteres for the lactone group in rofecoxib (top left, red). The highlighted region shows bioisoteres that score well using whole molecules but poorly when compared at the fragment level. For example, the known COX2 active series containing the thiazolotriazole replacement for the lactone (top middle, green) appears some 250 places lower in the results using fragment only scoring compared with whole molecule scoring




fragment to influence the properties of the final molecule. Moreover, as the final molecular context of the fragment is not considered, fragments that might represent only a small change to the final molecule in its entirety may be scored poorly because they represent a large change when scored at the fragment level.

## NEW APPROACHES

In this journal in 2007, we outlined the Field Point approach to describing molecules (1). We showed how, by encoding the electrostatic environment surrounding a ligand, it could be viewed in a more insightful and informative way. We showed how distilling the full field down to a series of 'hot spots' around the molecule – termed 'Field Points' – provided both a powerful insight into the behaviour of molecules and a mechanism by which molecules could be compared in a computationally efficient and therefore rapid manner.

Field Point descriptions of molecules have been used widely to provide richer, more informative views of the way in which ligands interact with proteins, to inter-relate compounds from different chemical series that act at the same protein site, to find novel chemical series through virtual screening, and to decode Structure Activity Relationships (SAR) by comparing molecules as proteins 'see' them (2). Field Point technology can also provide a much more accurate basis on which to identify novel, chemically relevant bioisosteres.

### Application of Molecular Fields to Bioisostere Finding

The principle behind fragment replacement methods to identify bioisosteres is simple: remove a portion of an active molecule, search a fragment database for a replacement moiety that will physically fit into the vacated space, and score the replacement for similarity to the original. In practice, a number of factors contribute to the effectiveness of the method. Primary amongst these are the accurate scoring of potential replacements, the relevance of the fragment database, and the originality and synthesisability of the suggested bioisosteres.

In scoring replacement fragments, it is essential to remember that the molecular fields around them are a property of the whole molecule and not of the isolated fragment. Replacement fragments cannot be assessed accurately in isolation from the whole molecule as they can have a significant effect on the retained portions of the molecule. Not only will fragments that are strongly electron donating have different effects from ones that are electron withdrawing, but the context of the molecule into which they are placed will determine the extent and character of those effects.

To this end, we chose to join replacement fragments into the retained portions of the target, minimising the energy of the result to ensure sensible geometry, before scoring the whole of the proposed new molecule against the original molecule using Field Points. This approach, using the fields of the whole molecule, is only tractable because of the significant computational advantages provided by Field Point representation.

Because the score is based on the whole molecule, any effects that the new fragment may have on the original molecule are automatically considered. This process gives a results list that is significantly richer in non-obvious bioisoteres than would be the case had we only considered the isolated fragment (see Figure 2). Using the whole molecule has an additional benefit in that the medicinal chemist is presented with a list of potentially active molecules rather than partial fragments. This allows them to select molecules for synthesis more easily without mentally having to construct and retrosynthesise the final molecule.

### Choosing Fragments

Commercial screening collections provide molecules with good physicochemical properties, a high degree of synthetic accessibility and reasonable diversity.

**Figure 3:** The results of a FieldStere experiment to find bioisosteres for the lactone portion of rofecoxib (left, red). They include 'obvious' bioisosteres (Parecoxib and Valdecoxib analogues, dark blue, top), 'less obvious' bioisosteres (Etoricoxib and Celecoxib analogues, middle), and non-obvious bioisosteres such as the 12nM inhibitor triazolotriazole (green, bottom)

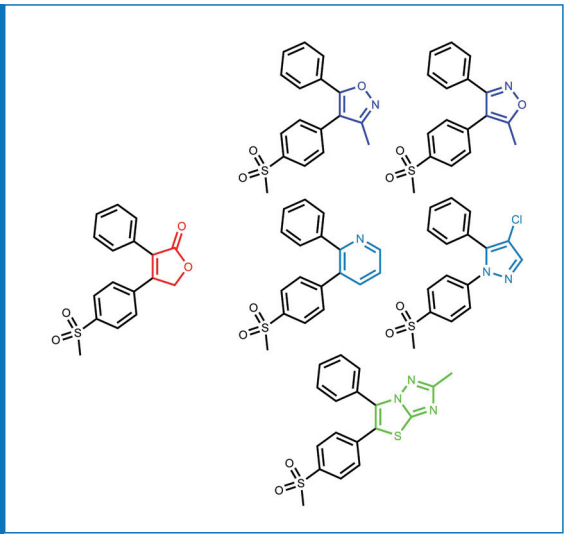
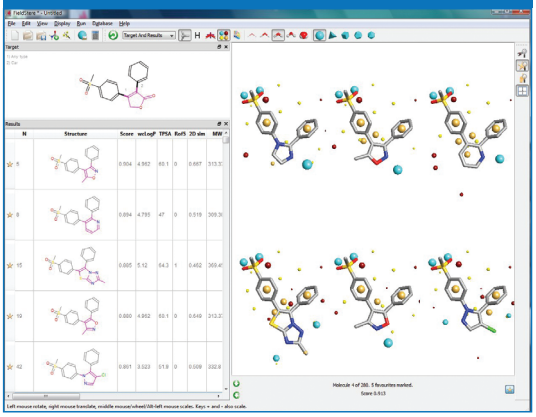


**Figure 4:** FieldStere screenshot showing the COX-2 results. Each bioisostere is presented as a 3D alignment to the original target and also in a table as a 2D picture together with calculated LogP and TPSA values



Fragmenting 2 million molecules by breaking bonds to heteroatoms and rings resulted in a database of 700,000 fragments. To aid the medicinal chemist further, we subdivided the database according to the frequency with which fragments are observed, arguing that more frequently occurring fragments were more synthetically accessible. Since Field Point patterns around fragments are dependent on the exact conformation of the fragment, each fragment is stored as a set of conformations, and the angles and distances between connection points in a fragment are recorded for each conformation that is stored.

It is important to have control over the nature of the connection points in the replacement fragment to reflect the types of chemistry that can be performed on the target molecules. To this end, each fragment is annotated with the nature of the connection points available. The user can simply specify that the connection point is, for example, an aromatic carbon to significantly restrict the search and chemistry space.

The originality and synthesisability of compounds produced using the Field Point method is best demonstrated through an example. Using rofecoxib as a starting point, we requested bioisosteres for the lactone moiety. We limited the chemistry space that was searched by specifying the replacement contain an aromatic carbon attached to the remaining phenyl group. The results (see Figure 3) contain an impressive range of bioisosteres from the 'obvious' – such as Valdecoxib and Parecoxib – to more interesting actives, such as an analogue of Etoricoxib. However, the results also contain 'non-obvious' bioisosteres such as the 12nM thiazolotriazole compound shown.

## CONCLUSION

The use of Field Points represents a major advance in the computational methods available for finding novel bioisosteres. Unlike previous methods, it scores the suggested bioisosteres in the context of the final molecule rather than as an isolated fragment. The results are significantly more diverse than previous methods with both known bioisosteres and novel replacements being suggested in most cases. Application of Field Point based bioisostere technology to drug discovery projects has already been shown to shorten development time and increase novelty.

### References


1. **Vinter A and Rose S, Molecular Field Technology and its Applications in Drug Discovery, *Innovations in Pharmaceutical Technology*, 23, pp14-18, 2007**
2. **Bellenie BR, Barton NP, Emmons AJ, et al, Discovery and optimization of highly ligand-efficient oxytocin receptor antagonists using structure-based drug, *Bioorganic and Medicinal Chemistry Letters*, 19, pp990-994, 2009**
3. **Dart MJ, Wasicak JT, Ryther KB, et al, Structural aspects of high affinity ligands for the a4b2 neuronal nicotinic receptor, *Pharmaceutica Acta Hevetiae*, 74, pp115, 2000**



**Tim Cheeseright** is currently Director, Products, at Cresset BioMolecular Discovery Ltd. He trained as a synthetic chemist at the University of Oxford, and then undertook post-doctoral studies at the University of Manchester and the Institute of Cancer Research before joining the Combinatorial Chemistry Group of Peptide Therapeutics Ltd (later Acambis plc). Tim moved into computational chemistry following the acquisition of Peptide's small molecule business by Medivir AB. He joined Cresset as Chief Application Scientist in 2002. Email: tim@cresset-bmd.com