Beyond Markush – Protecting Activity not Chemical Structure

By Steve Gardner and Andy Vinter at Cresset-BMD Ltd

By providing a more accurate and useful representation of the activity and properties of compounds, field-based systems have the potential for widespread use in patent applications, taking their place alongside, and possibly in time replacing, Markush structures.

The use of Markush structures to protect chemical series in patent applications is based on a convenient but obviously false premise – that combinations of different groups of substituents around a common core generate molecules that have the same activity and biological properties. The use of Markush structures can leave companies unprotected when structurally diverse compounds with the same activity (bioisosteres) are identified. New field-based methods allow the routine detection of such bioisosteres and can be used to evaluate patent positions, strengthen new filings and choose innovative chemistry to overcome existing chemical patents.

BACKGROUND

In the 1920s, chemists seeking to protect their inventions wished to find a way to avoid having to patent individually each member of a class of compounds that would have a similar function. A number of chemists had begun to draft claims in patent filings using structural notations that described a small number of structurally related compounds. These claims were routinely rejected until the landmark appeal of Eugene Markush was upheld by the US Patent Office (1). Markush was awarded a patent for "The process for manufacture of dyes which comprises coupling with a halogen-substituted pyralazone, a diazotized unsulphonated material selected from the group consisting of aniline, homologues of aniline, and halogen substitution products of aniline". Critically, the patent was for processes to produce a range of compounds, including ones that had not actually been synthesised or tested.

From 1925, the USPTO officially sanctioned claims of this type where a 'virtual set' of compounds is considered with R-groups at specified positions selected from a group consisting of a closed listing of potential radical substituents (as shown in Figure 1). A typical Markush claim might be constructed using language such as "an alcohol of the formula R-OH, wherein R is selected from the group consisting of CH3-, CH3CH2- and (CH3)2CH-". Crucially, for the purposes of the claims, each of the potential combinations of substituents is considered to be equivalent and have 'unity of invention'. In other words, in the case of a drug, all of the potential structures are assumed to have the same activity, side effects and other biological properties – something that we know clearly through experience to be untrue.

While Markush structures were a great convenience in 1925, saving each structure from being claimed independently, with improvements in medicinal and combinatorial chemistry it has become routine to have multiple R-groups each with hundreds of defined substituents, generating millions (or billions) of potential compounds. The problems with the inherent lack of specificity of Markush claims have been apparent for many years and Patent Office practice relative to Markush type claims has been reviewed many times since, as claims have become broader and harder to substantiate (2,3). In modern use, compounds defined by Markush claims are directed to a single invention when they

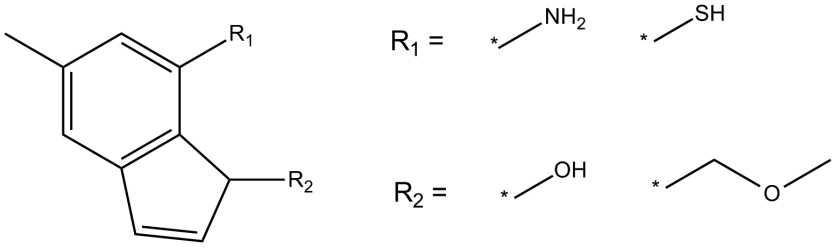

Figure 1: Simple example of Markush structure representing four specific compounds

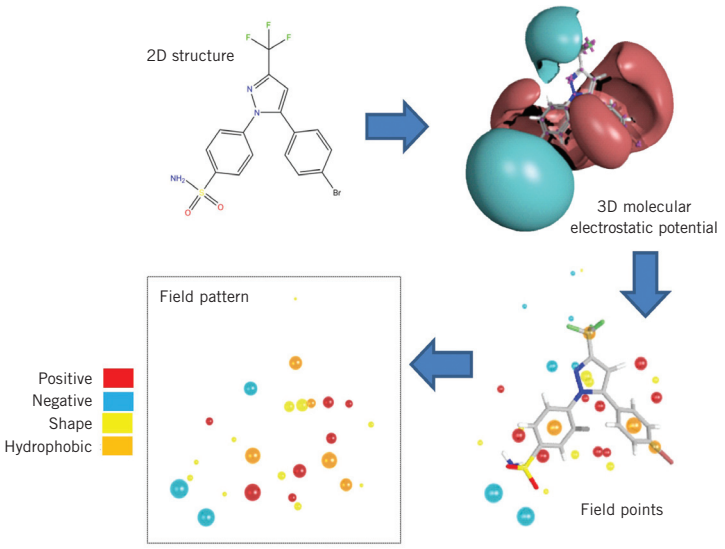

Figure 2: The generation of field patterns

share both common utility and a substantial structural feature essential to that activity.

The most serious problem with Markush-based patenting is, however, much more fundamental – while it provides protection for a region of related chemical structure space, it does not provide meaningful protection for the activity that a drug is targeting. This can be disastrous commercially. There have been many examples over the years where competitors have been able to circumvent a company's patent position by identifying compounds with the same or better activity at the target, whose structures are sufficiently different that they are not covered by the initial patent. This has contributed to a considerable reduction in the market exclusivity time for new therapeutics over the last 15 years (4).

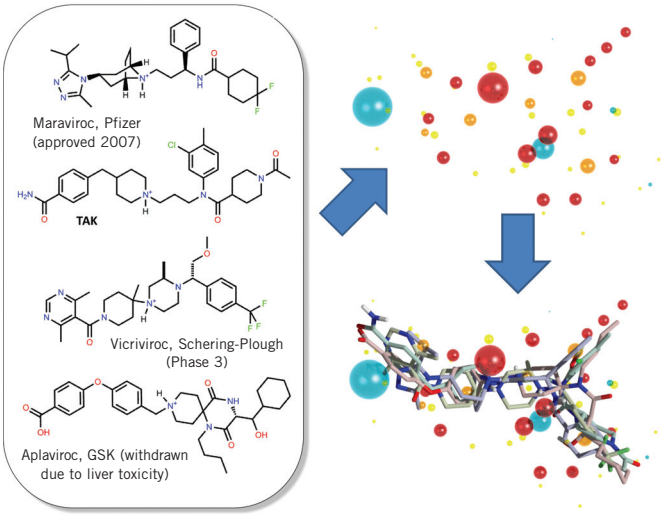

Figure 3: 3D alignment of diverse CCR5 inhibitor compounds sharing the same field patterns and biological activity

STRONGER PROTECTION USING FIELDS

It has been realised since the early days of drug discovery that molecules of different structural types can elicit the same biological action. Such compounds are known as bioisosteres. Moreover, X-ray crystallography has shown that these structurally diverse bioisosteres bind to the common target in the same way. This similarity of activity is known to be independent of 2D molecular structure, because a protein target does not 'see' the atoms and bonds of a drug (its 2D 'skeleton'), but instead interacts with the electron cloud around the molecule and the physicochemical properties on the surface of the compound. The 2D structure, which is the basis of Markush structures, is often a very poor indicator of the likely biological activity and properties of a molecule. This means that Markush structures may be incapable of providing enough protection for a given activity. This has been recognised by a number of groups who have been trying to escape from a simple 2D structure representation (5). We have demonstrated that molecular fields offer significant benefits to researchers seeking to search and evaluate the patent protection around a specific area of activity and write stronger patents (6-8).

Instead of relying on the 2D structure alone, we can now use the fields around molecules to assess their likely activity and properties, regardless of structural similarity. Four molecular fields can be used to describe the properties on the surface of a compound that are the main contributors to molecular interactions between drugs and their protein targets: electrostatic (positive and negative), steric (shape) and hydrophobic ('fat-loving'). The most important regions on the fields (the maxima, where interactions with another molecule are strongest) can then be identified and that portion of the field surface substituted with a Field Point.

As shown in Figure 2, Field Points provide a highly condensed but accurate representation of the nature, size and location of the critical properties required for binding and instigating a specific therapeutic effect. In Figure 2, the Field pattern (framed) represents what a protein target (or any other molecule) sees and feels of the compound shown at top-left. This pattern contains no information about the structure (bonds and angles) that generated it, and in fact many different structures could potentially generate a similar pattern. Crucially, any molecule that can present that same configuration of Field Points is

likely to have the same activity (subject to being able to fit into the active site).

Fields can be used to discover bioisosteres with diverse chemotypes and to explain their common biological activity and properties. For example, the four chemically diverse CCR5 inhibitor molecules shown in Figure 3 have been discovered by four different companies to act at the same biological site and have the same therapeutic effect. All four have very divergent structures, and the CCR5 inhibition activity has obviously been impossible to protect in any meaningful way using Markush claims. The structures do, however, show very similar field patterns in their bioactive conformation, and when clustered on the basis of field patterns, their functional similarity becomes more obvious.

There are many commercially valuable questions that can be answered once we have the ability to go beyond Markush to identify structurally diverse bioisosteres to a known active compound. How much protection does a given Markush structure provide around a given activity? How many Markush structures are needed to protect an activity fully? How inventive is a new claim? Do we already own compounds likely to have similar activity? Can we find new, unprotected chemistry for a given target?

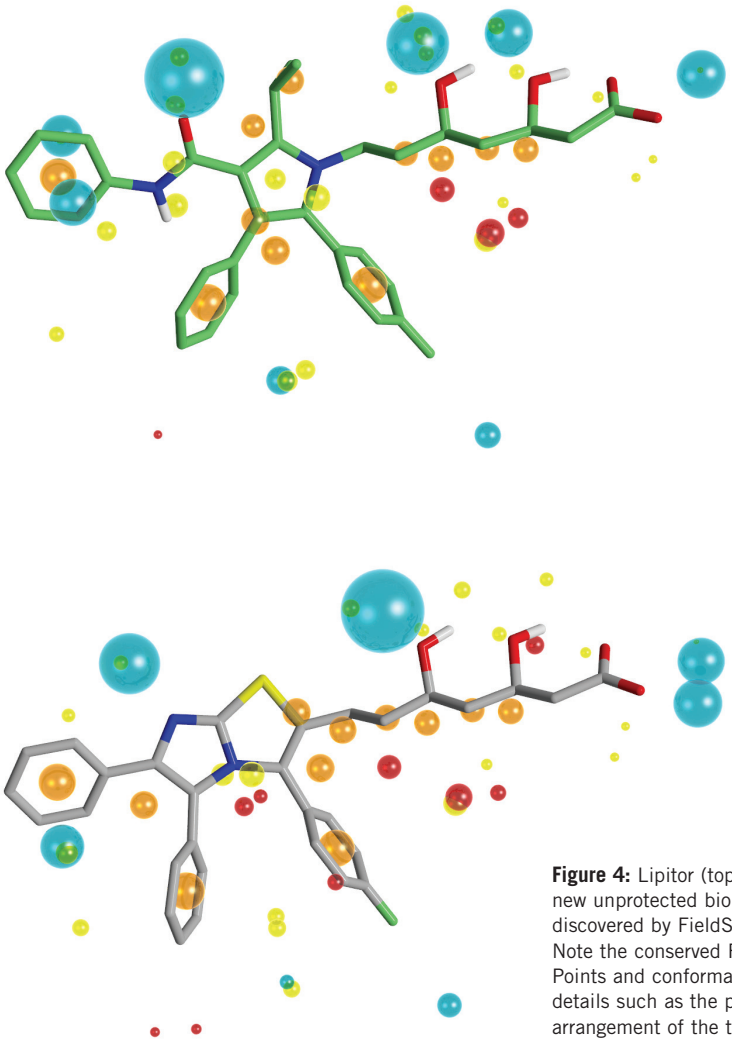

Figure 4: Lipitor (top) and new unprotected bioisostere discovered by FieldStere. Note the conserved Field Points and conformational details such as the propeller arrangement of the three outer ring groups

Identification of Novel Bioisosteres

An example of the potential value of the Fields-based approach is shown in Figure 4 where a bioisostere to Lipitor (atorvastatin, $13.6 billion revenue in 2006), which is not covered by Pfizer's patents, was discovered by Cresset's FieldStere package. In this study, the central ring of the Lipitor structure was substituted with fragments from a database of known structures. Fields were then generated around the novel molecules and these were scored for similarity to the Lipitor fields. Comparing the fields between whole molecules is more likely to find larger non-obvious replacements that result in novel, unprotected molecules that still share the same activity.

Evaluation of Existing Patent Landscape

One of the challenges of modern drug discovery is the lack of innovation and prevalence of me-too compounds. Projects often start from the screening of a corporate compound collection, which typically contains 10^6 compounds, a tiny fraction of the estimated 10^{48} drug-like chemotypes (9). As different companies' collections tend to be relatively similar to each other, the starting points for drug discovery projects are often closer than might be expected in the huge available chemical space. Therefore it becomes crucial to understand how much of the potentially active chemistry around a target has already been protected. In field terms, this question becomes: how many of the suspected bioisosteres do the protected Markush structures cover, and are there novel unprotected bioisosteres available? For in-licensing, it is important to identify quickly which compounds might be bioisosteric to a potential in-licensing candidate (including any in a company's own corporate collection) and who might have claimed protection on them.

Development of Stronger Patent Filings

When developing a new chemical patent, it would be very useful to understand the likely diversity of chemotypes that will display the specific activity to be protected. In field terms, the question becomes: how many different Markush structures are needed to cover

the important bioisosteres that are predicted for a given activity? A related question that will require deeper consideration now bioisosteres can be routinely identified is how obvious a new claim might be if it is predicted to be a highly-scoring bioisostere to a known active structure.

LOOKING AHEAD – A FUTURE BEYOND MARKUSH

Markush structures have certainly been helpful – they are simple to understand, they can be written with pen and paper, and because they are a closed set, it is relatively easy to say if a structure is covered or not. But 85 years on, their major disadvantages are all too apparent – the premise of their use (equivalence of activity) is obviously flawed, they are subject to misuse and over-claiming, and they are difficult to search and examine. Most importantly to the industry, they offer only limited protection to increasingly expensive chemical inventions, and with the widespread availability of bioisostere detection tools this protection is further undermined.

Field patterns on the other hand are very specific to a given target and activity, and any compound that matches a given field pattern sufficiently closely can reasonably be expected to have the desired activity. The disadvantages of field patterns are that the measure is a similarity score and therefore not a simple in/out test, and the 3D field pattern needs software tools to be viewed.

Interestingly, there is a precedent for the filing of a field pattern to protect a specific activity. In 1996 Xenova Ltd filed a patent application (GB2317030A) entitled 'Defining a pharmacophore for the design of MDR (Multi-Drug Resistance) modulators'. The main claims of this patent were based around the definition of a field pattern as described above and its use to distinguish molecules that were likely to possess MDR modulatory activity. This filing was eventually abandoned as the company moved away from MDR as a research area, but demonstrated the principle of the technique. Field-based systems have developed significantly since 1996 and have now evolved into mainstream, desktop PC tools with the potential for widespread use in patent applications.

In the future, as we approach the centenary of the first Markush structure, we can look forward hopefully to a more accurate and useful representation of the activity and properties of compounds that we wish to protect, with field patterns taking their place alongside – and possibly in time replacing – Markush structures.

References

1. **USPTO, Federal Register 72, 154 44,992-45,001**
2. **USPTO, MPEP 803.2 Markush Claims [R-5] – 800 Restriction in Applications Filed Under 35 USC111; Double Patenting**
3. **In re Harnisch, 631 F.2d 716, 206 USPQ 300 (CCPA 1980)**
4. **The Pharmaceutical Market Outlook to 2010, Business Insights**
5. **Fliri A, Development of Technology for Transforming Analysis of Patent Information, Presentation at The International Conference on Trends for Scientific Information Professionals, Nice, October 2008**
6. **Cheeseright T, Mackey M, Rose S and Vinter JG, Molecular field extrema as descriptors of biological activity: Definition and validation, *J Chem Inf Model* 46(2), pp665-676, 2006**
7. **Rose S and Vinter A, Molecular Field Technology and its Applications in Drug Discovery, *Innovations in Pharmaceutical Technology* 23: pp14-18, 2007**
8. **Cheeseright T, The Identification of Bioisosteres as Drug Development Candidates, *Innovations in Pharmaceutical Technology* 28: pp22-26, 2009**
9. **Bohacek RS, McMartin C and Guida WC, The art and practice of structure-based drug design: a molecular modeling perspective, *Med Res Rev* 16(1): pp3-50, 1996**

Steve Gardner, PhD, is Chief Operating Officer of Cresset-BMD Ltd. He has built a number of innovative life science informatics products and companies. Previously Chief Technology Officer of BioWisdom, Chief Executive Officer of Synomics and Astra's Director of Research Informatics, he has designed and deployed enterprise informatics and computational chemistry/biology applications in many biopharma companies. In all, Steve has worked on over 25 projects in safety, pharmacovigilance, in-licensing, biobanking and electronic patient records. He consults and publishes widely, and has several patents in semantic data integration. Email: steve@cresset-bmd.com

Andy Vinter, PhD, CChem, FRSC, is Chief Scientific Officer of Cresset-BMD Ltd. From 1964 to 1990, he worked in the pharmaceutical industry as an organic and computational chemist. In 1990, he left his position as Director of Molecular Modelling and Head of IT at SmithKline & French to pursue his research into the Fields technology. Andy was a Research Associate in the Chemistry Department of the University of Cambridge and UCL, and a consultant for the industry. His research culminated in the founding of Cresset with a grant from the Wellcome Trust in 2001. Email: a.vinter@cresset-bmd.com