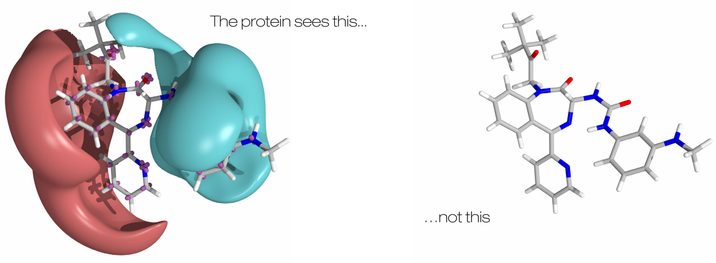
New Leads for GPCR Projects: A Real Breakthrough in Virtual Screening.

T.J. Cheeseright, J.G. Vinter, M.D. Mackey
Cresset BioMolecular Discovery Ltd, Spirella Building, Bridge Road, Letchworth SG6 4ET, UK

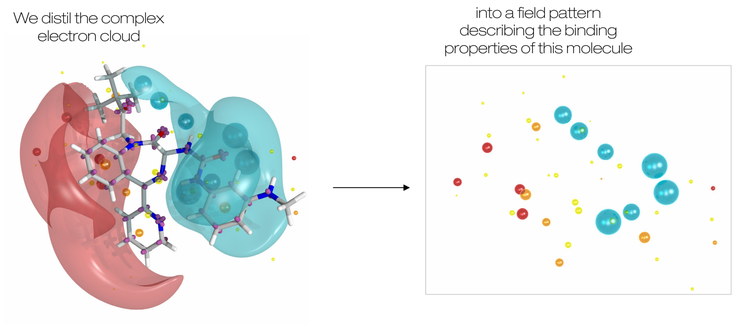
Cresset
BioMolecular Discovery Limited

Introduction

Proteins respond to the **electron cloud** round a molecule rather than to the 3D arrangement of its individual atoms. If we want to describe how a molecule appears to a protein, we need to describe it as a set of surface properties, not as a collection of atoms and bonds. A good description of these steric and electrostatic properties is vital if we are to understand activity.



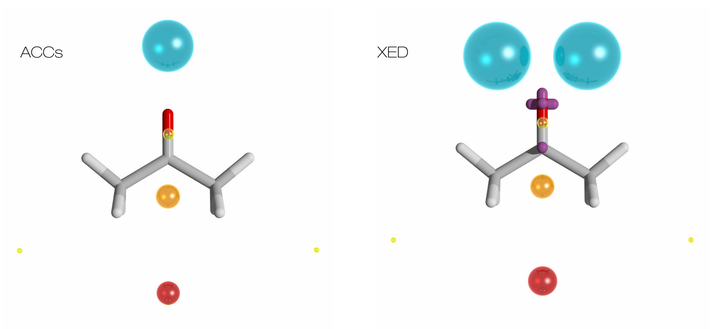
A full surface description of a molecule over all of its accessible conformations is too complex to handle. We solve this problem by condensing the complex three-dimensional fields down to their local extrema, or 'field points'.



Two molecules with different structures but similar biological activities present similar electron clouds to their common target. As a result, they have similar sets of field points. This means that field patterns can be used to align molecules, to score active molecules and to search through databases of compounds looking for potential hits. As the pattern is not directly related to the 2D connectivity of the molecule, but rather its 3D properties, fields can be used to compare molecules from completely different structural classes.

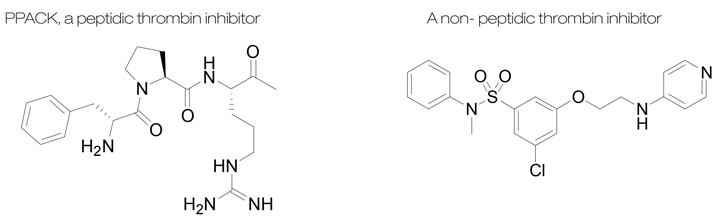
Getting the fields right

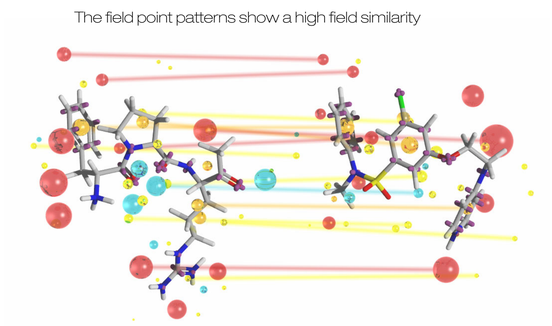
Without a good description of electrostatics, the field points are incorrect. Most standard molecular mechanics force fields use the Atom-Centred Charge (ACC) approximation, which assumes that the charge distribution around each atom is spherical. The **XED** force field (Vinter, 1994) explicitly models the electron distribution around atoms rather than placing charge at the centre of atoms. This significantly increases the accuracy of the Molecular Electrostatic Potential (MEP), giving QM-like electrostatic potentials at MM costs.



Field similarities

Field point patterns can be used to align molecules in three dimensions and to score these alignments in terms of the overall similarity of the two sets of molecular fields in a given orientation.

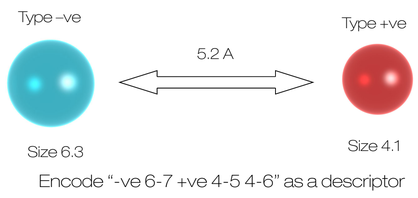




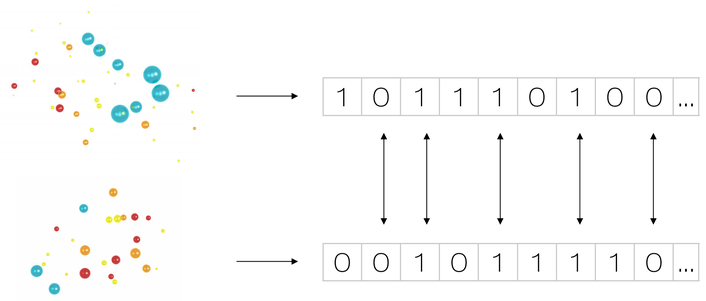
The specificity of field point patterns combined with the fact that they do not map 1-to-1 with functional groups means that meaningful comparisons can be made between peptides and non-peptides. Field overlays have been found to be extremely useful in SAR analysis and the design of new bioactive molecules (McDonald et al. 2000).

A fast field-based search method

Comparison of fields patterns in three dimensions is computer-intensive. A faster comparison can be made by computing the distance matrix for all field points in a conformation of a compound. Each pair of field points (their types, their sizes and the distance between them) then represents a descriptor for that molecule.

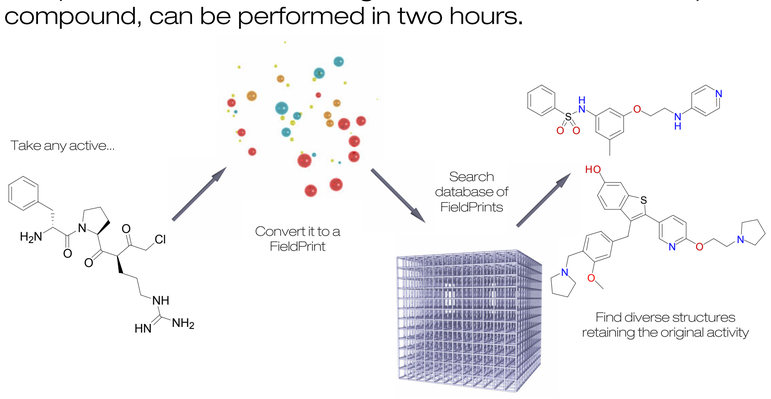


The full set of these descriptors, called a FieldPrint, encodes the three-dimensional field pattern in a one-dimensional vector. A similarity metric can then be defined across these vectors based on how many field pair descriptors two molecules share.



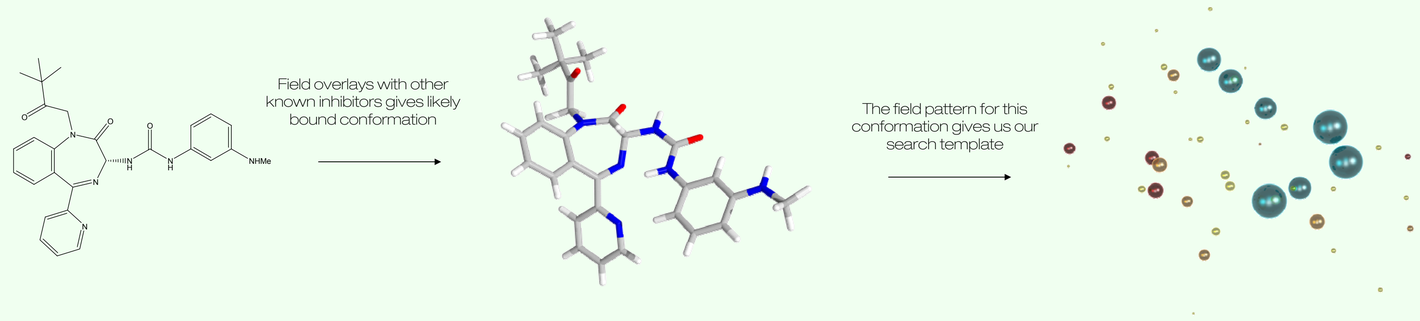
Note that FieldPrints are specific to conformations, so that the database we search against must be pre- populated from a conformational search. We have developed our own fast conformation generator which performs a full conformational search in 30s for an average drug-sized molecule. Every conformation generated is minimised using the XED force field. Our conformation generator performs well when compared to other methods (cf. Bostrom 2001).

Using FieldPrints, a search over a database of 2.5 million compounds, with an average of 35 conformations per compound, can be performed in two hours.

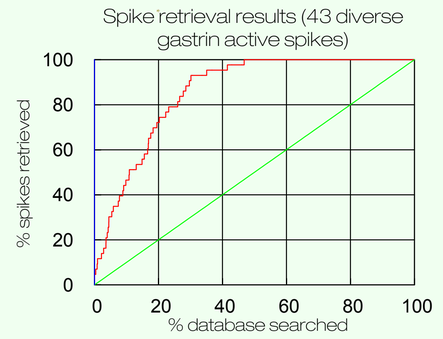


Case Study in collaboration with the James Black Foundation

A validation exercise was performed in collaboration with the James Black Foundation (supported by Johnson & Johnson). The target was gastrin (CCK 2), a peptide GPCR. No crystal structure was available. The likely bound conformation of a known antagonist (a Yamanouchi compound, below) was determined by field-point-based overlays with two other known structurally diverse antagonists.



This template was then used to search through a database of 600,000 commercially-available compounds. Each compound had been subjected to a full conformational search and up to 50 diverse conformations were kept for each. The initial search ranked the whole database according to FieldPrint similarity to the search molecule: the top 50,000 compounds were then refined by performing full 3D field overlays. The top 1000 compounds were then filtered for druglikeness and chemical tractability, giving a final selection of 100 compounds (0.017% of the database). 88 of these compounds were purchased and tested. The activity criterion set by JBF for a compound to be 'interesting' was 10uM.



600,000 compounds
↓ FieldPrint
50,000 compounds
↓ 3D Field overlays
1,000 compounds
↓ Selection for druglikeness
100 compounds
↓ Actually available from vendors
88 compounds

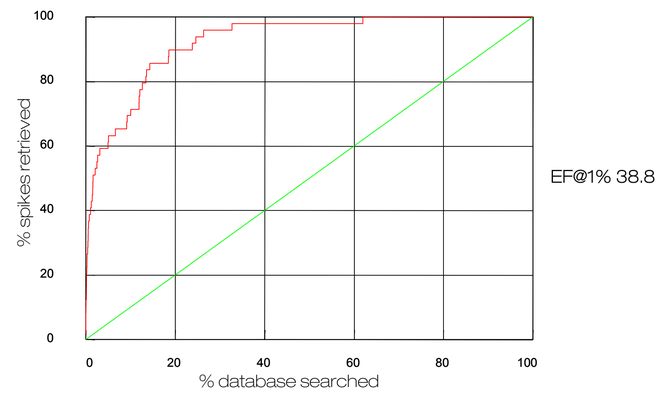
⇒ 30% hit rate

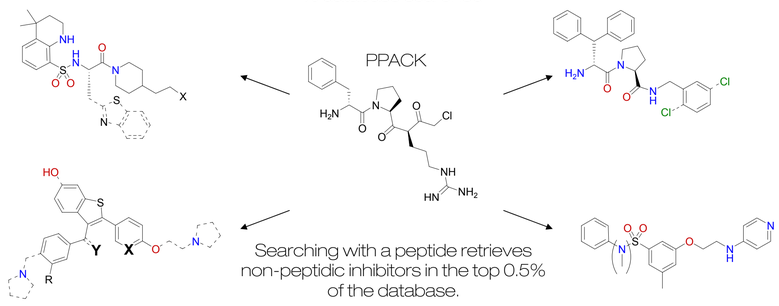
- 27 out of 88 compounds were active (better than 10uM)
- 4 of these were highly active (better than 1uM)
- Most of the hits had no structural similarity to any known gastrin inhibitor
- MW of the hits ranged from 350 to 600: the MW of the search molecule was 500.

Validation: spiked data sets

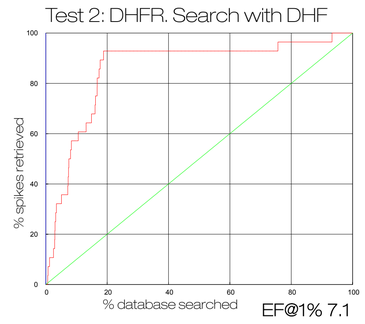
A number of known active molecules from literature sources are 'spiked' or added into a database of commercially-available compounds. We then test how efficiently these spikes are retrieved using the FieldPrint of a known active in its bound conformation as the search template.

Test 1: Thrombin. Search with PPACK (D-Phe-Pro_Arg-chloromethyl ketone)

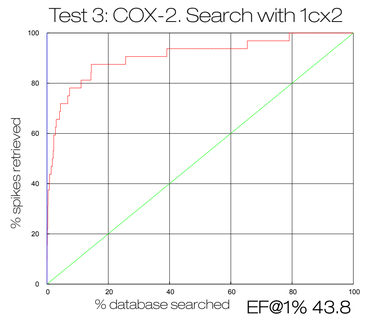




Test 2: DHFR. Search with DHF



Test 3: COX-2. Search with 1cx2



Here, searching with the natural ligand for DHFR retrieves known druglike inhibitors. This demonstrates that the FieldPrint technique could perform well even if no druglike inhibitors were available as search molecules.

Roughly 2/3 of the spikes are retrieved extremely rapidly in the COX-2 search. Examination of the compounds which are retrieved poorly reveals that that all have a significantly different binding mode to the search molecule.

J. G. Vinter, (1994) 'Extended electron distributions applied to the molecular mechanics of intermolecular interactions', J Comp-Aid Mol Design, **8**, 653-668
M. McDonald, D. J. Dunstone, S. B Kalindjian, I. D. Linney, C. M. R. Low, M. J. Pether, K. I. M. Steel, M. J. Tozer and J. G. Vinter. (2000) '2,7-dioxo-2,3,4,5,6,7-hexahydro-1H-benzo[h][1,4]diazonine as a new template for the design of CCK2 receptor antagonists', J. Med. Chem. **43**, 3518-3529
J. Bostrom, (2001) 'Reproducing the conformations of protein-bound ligands: A critical evaluation of several popular conformational searching tools' J Comp-Aid Mol Design **15**, 1137-1152