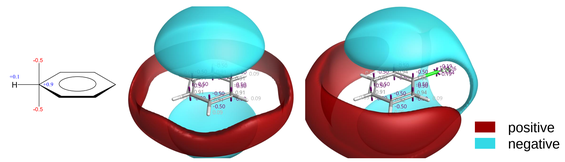
Putting electrostatics and water at the heart of structure-based design

Tim Cheeseright, Mark D. Mackey, Giovanna Tedesco, Paolo Tosco, Susana Tomasio

Cresset, Cambridgeshire, UK tim@cresset-group.com cresset-group.com

XED force field

Cresset's XED force field[1] provides a detailed description of molecular electrostatics through the use of off-atom centre charges. Critical to the XED molecular mechanics approach is the ability to separate partial charges into π- and σ-components.



Calculation of the electrostatic environments of ligands[2] using the XED force field has received significant attention. However, the application of this unique force field to proteins is largely unreported. We were interested to use the XED force field to calculate the electrostatic environment of the active site of a protein. This analysis could prove invaluable for the understanding of ligand binding, SAR and the design of new molecules that target the protein.

Protein interaction potentials

The calculation of protein interaction potentials (aka protein fields) presents many challenges with non-trivial solutions. In particular, detailed attention needs to be given to protein preparation issues, especially residue charge states and hydrogen orientations, the handling of multiple charged residues that are not directly involved in ligand binding but which can swamp the subtleties of the active site electrostatics, and the inclusion or exclusion of water.

The approach we used is similar in principle to the calculation of ligand fields: the protein's active site is flooded with a probe atom and interaction potentials are calculated at each point using a complex dielectric with a distance dependent element. Potentials are then contoured as a surface. Positive and negative surfaces result from the use of a charged probe atom with the van der Waals parameters of an oxygen atom.

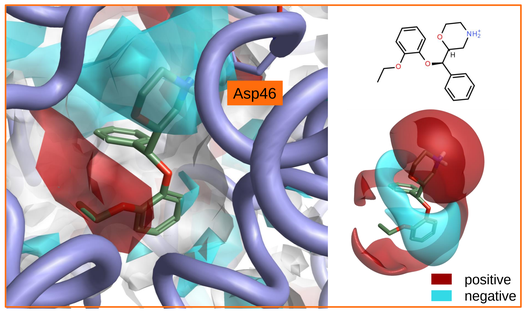
Method

Ligand-protein complexes were downloaded from the Protein Data Bank into Flare™,[3] our new structure-based design application. Ligand and protein structures were carefully prepared using Build Model,[4] and residues lining the active site minimized with the XED force field.

Positive and negative protein interaction potentials were calculated and displayed as iso-surfaces, both for the 'dry' protein structure and including stable crystallographic water molecules according to the results of 3D-RISM analysis as implemented in Flare. Ligand fields were also calculated using the XED force field and compared with the protein interaction potentials of the 'dry' and hydrated' protein active site.

DAT

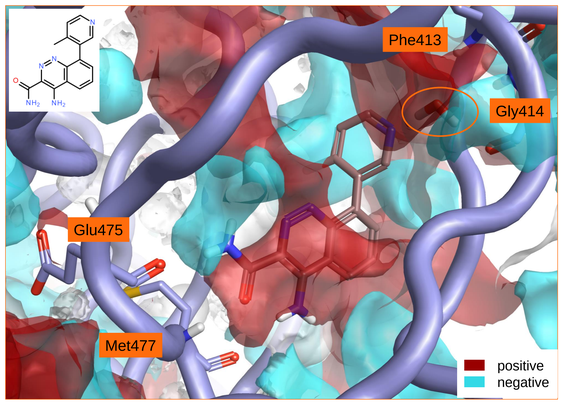
The protein interaction potentials for the dry active site of Drosophila dopamine transporter bound to reboxetine (PDB: 4XNX[5]) shows an excellent match with the ligand electrostatics.



BTK

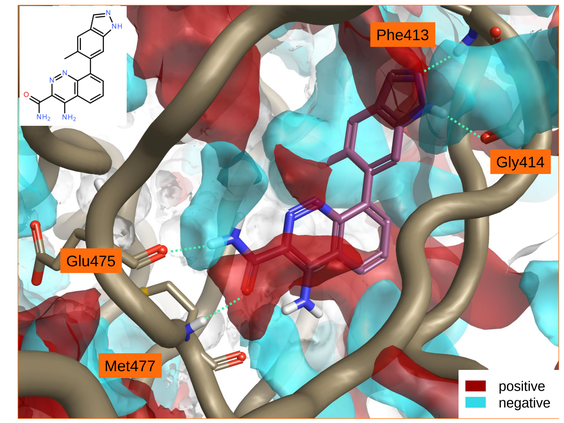
In PDB: 4ZLZ and 4Z3V[6] compounds 8 and 11 interact with the active site of Bruton's Tyrosine Kinase (BTK) by making H-bonds both with the hinge region and P-loop backbone residues. For compound 8, the interaction with the P-loop is mediated by a water molecule.

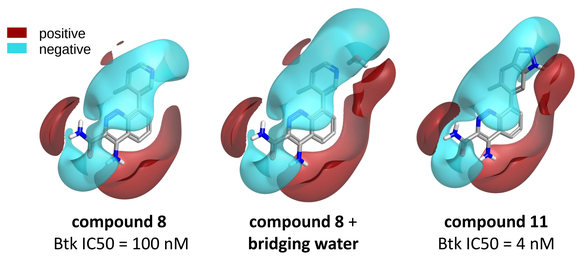
For 4ZLZ, the protein interaction potential of the dry active site shows a good, but imperfect, match of the ligand fields.



The electrostatics of the dry active site of 4Z3V instead match the ligand fields in an a much more precise manner.

The imperfect characterization of the 4ZLZ active site is due to the exclusion of the bridging water molecule from the calculation of protein interaction potentials. It is known from the SAR for this series that substituents mimicking the electrostatics of the combined compound 8 and bridging water molecule (left) lead to compounds such as 11 with improved BTK activity (right).





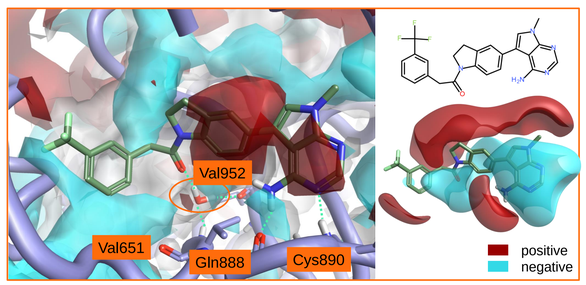
compound 8
Btk IC50 = 100 nM

compound 8 + bridging water

compound 11
Btk IC50 = 4 nM

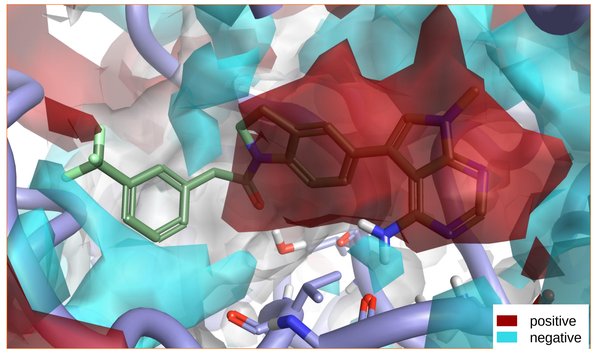
PERK

In PDB: 4G31[7], GSK6414 interacts with the active site of Protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK) by making two direct H-bonds with Gln^{888} and Cys^{890} as well as a water mediated interaction with Val^{952} and Val^{651}.

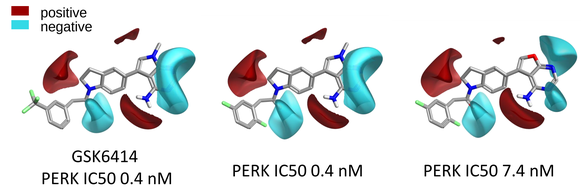


The protein interaction potentials for the dry active site provide a non-optimal description of the protein-ligand electrostatic interactions, with the carboxamide carbonyl pointing to a region of negative electrostatics where the bridging water molecule sits.

A more accurate description is obtained including the stable water molecules according to a 3D-RISM analysis in the calculation of protein interaction potentials.



The SAR for this series is fully consistent with the protein interaction potentials for the PERK active site.



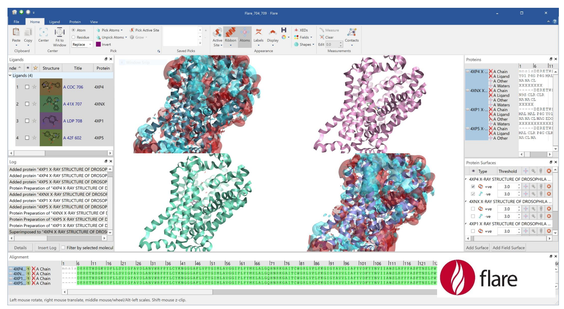
GSK6414
PERK IC50 0.4 nM

PERK IC50 0.4 nM

PERK IC50 7.4 nM

Conclusion

Protein interaction potentials and ligand fields, as implemented in Flare, are a powerful way of understanding the electrostatics of ligand-protein interactions. The inclusion of stable water molecules following a 3D-RISM analysis dramatically improves the precision of the method for the characterization of protein active sites. The information gained from protein interaction potentials can be used to inform ligand design, compare related proteins to identify selectivity opportunities, and understand SAR trends and ligand binding from the protein's perspective.



References

1. *J. Comp. Aided. Mol. Des*. **1994**, 8, 653-668
2. *J. Chem. Inf. Model.* **2006**, 46, 665-676
3. http://www.cresset-group.com/flare
4. *Proteins* **2011**, 79, 2693-2710
5. *Nat.Struct.Mol.Biol*. **2015**, 22, 506-508
6. *J. Med. Chem.* **2015**, 58, 5437-5444
7. *J.Med.Chem.* **2012**, 55, 7193-7207