

# Deciphering Complex SAR

T. Cheeseright<sup>†</sup>

<sup>†</sup> Cresset, New Cambridge House, Bassingbourn Road, Litlington, Cambridgeshire, SG8 0SS, UK

## Abstract

The substitution of aromatic groups provides a unique tool to moderate the potency and physicochemical properties of drug like molecules. However, the huge variety of substitutions that are possible can give rise to SAR that is almost impossible to understand, with small changes resulting in large shifts in potency. In these circumstances the understanding of the causes of the observed activity cliff is critical to progressing the project aims. This is an area where we at Cresset have always felt that using molecular interaction fields gives you a head start as you can model the electrostatic and shape properties of the molecule accurately. The release of the [Activity Miner](#) module for [Forge](#) and [Torch](#) significantly improves this process by detecting automatically activity cliffs in the SAR. Below we present a case study on a small set of changes around a set of reported DPP-IV inhibitors and show how the activity miner interface helps find the root causes of the changes in activity.

## Method

A set of DPP-IV inhibitors related to the ligands from PDB codes 2QOE and 2P8S were extracted from [bindingdb](#) together with IC50 values for enzyme inhibition. Using Forge, PDB 2QOE was downloaded and split into reference ligand and protein. The ligand from PDB code 2P8S was downloaded as a fixed conformation and aligned to the 2QOE reference using the default 'normal' settings then added as an additional reference molecule. The remaining 33 compounds in the dataset were aligned using the 'Substructure' method to these references with the maximum score against any reference being used to choose the alignment. The resulting alignments are shown in figure 1.

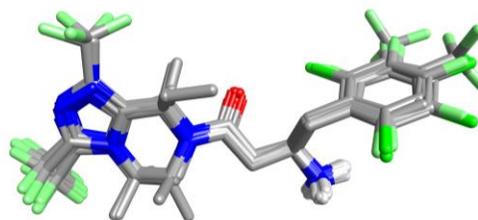
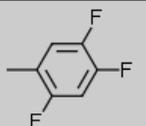
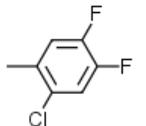
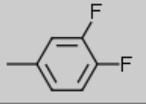
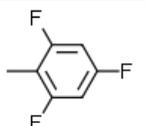
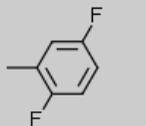
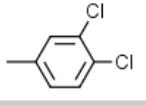
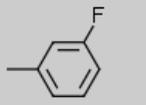
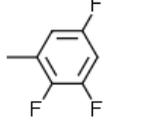
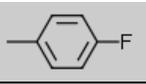


Figure 1. Aligned structures.

The aligned dataset was transferred to the Activity Miner module to study the SAR around the terminal phenyl substituent. The 'activity view' focused on the most active compound (shown below) highlights that the SAR around this substituent is complicated with many small changes resulting in significantly worse IC50 values. The activity view presents a central (focus) molecule, with the most similar molecules to the focus compound displayed in a wheel around it. The size of the segment represents the distance between the two molecules and the segment is colored by the disparity between the pair. Highly colored segments represent changes that result in disproportionately high changes in activity (color: red is worse activity, green is better), figure 2.



Table 1. Classic SAR table.

Row	Phenyl Substitution	Activity (pIC50)
1		8.2
2		7.1
3		6.9
4		7.1
5		7.6
6		5.8
7		6.9
8		6.1
9		6.6

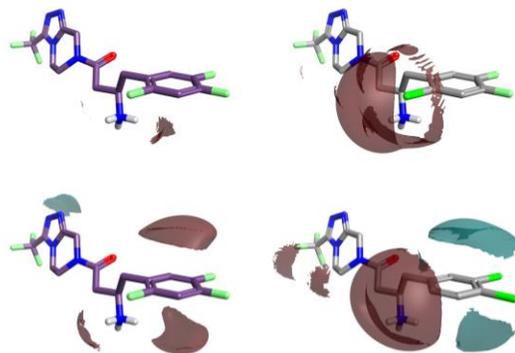


Figure 3. Field difference of row 1 of table 1 with row 2 (top) and row 3 (bottom).

### 2. The 4 position prefers negativity at the end

Removing the 4-F from row 1 gives row 5. Moving the fluorine atom in this position round the ring one position gives row 8 (figure 4). In both cases the activity is reduced by the change. The smaller change in activity when going from F→H suggests that introducing a negative region in the 3 position is additionally unfavorable. Neither of these hypotheses are obvious from the protein crystal structure where both the 3 and 4 positions interact with a number of residues of various types.

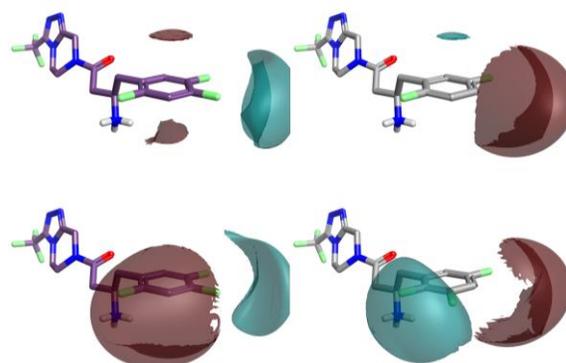


Figure 4. Field difference of row 1 of table 1 with row 5 (top) and row 8 (bottom).

### 3. The 5 position must be negative at the end

All the changes that remove the negativity from the end of the 5 position result in significant drops in activity whilst those that retain the negativity, even in the absence of other

favorable interactions retain some activity. For example row 4 has both the 2 and 4 fluoro atoms but is only pIC50 7.1.

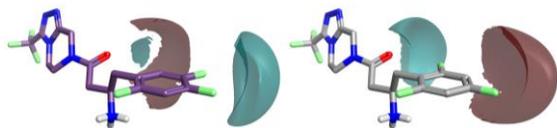


Figure 5. Field difference of row 1 of table 1 with row 4.

The reason for this becomes evident on examination of the protein crystal structure. This atom points directly at the edges of the indole from Trp<sup>659</sup> and the phenyl of Tyr<sup>670</sup> (figure 6, numbers from PDB 2QOE).

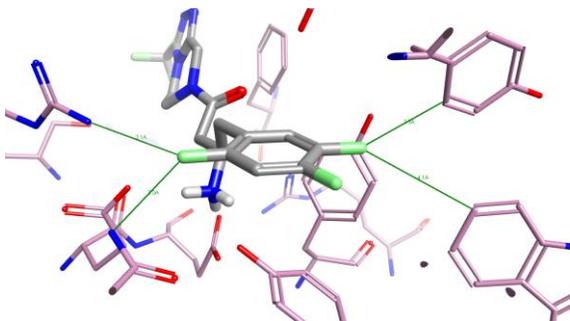


Figure 6. Protein interactions of 2- and 5-fluoro substituents.

#### 4. The electron density of the phenyl substituent is important

This hypothesis is harder to establish as it comes from many observations. The most obvious is the change from row 3 to row 6 (figure 7) where there is a drop in activity from pIC50 6.9 to 5.8. Clearly this could be due to the increased size of the chlorine atoms but equally likely is the change in the electronic properties of the phenyl ring where highly electron poor rings have higher activity.



Figure 7. Field difference of row 3 of table 1 with row 6.

This change is also observed where any of the fluorines of row 1 are deleted or where any atom is switched from fluorine to chlorine. Again the protein crystal structure helps to validate this hypothesis as the catalytic serine together with a couple of tyrosine residues point their respective alcohol oxygen atoms at the face of this ring (figure 8).

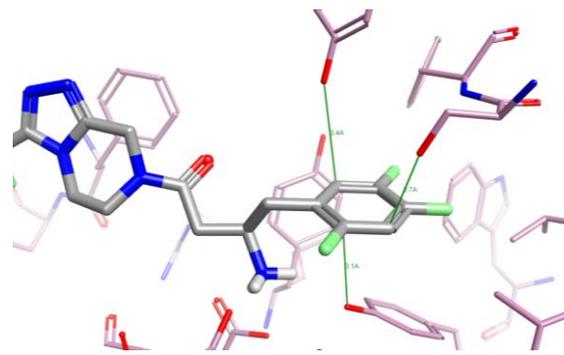


Figure 8. Protein interactions of the face of the phenyl substituent.

## Conclusion

Many of our hypotheses could have been guessed at from studying the crystal structure of the 2,3,5-tri-fluorophenyl analogue in detail. However the use of the field difference mode in activity miner brings the interactions into sharp focus and helps us rationalize the observations that we have. Subtle effects such as the difference between electron-rich aromatic and electron-poor aromatic rings are clearly visualized, explaining difficult and complex SAR in a way that is easy to interpret.

Our hypotheses can now be used in the design of new ligands with better IP or

physicochemical properties with each design being validated against the regions of positive or negative field that we conclude to be important. Equally we could look for new ideas for this section of the molecule by using Spark together with the new reagent databases to suggest compounds that are readily synthesizable and that would retain the activity we have in this series while driving us into new regions of chemical space.

## References

1. <http://cresset-group.com/activityminer>