Fragment hopping with Blaze

T Cheeseright[†]

[†] Cresset, New Cambridge House, Bassingbourn Road, Litlington, Cambridgeshire, SG8 0SS, UK

Abstract

In this case study we look at the performance of Blaze[1] (our 3D ligand-based virtual screening application) on a fragment data set. Comparing the Blaze default settings with an alternative shape-only screen, we found good retrieval rates of known actives in both cases. The two methods are found to be surprisingly complementary as distinctly different hits are obtained from the different methods. Examination of the results shows that Blaze is capable of finding fragments that are similar to known actives as well as completely novel suggestions. Unlike many virtual screening methods, there is no evidence that the performance of Blaze decreases with smaller search queries.

Introduction

In a recent paper Keserű and co-workers reported[2] the results of a fragment screen for adrenergic α_{2c} agonists. They used a combination of cell based wet screening and docking to a homology model to identify 17 novel hits with varying levels of inhibitory and agonist activity. The published structures are a valuable resource for exploring the performance of virtual screening methods on fragments. Fragment VS is known to be difficult: 2D fingerprints perform poorly on small molecules and there have been many publications on the difficulties of docking fragments compared to drug-sized molecules.

In this study we used one of the Keserű docking results as a query molecule for exploring the retrieval of the other published hits using two ligand-based virtual screening methods that are both available within Blaze, Cresset's VS platform.

Method

The published hits were uploaded to Blaze as a spike set. Where the stereochemistry was unknown we uploaded them with unspecified stereochemistry and relied on Blaze to enumerate chiral centres at the same time as it generated conformation populations. The final collection 'Adrenergic_a2C_fragments' was included in the search and Blaze monitored and reported the retrieval rates and enrichment factors automatically.

Blaze requires an active search query in its bioactive conformation. We wished to use one of the hits obtained from docking as a query to mimic the effect of a combined structure-based and ligand-based screen. Of the two hits, compound 10 is the more active and hence was chosen as the primary search query. The original paper details docked poses for both hits that they obtained but did not provide the coordinates. We were able to approximately reproduce the conformation for compound 10 from the published picture and used this as our query (Figure 1).

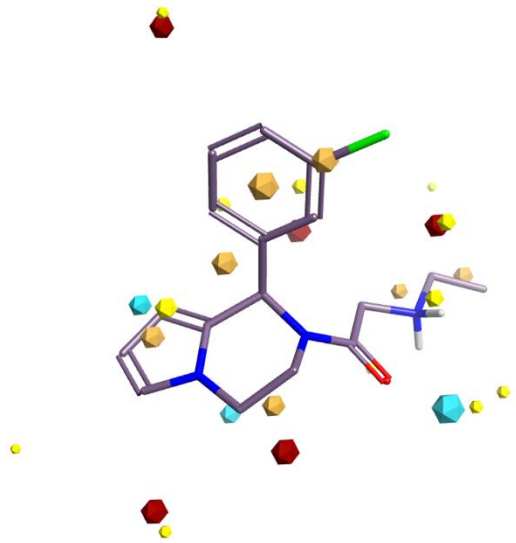


Figure 1.

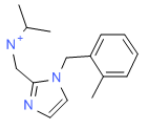
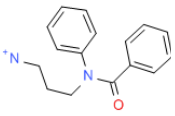
We ran the Blaze experiment on the public Blaze server[3] using Chembl compounds as decoys. We limited the search to compounds with 11 to 20 heavy atoms so as to retrieve fragment-like hits. Furthermore we applied a filter to ensure that all Chembl compounds retrieved contained a positive charge, as this is a dominant feature of the published hits. Note that all spike molecules were allowed to pass the filter even if they were not charged. It is possible that this reduced the retrieval rates as it applies a charge bias between spikes and decoys.

Two search methods were applied. In the first approach the default Blaze conditions were used to score hits. This uses a combination of electrostatic (field) similarity and an explicit shape similarity in equal measure. The second approach used only the shape similarity to score hits. In both cases we applied the highest level of calculation ('simplex') to the complete filtered dataset of 10,680 compounds that lie within the heavy atom limits.

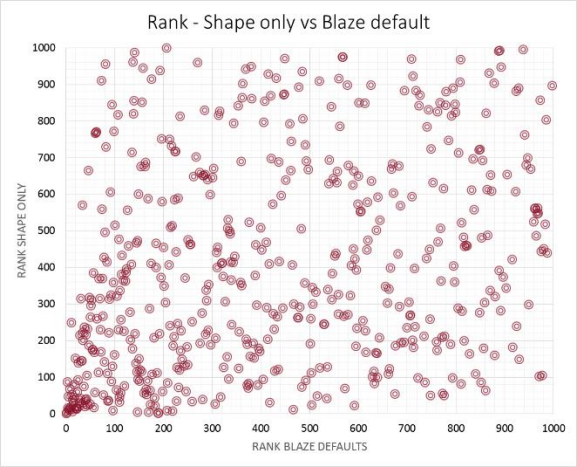
Results Introduction

Performance metrics

The results from each experiment – 'Blaze defaults' and 'Shape only' are summarized in the table below. Note that the query molecule was present in the data set and was retrieved at position 1 in both cases.

Method	ROC AUC	BEDROC20	Spikes in top 5%	First new chemotype
Shape only	0.746	0.245	5	
Blaze defaults	0.795	0.392	8	

On first pass the results from the two methods seem fairly similar, although the Blaze default search gets a slightly better early enrichment. However, as always, the devil is in the details. If we take the rank of each molecule in the dataset and compare them between the two methods we can see that they have a low correlation (Figure 2).



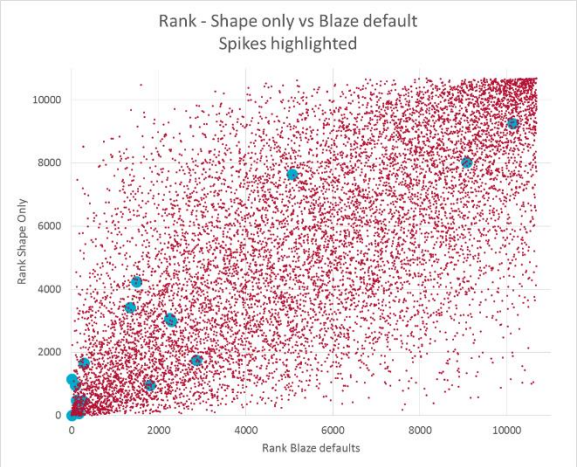


Figure 2: Rank of compounds using Shape only vs Blaze defaults: (a) zoomed to top 1000 results; (b) overall highlighting spike molecules in blue.

Figure 2 shows that the rank for any particular compound can be very different between the two methods. There are a few compounds at the top of the both lists that are retrieved by both methods, but using a typical cutoff of e.g., 1% of the database (106 compounds) then there are large numbers of compounds that are exclusive to each method. This applies to both the spikes and the decoys. Although the Blaze defaults perform better in this particular case, the Shape-only method still finds spikes in the top 2000 compounds that are missed by the Blaze defaults.

Fragment hopping

As well as the known ligands a number of new structures are returned that are pharmacophorically similar yet structurally very different to the known actives. Figure 2 shows some selected molecules from the top 35 results in the Blaze-default search. As can be seen, Blaze finds a very wide variety of different chemotypes with excellent shape and pharmacophoric matches to the query.

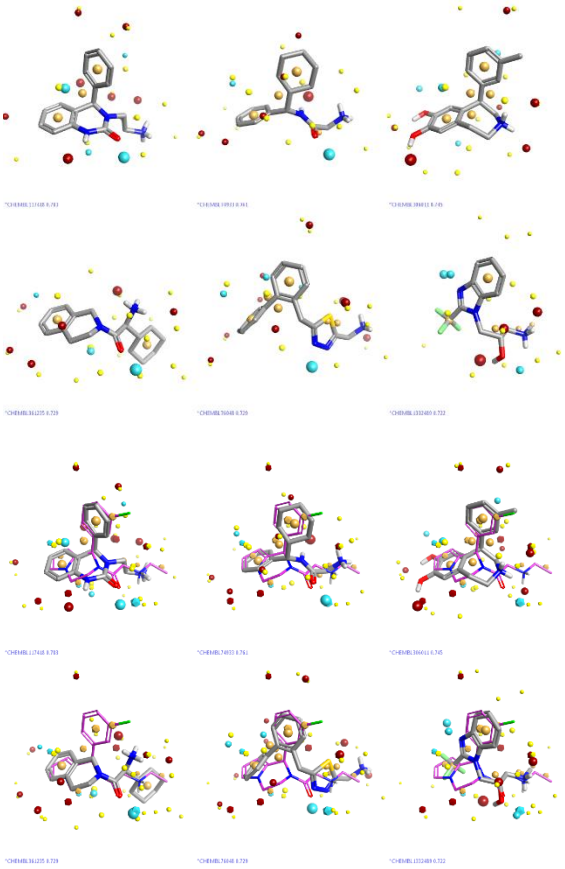


Figure 2 (a). Selected results from Blaze results; top row: CHEML117418 (rank 2); CHEMBL74933 (rank 5); CHEMBL306011 (rank 7); bottom row: CHEMBL361235 (rank 18); CHEMBL76048 (rank 20); CHEMBL1332489 (rank35). (b) The same results as (a), shown overlaid on the query molecule (shown in pink).

The amino acid moiety present in CHEMBL74933 is a common motif in the full list of top-ranked compounds and interestingly is also represented in the other docking hit found by the original authors. In addition, CHEMBL306011 is a close analogue of one of the more active fragments reported in the paper and is found by both methods along with a large number of structurally similar molecules.

Conclusion

This experiment shows that Blaze can produce excellent results for virtual screening of fragments. Indeed the enrichments are similar or greater than those that are obtained on larger molecules. The results also reinforce our contention that the Blaze hit lists are often complementary to those obtained using other methods, such as pure shape.

In this case, once the Keserű group had obtained their first few hits from docking, a Blaze search to expand the hit list around these would have provided a highly cost-effective alternative to performing a large physical screen.

References and Links

1. http://www.cresset-group.com/products/blaze/
2. E. Szőllősi, A. Bobok, L. Kiss, M. Vass, D. Kurkó, S. Kolok, A. Visegrády, G.M. Keserű; Bioorg. Med. Chem. 23 (2015) 3991–3999. doi: 10.1016/j.bmc.2015.01.013.
3. http://www.cresset-group.com/products/blaze/sign-up-for-access-to-blaze-demo-server/