Prioritizing the Most Promising Virtual Screening Hits

Morgan Morris, Jessica Plescia, Sílvia Tuset, Andy Smith, Stuart Firth-Clark
Cresset, Cambridgeshire, UK morgan.morris@cresset-group.com.com

Introduction

Non-receptor tyrosine-protein kinase TYK2 is an enzyme involved in the IFN-α pathway and various interleukin pathways regulating inflammation and immunity. Currently, it is under investigation for its therapeutic modality in diseases such as rheumatoid arthritis and psoriasis.[1]

Much focus has been given to the development of inhibitors which selectively target TYK2 while avoiding and preserving the functions of the structurally similar tyrosine kinases JAK 1, 2 and 3. A workflow incorporating the Blaze™ [2] virtual screening platform (Figure 1) was used to find virtual hits against TYK2.

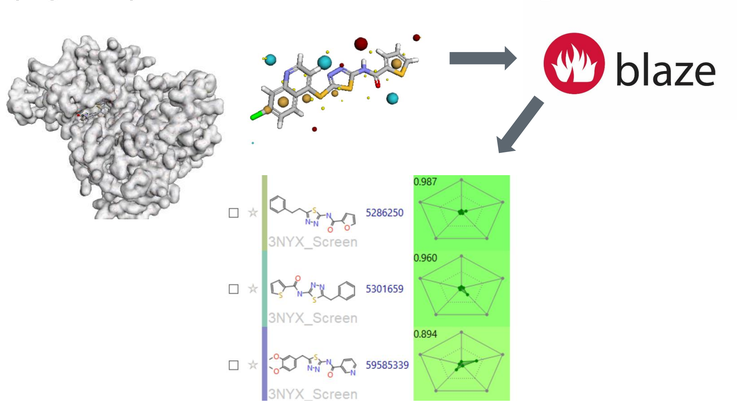


Figure 1: Workflow highlighting the input of TYK2 (PDB = 3NYX) entries with associated ligands into Blaze, followed by data triaging in Flare™ [3].

Methods

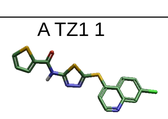
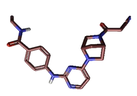
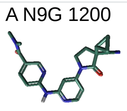
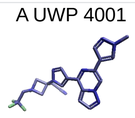
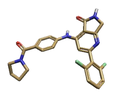
47 TYK2 PDB entries and corresponding co-crystallized ligands were assessed for clustering.

Selections were made based on:

- Stereogenic centers within ligands (<4)
- Resolution (≤ 2.5 Å)
- ≤ 1 activity overlaps for ligands with JAK series

Ligands from the 22 PDB complexes passing initial filters were clustered based on electrostatic and shape features described by the XED force field. The highest bioactivity representative of each cluster was then selected to be used as reference ligands in their validated bioactive conformation (Figure 2) in Blaze virtual screens.

Table 1: Information about crystallographic data associated with this study.

Ligand	PDB Code	Resolution (Å)	Ligand	IC_{50} or K_i (nM)	Stereo-centers	Screening Constraints
1	3NYX	2.50	A TZ1 1 	K_i = 32	0	None
2	6DBK	2.00	A G5D 4000 	IC_{50} = 23	2	Nitrile C≡N
3	6OVA	2.50	A N9G 1200 	IC_{50} = 1.1	1	Nitrile C≡N
4	6X8F	2.15	A UWP 4001 	IC_{50} = 10	0	Azetidinium C3H8N+
5	7UYU	2.05	A OV0 1201 	K_i = 0.51	0	Amide N-H

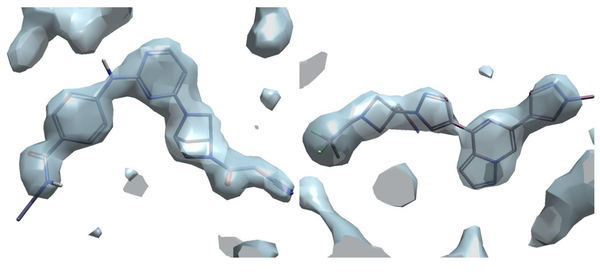


Figure 2: The crystal structure conformations of the ligands were checked against the electron density maps before selecting five diverse reference compounds using the XED force field. Left panel = **2**; Right panel = **4**.

Blaze Screening

Approximately 22 M commercially available compounds were screened using Blaze against 5 structurally diverse representative ligands which bind to the TYK2 JH1 pseudokinase domain,[4] using their associated proteins as excluded volumes. Screened compounds were conformationally sampled and aligned against the reference in each case, and the highest scoring 50% (clique refinement) followed by the highest 10% (simplex refinement) retained in each instance. The similarity score against the reference was weighted to 50% shape and 50% field-based (XED, Figure 3).[5]

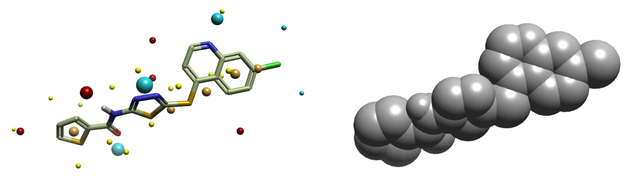


Figure 3: Overview of the similarity scoring in Blaze using ligand **1**. 50% comes from field description and 50% comes from shape. Red = positive; cyan = negative; orange = hydrophobic; yellow = highly accessible van der Waals surface.

To promote essential interactions during virtual screening, additional pharmacophoric constraints were added in Blaze to key functional groups (Table 1).

The top 1000 ranking compounds from each of the five virtual screens were analyzed using a structural filter and problematic compounds removed.

Data Triaging in Flare

Following substructure filtering, molecules containing more than one torsion that is present but with low frequency within the CSD were additionally filtered out.[6] The remaining virtual hits were then moved to the triaging stage.

Filtered molecules were rescored in Flare using the following orthogonal methods: (a) LeadFinder™ docking and scoring against the dry reference protein, and (b) Electrostatic Complementarity™ (EC) between ligand and protein.[8]

A multi-parameter function (known as radial plot function [RPF] in Flare) was then generated using: Docking score, EC scores, Similarity score.

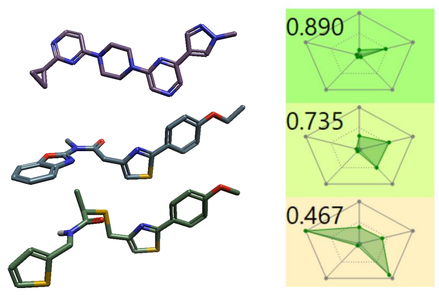


Figure 4: Representative RPF analyses for 3 virtual hits obtained from screening ligand **4** in Blaze. Cresset's Flare allows for custom radial plot design with the user's own parameters.

ROC Analysis

To test the validity of possible RPFs in Flare, 9 known hits closely related to **5** were incorporated into a test database of ca. 1400 compounds which have similar physical and structural properties. The test database was screened, and an ROC curve (Figure 5) was generated.

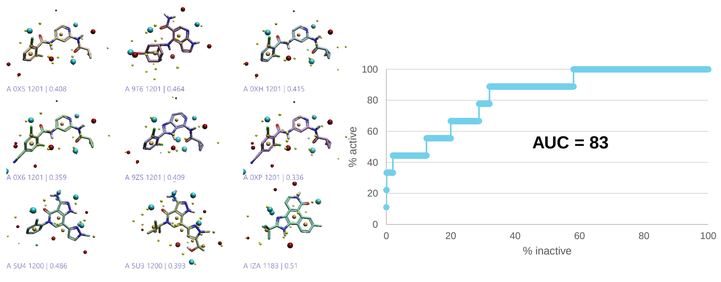


Figure 5: 3D structures of hits within the test database to test possible multi-parameter scoring functions. Data enrichment can be extrapolated through the area under the ROC curve (AUC).

For the selected RPF (a combined docking score, EC scores and similarity score), a threshold of the top 1% (≥ 0.46) was used as a cut-off score as part of the triaging analysis of the virtual hits.

Protein-Ligand Interactions

The virtual hits were further inspected visually after RPF scoring using 2 approaches:

- Ligand-only: a subset of the Watson-Bruns Eli Lilly Rules was used to filter out structures.[9]
- Protein-ligand: Hits were filtered out if they did not make at least one contact with the key hinge-binding residue, Val 981.

A final clustering analysis using FCFP6 was performed on the selected compounds to generate a diverse list of virtual hits for purchase.

Results

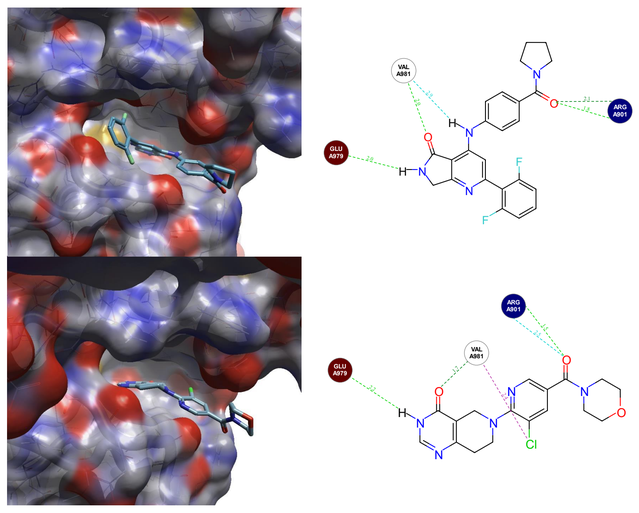


Figure 6: Top panel: Ligand **5** explores the TYK2 hinge region and glycine-rich loop. Bottom panel: A virtual hit found which exploits a novel halogen bond to bind to the hinge region via Val 981. Protein-ligand interactions: light/ dark green = strong H-bonds; light blue = weak H-bond; light purple = halogen bond.

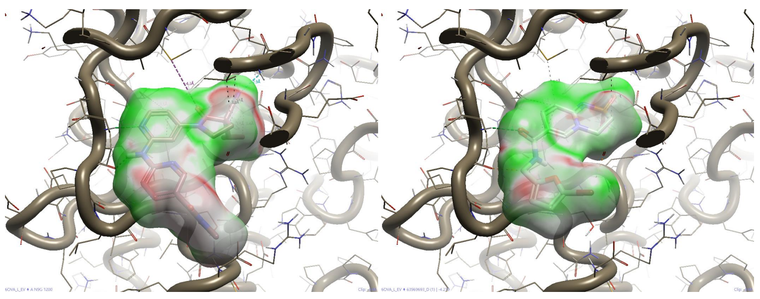


Figure 7: Flare's EC surfaces[8] highlight improved interactions in the bromofuran-containing screening hit (right-hand panel) versus the native ligand **3** (left-hand panel). Red = clashes with protein; green = protein-ligand complementarity; gray = neutral.

Conclusions

Virtual hits were obtained which displayed diverse chemistries and novel interactions using a Blaze-Flare workflow with the TYK2 JH1 pseudokinase domain. While the only way to truly validate a virtual screen is to acquire and biologically test virtual hits, this study illustrates the many successfully validated virtual screens executed by Cresset customers.

References

1. Leit, *et al*., Bioorg. Med. Chem. Lett. **2022**, *73*, 128891-128901
2. Blaze™, version 10.8, Cresset®, Litlington, Cambridgeshire, UK; http://www.cresset-group.com/blaze/
3. Flare™, version 7, Cresset®, Litlington, Cambridgeshire, UK; http://www.cresset-group.com/flare/
4. Tsui, *et al*., Proteins. **2011**, *79*, 393-401
5. Cheeseright, *et al*., J. Chem. Inf. Model. **2006**, *46*, 665–76
6. Penner, *et al.,* J. Chem. Inf. Model. **2022**, *62*, 1644-1653
7. Stroganov, *et al*., J. Chem. Inf. Model. **2008**, *48*, 2371-85
8. Bauer, *et al*., J. Med. Chem. **2019**, *62*, 3036-50
9. Bruns, *et al*., J. Med. Chem. **2012**, *55*, 9763-72