Fragment Oriented Molecular Shape (FOMS) Search: A Novel Shape-Based Virtual Screening Method

Ethan R. Hain[1,2], David R. Koes[3]

[1]*TECBio REU @ Pitt, Dept. of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA 15260*
[2]*Dept. of Chemistry and Biochemistry, St. Mary's College of Maryland, St. Mary's City, MD, 20686*
[3]*Dept. of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA 15260*

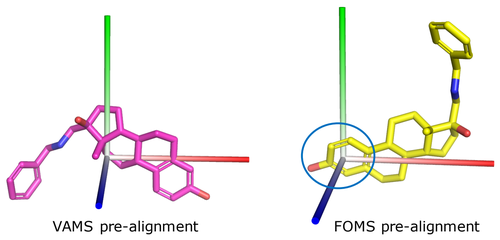
Abstract

Molecular shape is a useful tool for identifying small-molecules for therapeutics. Shape similarity can be calculated from the overlap of aligned molecular shapes. Methods that dynamically optimize the alignment to maximize the volume overlap are computationally intensive. The computational burden of alignment can be reduced by pre-aligning shapes to a single canonical alignment. The recently published Volumetric-Aligned Molecular Shapes (VAMS) method pre-aligns molecules to their moments of inertia and was shown to be effective as a pre-screen. The introduced method, fragment oriented molecular shape (FOMS) search, pre-aligns shapes to only a part of the shape in detail (e.g. a key interacting fragment). This optimizes the overlap of the shapes in the region of the fragment. We show that FOMS outperforms VAMS for most protein-ligand interactions in the challenging MUV dataset and is extremely efficient at using shape constraints to filter the dataset.

Background

Shape-based virtual screens have been used frequently to identify the most similar molecules in a database to a set of one or more active molecules (1).

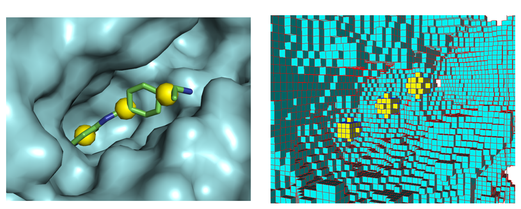
Volumetric aligned molecular shapes (VAMS) search has been shown to be an effective shape-based pre-screen (2). VAMS **pre-aligns** voxelized molecular shapes to a reference system defined by the molecules' moments of inertia, eliminating the need to optimize an alignment as in other methods and resulting in a substantially accelerated search.



VAMS pre-alignment | FOMS pre-alignment

Fragment oriented molecular shapes (FOMS) is an alternative pre-alignment method where molecules are aligned to a predetermined fragment. The user is required to identify the fragment structure's binding mode, and the search space is confined to compounds that contain only the specified fragment.

Shape constraints are filters that exactly specify a minimum and maximum volume. The minimum volume delineates where the desired ligand must be and the maximum volume prevents it from clashing with the receptor. When shapes are aligned, as with VAMS, indexing data structures (3) can be used resulting in sub-linear time search.

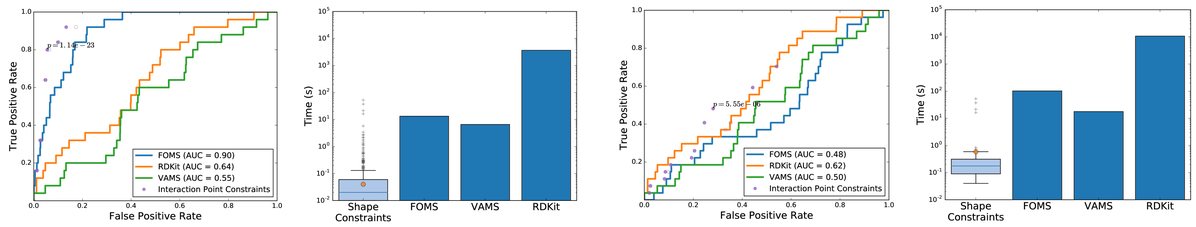


Shape constraints for a Rho-Kinase inhibitor (sticks) and receptor (surface) from PDB 2H9V derived from interaction points (spheres).

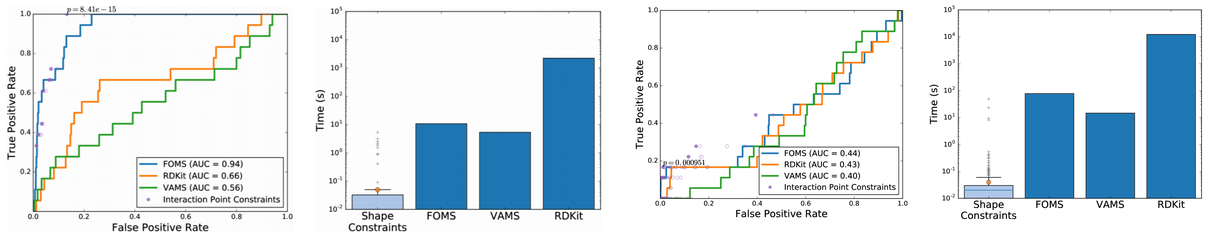
Evaluation of FOMS Performance

Target	Best p-value	Sig. Queries
CathG	4.23e-06	1%
ERα	0.0505	0%
ERα agonist	0.0106	0%
ERβ	0.0328	0%
FAK	0.000289	4%
FXIa	0.387	0%
HIVrt	0.000951	5%
HSP90	0.000755	2%
PKA	1.14e-23	80%
Rho	8.41e-15	80%

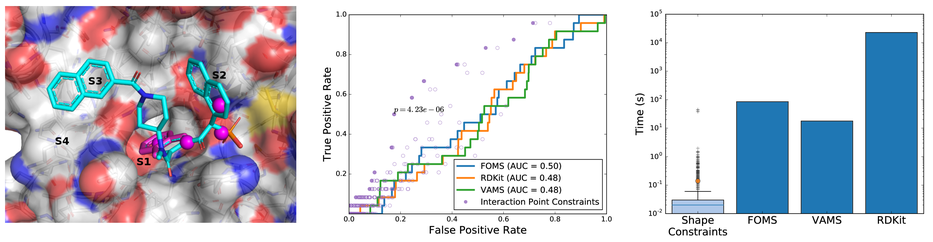
Significance of the best shape constraint queries and the percentage of queries with a Bonferroni-corrected p-value <0.01.



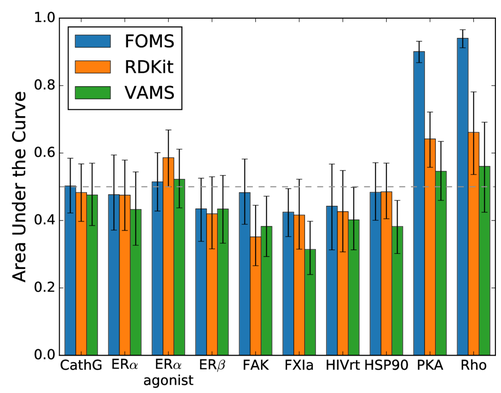
Protein kinase A virtual screening performance for various methods. Alternative fragment choice shown on the right; fragment choice can greatly affect the performance of FOMS.



Rho-kinase 2 virtual screening performance. HIV-rt virtual screening performance.

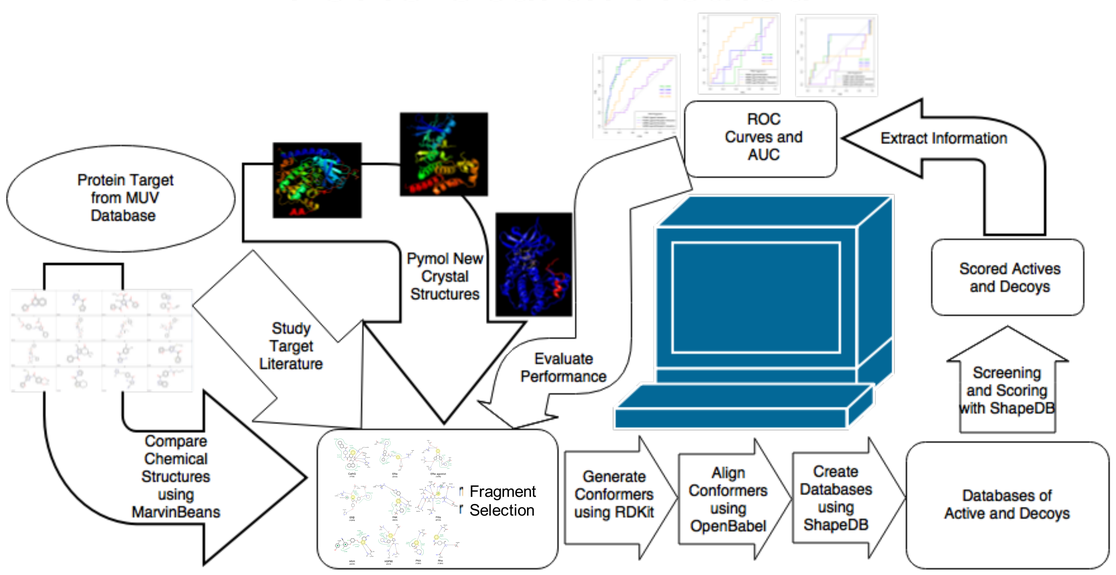


Cathepsin G. Interaction points (magenta spheres) that defined the shape constraint query with the most significant enrichment delineate only half the molecule and are located near a charge-charge interaction and aromatic interaction. Unlike molecular similarity, shape constraints can select molecules that sterically match these features while ignoring the remainder of the query ligand and avoiding significant steric clashes with the receptor.



AUCs for shape similarity methods across all targets shown with their 95% confidence intervals. The MUV dataset proved challenging for all the methods, but FOMS generally outperformed VAMS and was competitive with Open3DAlign from RDKit.

FOMS Search Method



Flowchart detailing workflow for producing and evaluating ligand screening

Conclusions

FOMS dramatically accelerates molecular shape queries and efficiently supports shape constraint queries. We anticipate that FOMS will prove a useful tool in both fragment-based drug discovery workflows, which are naturally centered around a fragment, and lead optimization workflows, where the core scaffold of the lead compound can serve as a fragment.

Acknowledgments

TECBio REU@Pitt is supported by the National Science Foundation under Grant DBI-1263020 and is co-funded by the Department of Defense in partnership with the NSF REU program.

References

1. Nicholls, A, McGaughey, GB, Sheridan, RP, Good, AC, Warren, G, Mathieu, M, Muchmore, SW, Brown, SP, Grant, JA, and Haigh, JA (2010) Molecular shape and medicinal chemistry: a perspective. *Journal of medicinal chemistry* **53**, 3862-3886

2. Koes, DR, and Camacho, CJ (2014) Shape-based virtual screening with volumetric aligned molecular shapes. *Journal of computational chemistry* **35**, 1824-1834

3. Koes DR, Camacho CJ. Indexing Volumetric Shapes with Matching and Packing. *Knowledge and Information Systems*. 2015 Apr 1;43(1):157-180.