Molecule Alignment using *Shapelets*

E. Proschak[1], M. Rupp[1], S. Derksen[1], G. Schneider[1]

proschak@bioinformatik.uni-frankfurt.de

1: Johann Wolfgang Goethe-Universität,
Beilstein Endowed Chair for Cheminformatics
Siesmayerstr. 70,
60323 Frankfurt am Main, Germany

Introduction

Complementarity of molecular surfaces is crucial for molecular recognition. A method for representation of molecular shape is presented[1]. We decompose the molecular surface into commensurate patches with defined shape by fitting hyperbolical paraboloids onto a triangulated isosurface of the Gaussian model of a molecule. As a result of this decomposition we obtain a 3D graph representation of the molecular shape, which can be used for shape matching.

Surface Calculation

The first step in our shape matching procedure is the calculation of the molecular surface. We describe a molecule *M* with a continuous spatial function. Specifically, we followed the established concept of representing a molecule as a sum of modified Gaussian functions (Eq. 1)[2,3]

$$M(\vec{x}) = \sum_{n=1}^{N} e^{\frac{-2(\vec{x}-\vec{c}_n)^2}{r_n^2}} \text{ with}$$

$\vec{x}$: spatial coordinates
N : number of atoms
$\vec{c}_n$: center coordinates
r_n : VdW-radii (1)

We used the Marching Cubes Algorithm[4] to extract the isosurface and reduced the number of surface points using a mesh decimation technique known as “Welding Vertices”[5]. We obtained a smooth isosurface representation of a molecule.

Fitting hyperbolic paraboloids

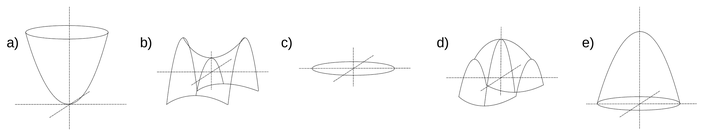
We adopted a technique used by Zachmann[6] to approximate local surface curvature. We fitted a hyperbolical paraboloid (Eq. 2) into a surface patch with a given radius using least mean squares (Eq. 3). The obtained hyperbolic paraboloids can be classified according to the shape index (SI) (Eq. 4).

$$p(u,v) = \begin{pmatrix} u & v \end{pmatrix} \mathrm{H} \begin{pmatrix} u \\ v \end{pmatrix} \quad (2)$$

$$E(\mathrm{H}) = \sum_i \left(n_i - p(u_i, v_i, \mathrm{H})\right)^2 \quad (3)$$

$$SI(p_r) = \operatorname{atan}\left(\frac{k_1 + k_2}{k_1 - k_2}\right) \text{ with} \quad (4)$$

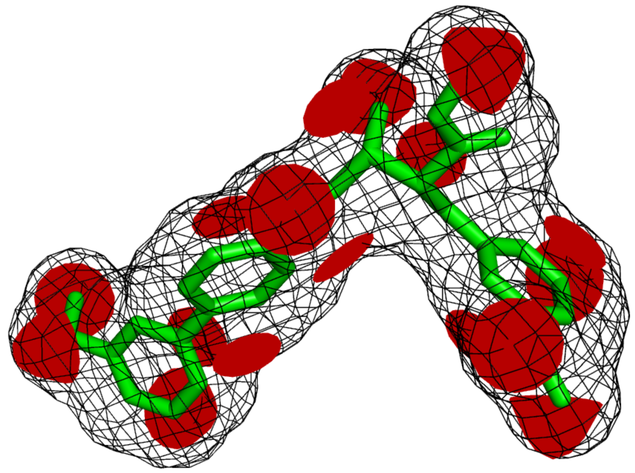
k_1, k_2 : Eigenvectors of H



Typical hyperbolical paraboloids. a. Bag (S)≈-1). b. Cleft (SI≈-0.5). c. Plane (SI≈0). d. Ridge (SI≈0.5). e. Knob (SI≈1)

Surface Decomposition

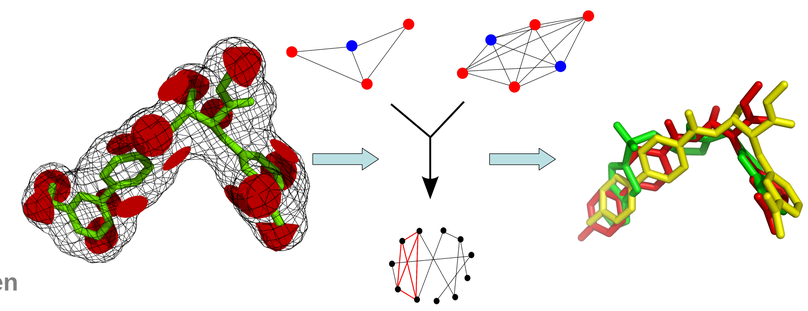
- Fitting of a hyperbolic paraboloids at every point of the surface
- Calculation of RMSD of fitted hyperbolic paraboloid
- Computation of local RMSD minima on the surface
- Filtering minima with a curvature cutoff
- Cutting out surface patches with minimal RMSD



Surface decomposition of a Factor Xa ligand FXV. Red shapes indicate planes and knobs of the molecular surface.

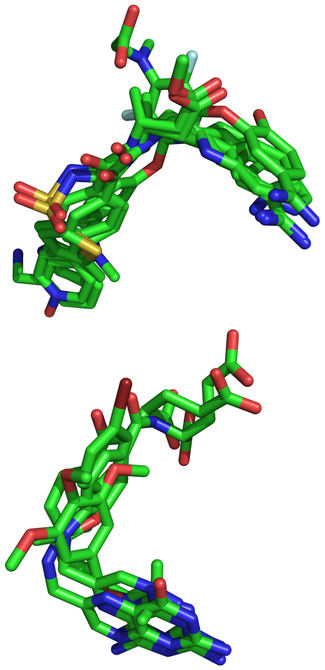
Matching Procedure

- Knobs and planes are vertices of a fully connected graph
- Edges are labled with spatial distance between the vertices
- Corresponding shapelets found by clique detection (Bron-Kerbosch algoritm)[7] on association graph.
- Transformation matrix calculated with Kabsch alignment[8] of corresponding shapelets



Alignment results

We evaluated the performance of *Shapelets* on two datasets of X-ray structures of Factor Xa and dehydrofolate reductase (DHFR). We reached at least comparable performance to previously published surface-based alignment tools[9,10].



Query	Target	RMSD [Å]
815	FXV	1.64
	PR2	0.52
	RPR	1.14
	Z34	2.07
FXV	PR2	1.39
	RPR	0.86
	Z34	1.83
PR2	RPR	1.30
	Z34	3.35
RPR	Z34	1.52
mean		1.56
sd		0.774

Query	Target	RMSD [Å]
FOL	MTX	1.18
	TMP	3.12
	WRB	1.13
MTX	TMP	2.14
	WRB	0.85
TMP	WRB	1.63
mean		1.68
sd		0.840

Outlook

We currently develop a virtual screening tool based on *Shapelets* molecular alignment. We obtained promising retrospective and prospective results, to be published soon.

References

[1] Proschak *et al.* (2007) *J. Comput. Chem.* accepted.
[2] Blinn (1982) *ACM Transactions on Graphics* 1, 235-256.
[3] Grant & Pickup (1995) *J. Phys. Chem.* 99, 3503-3510.
[4] Lorensen & Cline (1987) *Computer Graphics* 21, 163-169.
[5] Dunn & Parberry (2002) *3D Math Primer for Graphics and Game Development*, 331.
[6] Zachmann *et al.* (1992) *J. Comput. Chem.* 13, 76-84.
[7] Bron & Kerbosch (1973) *Communications of the ACM* 16, 575-577.
[8] Kabsch (1976) *Acta Cryst.* A32, 922-923.
[9] Hofbauer *et al.* (2004) *JCICS* 44, 837-847.
[10] Cosgrove *et al.* (2000) *JCAMD* 14, 573-591.

http://www.modlab.de

Supported by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften