

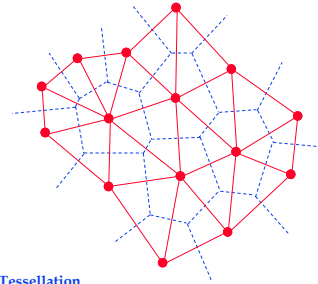
Protein Structure-Function Analysis Using Computational Geometry

Iosif Vaisman, *George Mason University*

email: ivaisman@gmu.edu

web: binf.gmu.edu/vaisman

Neighbor identification in proteins: Voronoi/Delaunay Tessellation in 2D

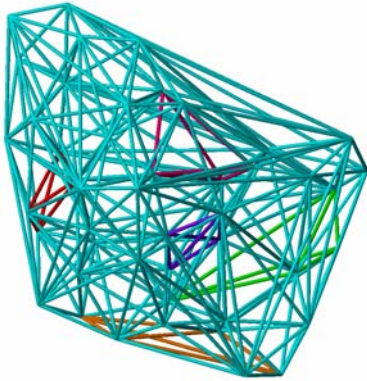


Delaunay simplex is defined by points, whose Voronoi polyhedra have common vertex

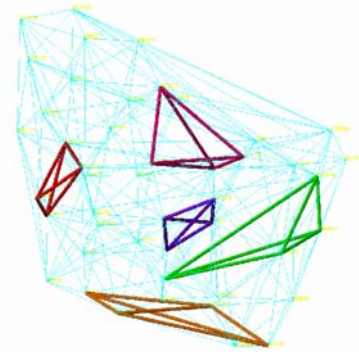
Delaunay simplex is always a triangle in a 2D space and a tetrahedron in a 3D space

Voronoi Tessellation
Delaunay Tessellation

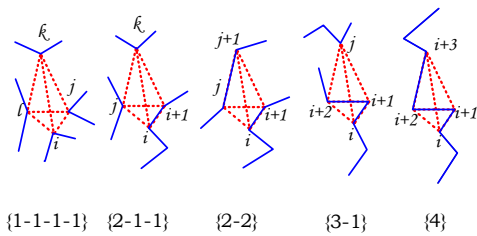
Delaunay tessellation of Crambin



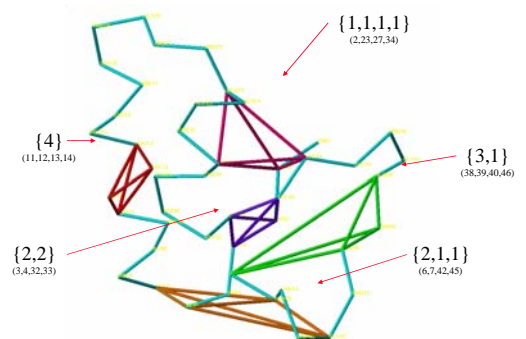
Delaunay tessellation of Crambin



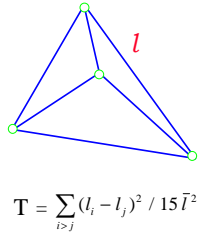
Classification of Delaunay simplices by sequential proximity



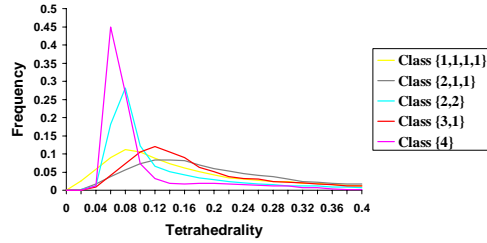
Types of Delaunay simplices in Crambin



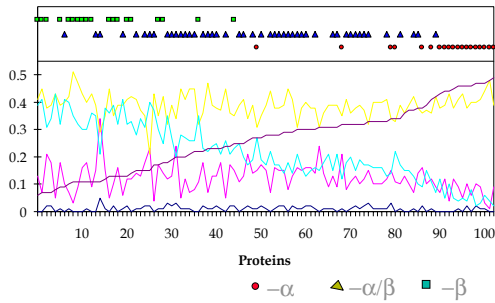
Tetrahedrality of Delaunay simplices



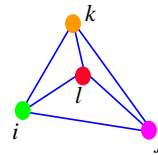
Tetrahedrality distribution of Delaunay simplices



Correlations between protein structure family assignment and relative content of classes of Delaunay simplices



Compositional propensities of Delaunay simplices



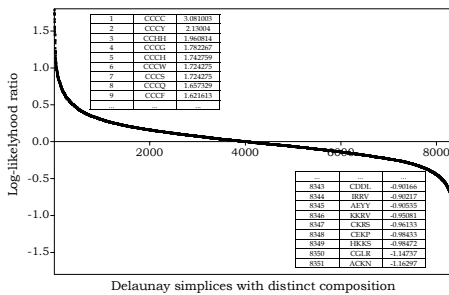
$$q_{ijkl} = \log \frac{f_{ijkl}}{p_{ijkl}}$$

f - observed quadruplet frequency,

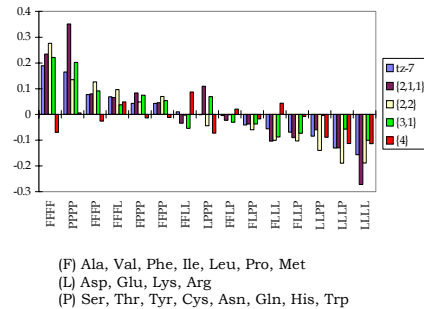
$p_{ijkl} = C a_i a_j a_k a_l$, a - residue frequency

$$C = \frac{4!}{\prod_i (t_i!)}$$

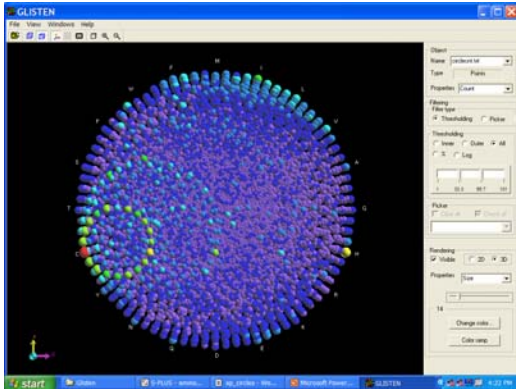
Log-likelihood of amino acid quadruplets with different compositions



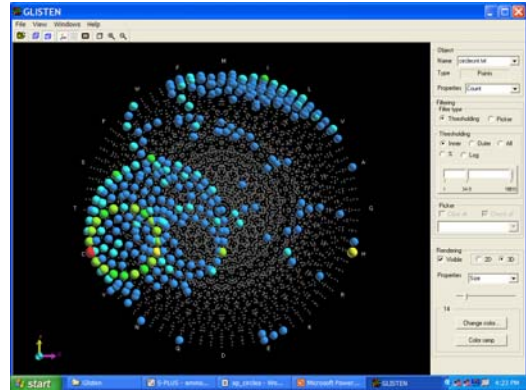
Log-likelihood of amino acid quadruplets (reduced alphabet)



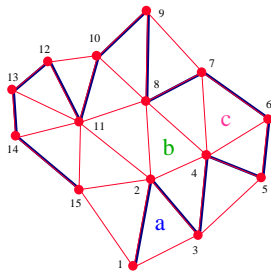
Log-likelihood of amino acid quadruplets



Log-likelihood of amino acid quadruplets



2D analogy for topological matrices

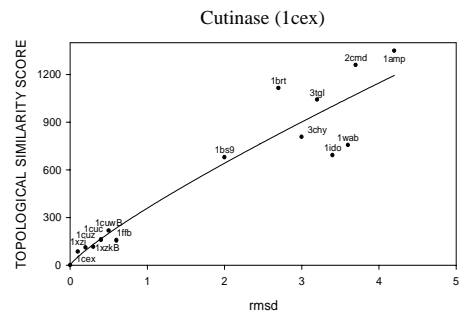


Protein backbone (C α only)
Delaunay tessellation

Simplex **a** : d1=1, d2=1
Simplex **b** : d1=2, d2=4
Simplex **c** : d1=2, d2=1

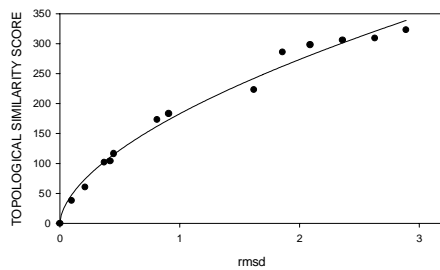
	1	2	3	4	d1
1	a	c			
2					
3					
4		b			
d2					

Topological similarity score between structures within FSSP families

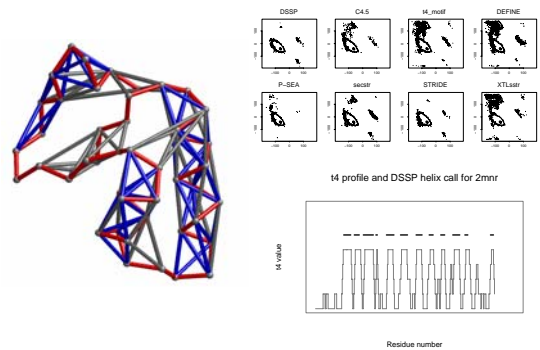


Topological similarity score between conformations from MD simulation

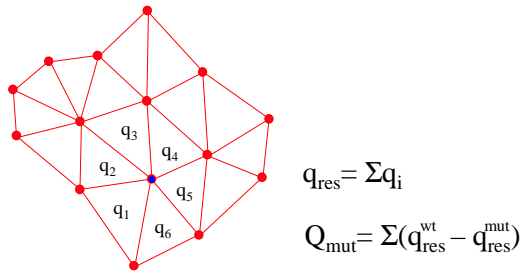
Acylphosphatase (2acy)



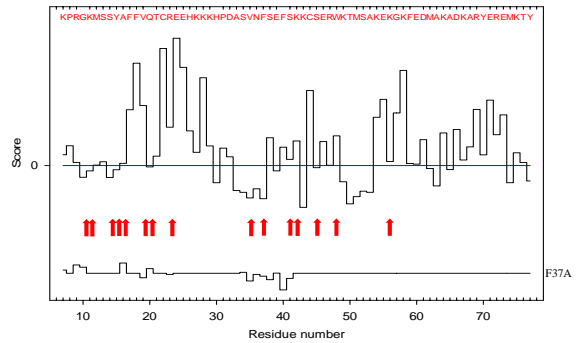
Secondary structure assignment



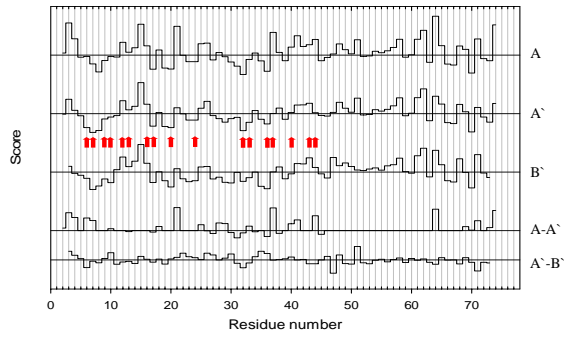
Residue and mutant score



DNA binding residues in HMG1

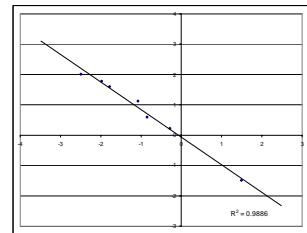


Protein-protein and protein-DNA interfaces (HMG-D)

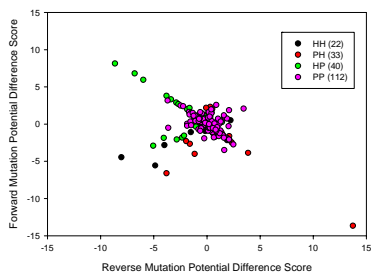


Computational mutagenesis of bacteriophage T4 lysozyme Reversibility of mutations

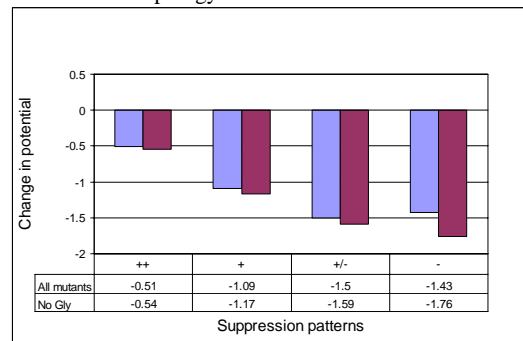
Protein	Mutation	Score change
1163	T26E	-2.49
1801	E26T	2.01
1163	A82S	1.49
1231	S82A	-1.49
1163	V87M	-0.28
1cu3	M87V	0.22
1163	A93C	-1.98
1381	C93A	1.78
1163	T152S	-1.08
1goj	S152T	1.12



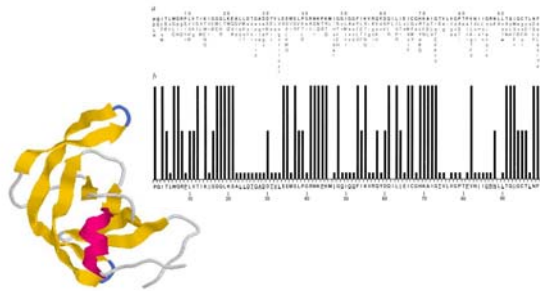
Reversibility of mutations



Computational mutagenesis of T4 lysozyme Topology-function correlation

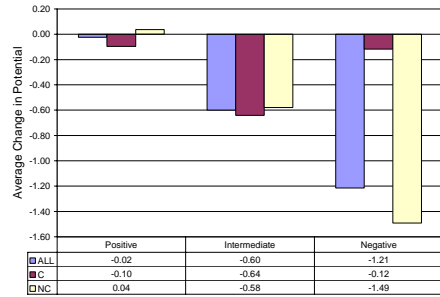


Complete mutagenesis of the HIV-1 protease Experimental data



Loeb DD et al., Complete mutagenesis of the HIV-1 protease, *Nature* 340:397

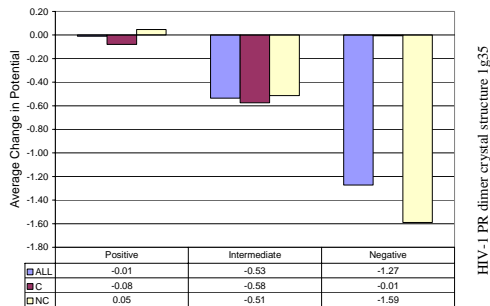
Computational mutagenesis of HIV-1 protease Topology-function correlation



HIV-1 PR monomer crystal structure 3phv

Experimental activity data from Loeb DD et al. *Nature* 340:397

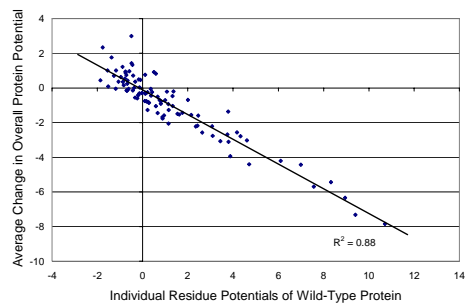
Computational mutagenesis of HIV-1 protease Topology-function correlation



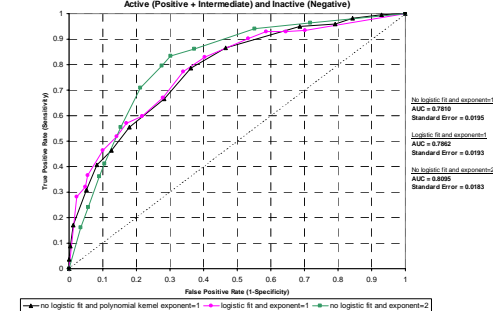
HIV-1 PR dimer crystal structure 1g35

Experimental activity data from Loeb DD et al. *Nature* 340:397

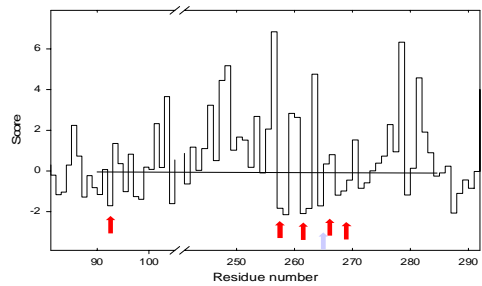
Comprehensive Mutational Profile vs. Potential Profile of HIV-1 protease (3phv)



ROC Curves for HIV-1 Protease Mutants Derived by Applying SVM to the Data Represented as a 2-Class System:

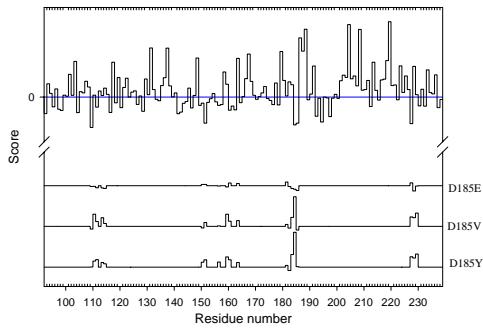


DNA binding residues in HIV RT



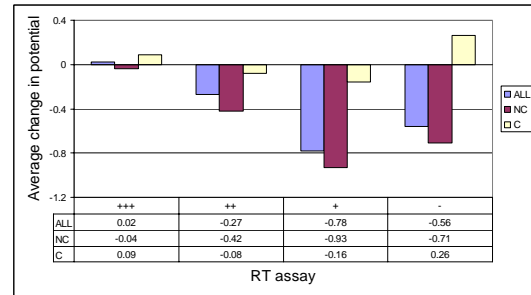
HIV RT MGBT: Ile94, Gln258, Gly262, Trp266, Gln269

Computational mutagenesis of HIV RT



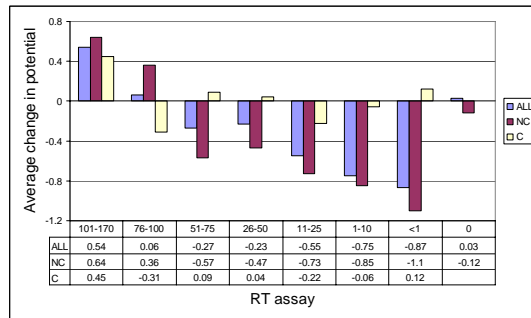
1rtj profile and D185 mutations

Computational mutagenesis of HIV RT Topology – function correlations



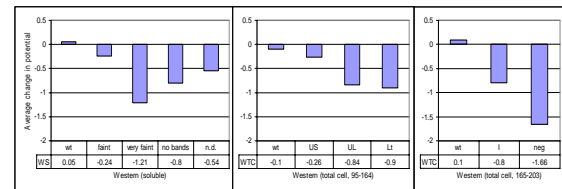
Experimental data from Wrobel JA et al. *PNAS* 95:638

Computational mutagenesis of HIV RT Topology – function correlations



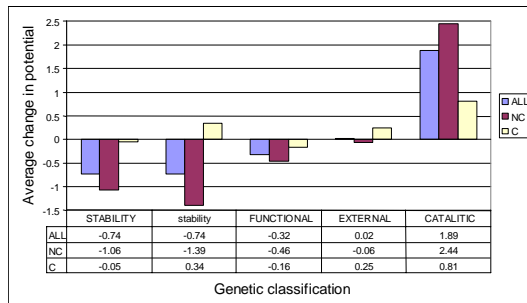
Experimental data from Wrobel JA et al. *PNAS* 95:638

Computational mutagenesis of HIV RT Topology – Western phenotype correlations



Experimental data from Wrobel JA et al. *PNAS* 95:638

Computational mutagenesis of HIV RT Residue classification



Experimental data from Wrobel JA et al. *PNAS* 95:638

Acknowledgements

Graduate students:

Tariq Alsheddi (structure alignment)
 David Bostick (topological similarity)
 Andrew Carr (functional sites, visualization)
 Sunita Kumari (structural genomics)
 Majid Masso (mutagenesis, HIV-1 protease)
 Ewy Mathe (mutagenesis, p53)
 Prabhu Raman (molecular recognition)
 Vadim Ravich (HIV RT mutagenesis)
 Greg Reck (hydration potentials, amyloids)
 Todd Taylor (statistical potentials, secondary structure, topology, protein stability)

Collaborators:

John Grefenstette (GMU)
 Curt Jamison (GMU)
 Dmitri Klimov (GMU)
 Dan Carr (GMU)
 Vladimir Karginov (IB)

Unpublished data:

Clyde Hutchison (UNC)
 Ron Swanstrom (UNC)

Funding:

NSF
 NIH-Innovative Biologics