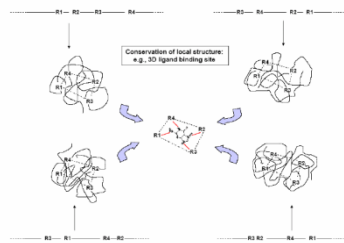
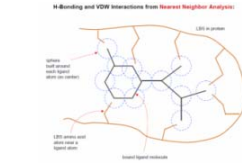


Vicente M. Reyes, Ph.D., Assistant Professor [E-mail: vmrsbi@rit.edu]
Department of Biological Sciences, School of Life Sciences, College of Science,
Rochester Institute of Technology, Rochester, NY 14623-5603

(1) Pharmacophore Concept:



(2) Nearest-neighbor Analysis:



(3) Training Sets Used:

Table 1 A. Training Set for GTP-Binding Proteins (Family 028) and CS2			Table 1 B. Training Set for ATP-Binding Proteins (Family 114)				
PDB ID	E.C. No.	Protein Description	Family	PDB ID	E.C. No.	Protein Description	Family
1C4H	4.1.1.17	Oxidative Deaminase Watanabe (Ox) (D17Y)	NA	1B36	2.7.3.37	Human Cyclin-Dependent Kinase 2	CS2
1E9H	NA	Structure of RAC1/PI3K Complex	NA	1B39	2.7.3.37	Human Cyclin-Dependent Kinase 2 Phosphorylated on Thr 160	CS2
1F9H	NA	Structure of E. coli MDR with Bound GTP and Bi	NA	1F96	2.7.1.1	Cyclin A-Cyclin-Dependent Kinase 2 Complex	CS2
1UPT	NA	Radical Structure of Alpha-Ras with Bound GTP and GTPase-Deficient Shown Stabilized with Taur	NA	1D0L	2.7.1.1	Rat MAP Kinase ERK2 with an Arg. Mutation at Position 52	MAPK
1NLC	NA	Crossed Structure of Human RAS4B GTP Bound Complexed with GTP	NA	1HCK	2.7.3.37	Human Cyclin-Dependent Kinase 2	CS2
1N1U	NA	Structural Evidence for Feedback Activation by Ras-GTP of the Ras Specific Guanine Exchange Factor SOS	NA	1V4T	2.7.1.1	Phosphorylated Cyclin-Dependent Kinase 2 Bound to Cyclin A	CS2
1P16	2.7.2.06	Structure of an mRiA Copying Engine Bound to the Phosphorylated Catalytic Terminal Domain of RNA Polymerase II	NA	1P9H	2.7.1.38	Ten Structures of the Catalytic Domain of Phosphorylase Kinase. An Active Protein Kinase Complexed with Nucleotide, Adenosine, Adenosine and Protein	Phos. Kin.
1T0B	NA	Tubulin Alpha-Delta Domain: Electron Diffraction	NA	1Q24	2.7.1.38	The Catalytic Mechanism of Phosphorylase Kinase Probed by Mutational Studies	Phos. Kin.
1A8C	3.8.6.18	GTP-Cytidylyltransferase (C1198 Mutant) in Complex with GTP	NA	1G8E	2.7.1.1	Phosphorylated CS2 Cyclin A-Substrate Peptide Complex	CS2
1C8M	2.7.2.06	Structure of a 2 Different Conformations of mRiA Copying Engine in Complex with GTP	NA	1P9H	2.7.1.38	The Crystal Structure of a Phosphorylase Kinase Peptide-Substrate Complex: Kinase Substrate Recognition	Phos. Kin.
1H9X	1.4.1.3	Crytal Structure of Bovine Liver Glutamate Dehydrogenase Complexed with GTP, NADH & Uru	NA	1G2K	NA	Structure: Bovine Glutamate Dehydrogenase Complexed with NADPH, Glutamate and GTP	CS2
1H8E	1.4.1.3	Bovine Glutamate Dehydrogenase Complexed with NADPH, Glutamate and GTP	NA	1G2K	NA	Structure: Bovine Glutamate Dehydrogenase Complexed with NADPH, Glutamate and GTP	CS2
1M00	3.4.4.4	Crossed Structure of the Bovine Muscle Adenylate Kinase Functional Ligands with GTP	NA	1G2K	NA	Structure: Bovine Glutamate Dehydrogenase Complexed with NADPH, Glutamate and GTP	CS2
1N1U	NA	Crossed Structure of RAS4B: Functional Implications	NA	1G2K	NA	Structure: Bovine Glutamate Dehydrogenase Complexed with NADPH, Glutamate and GTP	CS2
1G3P	NA	Crossed Structure of Mouse AMP (S46A) (T47L) GTP Form	NA	1G2K	NA	Structure: Bovine Glutamate Dehydrogenase Complexed with NADPH, Glutamate and GTP	CS2
3BAF	NA	The Small G-Protein RAS2A in Complex with GTP	NA	1G2K	NA	Structure: Bovine Glutamate Dehydrogenase Complexed with NADPH, Glutamate and GTP	CS2

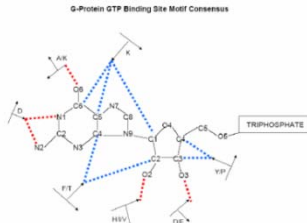
(4) Deriving Consensus Binding Site Motif (CBSM) from Nearest Neighbor Analysis Results:

Determining the 3D SM from the Hydrogen Bonds and van der Waals Interactions

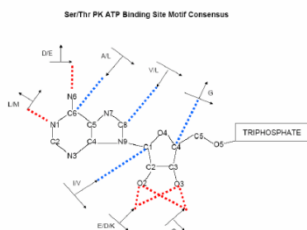
	TS1	TS2	TS3	TS4	TS5	TS6	TS7	TS8	TS9	TS10
I.A. 1										
I.A. 2										
I.A. 3										
I.A. 4										
I.A. 5										
I.A. 6										
I.A. 7										
I.A. 8										
I.A. 9										
I.A. 10										
I.A. 11										
I.A. 12										

Legend: TS, training structure; i.e., ligand atom

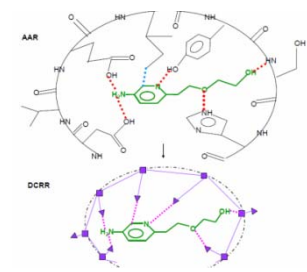
(5a) CBSM for GTP (small, Ras-type G-proteins):



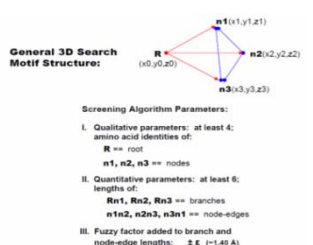
(5b) CBSM for ATP (ser-thr Protein Kinases):



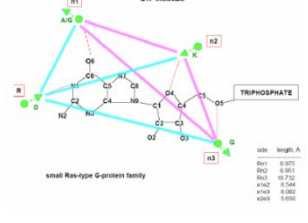
(6) Double-Centroid Reduced Representation (DCRR) of Proteins:



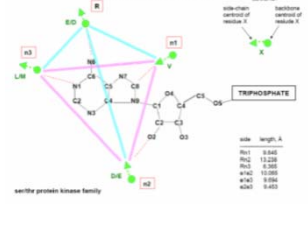
(7) Tetrahedral Data Structure of the 3D Search Motif (3DSM):



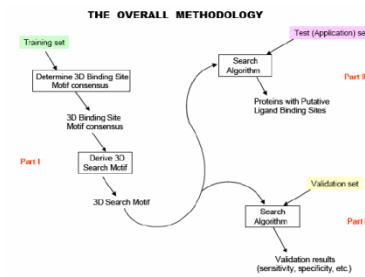
(8a) 3DSM for GTP (small, Ras-type G-proteins):



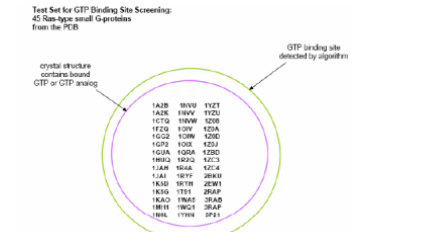
(8b) 3DSM for ATP (ser-thr Protein Kinases):



(9) The Overall Procedure:



(11a) Screening Results for the GTP Validation Set:



Pharmacophore Modeling Using a Reduced Protein Representation as a Tool for Structure-Based Protein Function Prediction

Vicente M. Reyes, Ph.D., Assistant Professor

Dept. of Biological Sciences, Rochester Institute of Technology, Rochester, NY 14623

Using the double centroid reduced representation (DCRR) of proteins, we have modeled the pharmacophores for ATP and GTP in ser/thr protein kinases (stPK) and small Ras-type G-proteins (RtGP). In DCRR, each amino acid in the protein is represented by two points, namely, the centroids of its backbone and sidechain atoms. The pharmacophore model, which we call the '3D search motif' (3D SM), is a tetrahedron with a unique root node, R, and three branch nodes, n1, n2 and n3; it also has three root-branch edges, Rn1, Rn2 and Rn3, and three branch-branch edges, n1n2, n1n3 and n2n3, all of specific lengths. These four nodes correspond to the four amino acids with the most dominant interactions (hydrogen bonds and van der Waals interactions) with the ligand atoms. We next developed an analytical algorithm (written in Fortran 90) for screening protein 3D structures for the 3D SM. The ATP and GTP 3D SMs were determined from sets of experimentally solved training structures, all of which contain the bound ligand. Validation tests performed on 'unseen' positive and negative structures reveal that the specificity of the method is nearly 100% for both protein families, and a sensitivity of 60% for the stPK family and approximately 93% for the RtGP family. Further tests reveal that our algorithm can distinguish effectively between GTP and GTP-like ligands, and between ATP- and ATP-like ligands. It is also shown that the method, which is local structure-based, works successfully in cases where global structure-based methods fail. These results show that the combined modeling and screening methods might be effective for the prediction of proteins belonging to the RtGP and stPK families. Finally, as a benchmark experiment, the method was applied to a set of protein 3D structures predicted by 123D threading and partially refined by Modeller v6.2 from the proteome of *Dictyostelium discoideum*, with promising results.