Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

Modeling of the potential coiled-coil structure of snapin protein and its interaction with SNARE complex

Ragul Gowthaman*, A. Johnwin Silvester, K. Saranya, K. S. Rathna Kanya and N. R. Archana

Department of Bioinformatics, School of Chemical and Biotechnology, Shanmugha Arts Science Technology & Research Academy (SASTRA), SASTRA UNIVERSITY, Tanjore, Tamilnadu, India;

Ragul Gowthaman* - Email: ragul@bioinfo.sastra.edu; * Corresponding author

received September 20, 2006; revised November 20, 2006; accepted November 27, 2006; published online December 02, 2006

Abstract:

Autism is a developmental disability causing learning and memory disorder. The heart of the search for a cure for this syndrome is the need to understand dendrite branch patterning, a process crucial for proper synaptic transmission. Due to the association of snapin with the SNARE complex and its role in synaptic transmission it is reported as a potential drug target for autism therapies. We wish to impart the noesis of the 3D structure of the snapin protein, and in this chase we predict the native structure from its sequence of amino acid residues using the classical Comparative protein structure modeling methods. The predicted protein model can be of great assistance in understanding the structural insights, which is necessary to understand the protein function. Understanding the interactions between snapin and SNARE complex is crucial in studying its role in the neurotransmitter release process. We also presented a computational model that shows the interaction between the snapin and SNAP-25 protein, a part of the larger SNARE complex.

Keywords: snapin; autism; memory; disorder; SNARE complex; modelling

Background:

Autism is a neurological disorder that affects the functioning of the brain causing lifelong developmental disability, including problems with communication, social interaction and physical activities. Dendrites are short fibers that conduct toward the cell body of the neuron where nerve cells receive and pass information to other nerve cells. The precise patterning of dendrites is important for determining how information is processed by a neuron. Abnormal decrease in the dendrite branches led to a state of difficulty where more hardly sites are available for receiving information and patients with autism disorder exhibit fewer dendrite branches and different dendrite patterns. Hence knowledge of the number of dendrites produced by a neuron and its patterning is very important. Intracellular protein cypin's role in increasing dendrite number by binding to tubulin heterodimers and by promoting microtubule assembly was made known recently. **[1]** Snapin, a protein first isolated as a SNAP-25 interacting protein **[2]**, was reported as a cypin-binding partner which may act to regulate dendrite number as part of a cypin protein complex. The binding of snapin to cypin crowded out the tubulin protein resulting in assembly of fewer dendrites and more branching. Previously snapin protein's role in the neuron was reported only as presynaptic **[2-6]** but latest findings reveals its role in regulating dendrite patterning in hippocampal neurons. **[7]** It was also reported that overexpressing snapin in hippocampal neurons decreases the primary dendrites growing out of the cell body and many secondary dendrites were also branched out from them. Snapin was pinpointed as a drug target protein for autism therapies because the aim is not to disturb cypin but to stop the binding between the cypin and snapin, so that cypin can still function with the other proteins it binds to. Since the protein sequence was not available directly for snapin the translated mRNA sequence was obtained from Genbank database. The secondary structure of snapin was determined and it was also checked for the presence of coiled-coil domains. A stretch containing heptad repeats typical for coiled coils was located in the C-terminal region of the molecule. **[8]** Knowledge on the three-dimensional structure of snapin protein will provide clue on its function and lead to structure based drug designing. Since there is no protein homologue of known 3D structure based on sequence similarity were found, a suitable fold for the snapin protein was found among known 3D structures by fold recognition methods. Comparative modeling method in which the three-dimensional structure of a protein sequence was modeled primarily based on its alignment to one or more proteins of known structure was applied to predict the structure of snapin protein. Molecular docking programs were used to model the interaction of snapin with SNAP-25 protein present in the SNARE complex.

Methodology:

The protein sequence for Snapin was obtained from Genbank. **[9]** The GenBank accession number for human Snapin protein is AF086837. Snapin was subjected to initial sequence characterization to infer the secondary and topological state of its amino acids.

Secondary structure prediction

The secondary structure of the snapin sequence was predicted by submitting the sequence to the leading GOR secondary structure prediction server. **[10, 11]** The GOR (Garnier–Osguthorpe–Robson) method uses both information theory and Bayesian statistics for predicting the secondary structure of proteins. GOR uses the evolutionary information of the PSI_BLAST to generate the multiple alignments after five iterations based on the non-redundant database. The sequence information of these multiple alignments was used to increase the information content for improved discrimination among secondary structures.

Coiled-coil prediction

Coiled-coil domains are characterized by a heptad repeat pattern in which residues in the first and fourth position are hydrophobic, and residues in the fifth and seventh position are predominantly charged or polar. The coiled-coil is formed by component helices coming together to bury their hydrophobic seams. As the hydrophobic seams twist around each helix, so the helices also twist to coil around each other, burying the hydrophobic seams and forming a supercoil. We have used the following three servers for coiled coil prediction: COILS, PAIRCOIL2, and MULTICOIL. COILS **[12, 13]** is a program that compares a sequence to a database of known parallel two-stranded coiled-coils and derives a similarity score which is compared to the distribution of scores in globular and coiled-coil proteins. A probability value for the sequence adopting a coiled-coil conformation was reported. Paircoil2 **[14]** predicts the parallel coiled-coil fold from sequence using pair wise residue probabilities with the Paircoil algorithm and an updated coiled coil database. The Paircoil program takes three arguments: a name for the sequence, a probability cutoff, and the amino acid sequence. The probability cutoff determines how stringently the program will screen the input sequence in detecting the existence of a coiled coil domain. The default value of 0.5 has been empirically determined to work well. MultiCoil program predicts the location of coiled-coil regions in amino acid sequences and classifies the predictions as dimeric or trimeric. The method is based on the PairCoil algorithm.

Fold Recognition

To predict the fold present in the snapin protein, we subjected the amino acid sequence to the 3D-PSSM server. **[15, 16]** This web-based protein fold recognition server uses the 1D and 3D sequence profiles coupled with secondary structure and salvation potential information. It exploits a list of Master Proteins of known structure within super families to create libraries of 1D- and 3D-profiles. It does this, first by iteratively PSI-Blasting each Master Protein against NRPROT and aligning relevant hits to create a 1D-PSSM. Next, 3D alignments between each Master Protein and all others in the same superfamily are created using the SAP algorithm. For each input sequence, the 3D-PSSM library is scanned using the global dynamic programming algorithm. The score for a match between a residue in the input sequence and a residue in the library sequence is calculated as the sum of the secondary structure, solvation potential and PSSM scores. 3DPSSM resulted in 20 proteins with folds similar to that of the snapin protein.

Protein modeling

Comparative modeling is a class of techniques in protein structure prediction that seek to construct a model of a protein's tertiary structure based on its amino acid sequence. Comparative modeling relies on a sequence alignment between the target sequence and the template sequence whose structure has been experimentally determined. We used the Crystal Structure of the coiled coil region 129-250 of the tumor suppressor gene product adenomatous polyposis coli (APC) (PDB: 1M5I) **[17]** as template for building a comparative model for snapin using MODELLER **[18]** that satisfied the spatial restraints. The alignment between snapin and the template obtained from 3DPSSM was implemented to build the global alignment in this modelling process.

Structure validation

The snapin models generated were evaluated using PROCHECK suite of programs **[19]** that provides a detailed check on the stereochemistry of a protein structure. Its output comprises a number of plots in PostScript format and a comprehensive residue-by-residue listing. The best snapin model was chosen based on the PROCHECK analysis, which includes checks on covalent geometry, planarity, dihedral angles, chirality, non-bonded interactions, main-chain hydrogen bonds, disulphide bonds, stereo chemical parameters, parameter comparisons, and residue-by-residue analysis. The Ramachandran plot for the models obtained from comparative prediction was studied in PROCHEK analysis. The Ramachandran plot is a way to visualize dihedral angles φ against ψ of amino acid residues in protein structure. It shows the possible conformations of φ and ψ angles for a polypeptide. In Ramachandran plot, the white areas correspond to conformations where atoms in the polypeptide come closer than the sum of their VanderWaals radii. These regions are sterically disallowed for all amino acids except glycine, which is unique in that it lacks a side chain. The red regions correspond to the allowed regions namely the alpha-helical and beta-sheet conformations where there are no steric clashes. The yellow areas show the partially allowed regions of left-handed helix wherein the atoms are allowed to come a little closer together. The snapin model was analyzed for the presence of coiled-coil structure by submitting it to the SOCKET server **[20]** with a packing cutoff of 8 A for center of mass distances and helix extension of 2 residues. SOCKET program finds the Knobs-into-Holes mode of

packing between alpha helices, which is characteristic of coiled coils. It unambiguously defines the beginning and end of coiled-coil motifs in protein structures and assigns a heptad register to the sequence. The packing-cutoff parameter specifies how close the knob side chain must be to all of the four hole side chains. The four distances between knob and each hole residue must all be within this cutoff for the interaction to be considered as a knob in a hole. The result shows the presence of 102 alpha helical residues in the structure.

Modeling the interaction of snapin with SNARE complex

The modeled snapin 3D structure was docked with the SNAP-25 protein present in the SNARE complex obtained from the Protein Data Bank [PDB ID: 1KIL] to study about its interaction using PATCHDOCK. **[21, 22]** PatchDock is a geometry-based molecular docking algorithm, which aims in finding docking transformations that yield good molecular shape complementarity. The PatchDock algorithm divides the Connolly dot surface representation of the molecules into concave, convex and flat patches. These complementary patches were matched in order to generate candidate transformations. Each candidate transformation was further evaluated by a scoring function that considered both geometric fit and atomic desolvation energy. Finally, an RMSD clustering was applied to the candidate solutions to discard redundant solutions. The surface area, atomic contact energy, various angle transformations and the score obtained from PatchDock program were reported.

Discussion:

We have predicted the secondary structure of snapin using GORV and the results as shown in TABLE-1 contains more of helices. The snapin sequence was observed for the presence of coiled coil regions using the following three popular servers COILS, PAIRCOIL, and MULTICOIL. The sequence when submitted to the coils server predicted the probability of the residues to be in the coiled coil structure and it was found clearly from the plot obtained in FIG-1 that there are regions with higher probability to be in the coiled coil domains in all the three window sizes with 14, 21, and 28 residues. It can also be inferred from the output of paircoil in FIG-2 that the residues 83 -119 have the probability of about 0.620 and hence show greater possibility of a coiled coil structure in snapin. The plot in FIG-3 between the residues and the probability of it being in the coiled coil structure revealed the fact that the probability of the protein to trimerize is high. Crystal Structure of the coiled coil region 129-250 of the tumor suppressor gene product adenomatous polyposis coli (APC) was predicted to have a coiled coil fold using 3DPSSM server. The alignment, consisting of snapin, APC protein and the predicted secondary structure of snapin along with PSSM scores is shown in FIG-4. Five models were generated using MODELLER and were validated and subjected to PROCHECK analysis by taking into consideration the percentage of residues in the core region in Ramachandran plot. **[23]** The validated best model for the snapin protein is shown in the FIG-5. The analysis showed that the best model has 89.7% of residues in the most favored region as shown in the FIG-6 of Ramachandran plot. As reported in the output of the socket program three major helical regions were found in positions 21-73, 76-95, 99-127 contributing to a total of 102 alpha helical residues in the structure. The snapin model was docked with SNAP-25 of SNARE complex as shown in Fig-6 to study the interaction between them using PatchDock server. The input comprised of SNAP-25 protein and snapin. The best interaction model of snapin and SNAP-25 as shown in Fig 7,8 had an approximate interaction surface area 2440 and atomic contact energy 454.06.The transformation of the three rotational and the three translational angle parameters obtained were 2.42 -1.41 0.28 16.16 51.39 77.88.The geometric shape complementarity score was 16798.

Snapin sequence
MAGAGSAAVSGAGTPVAGPTGRDLFAEGLLEFLRPAVQQL DSHVHAVRESQVELREQIDNLATELCRINEDQKVALDLDP YVKKLLNARRRVVLVNNILQNAQERLRRLNHSVAKETARR RAMLDSGIYPPGSPGK
Secondary structure
CCCCCCCHHHHHCCCCCCCCCCCHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCHHH HHHHHHCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHCCCCCCCCCCCCCC

Table 1: Snapin sequence & secondary structure

Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

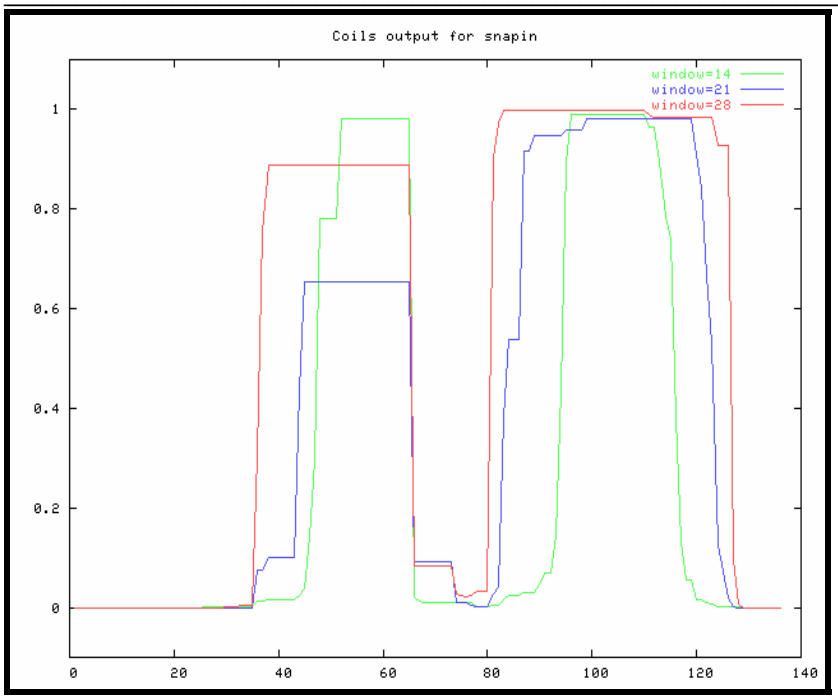

Figure 1: Graphical plot of coiled coil structure [coils]

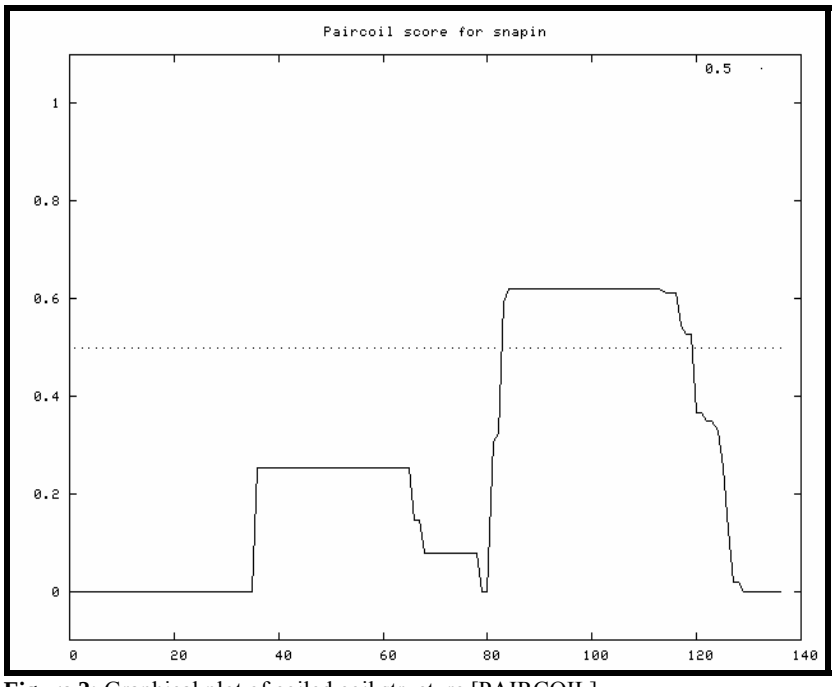

Figure 2: Graphical plot of coiled coil structure [PAIRCOIL]

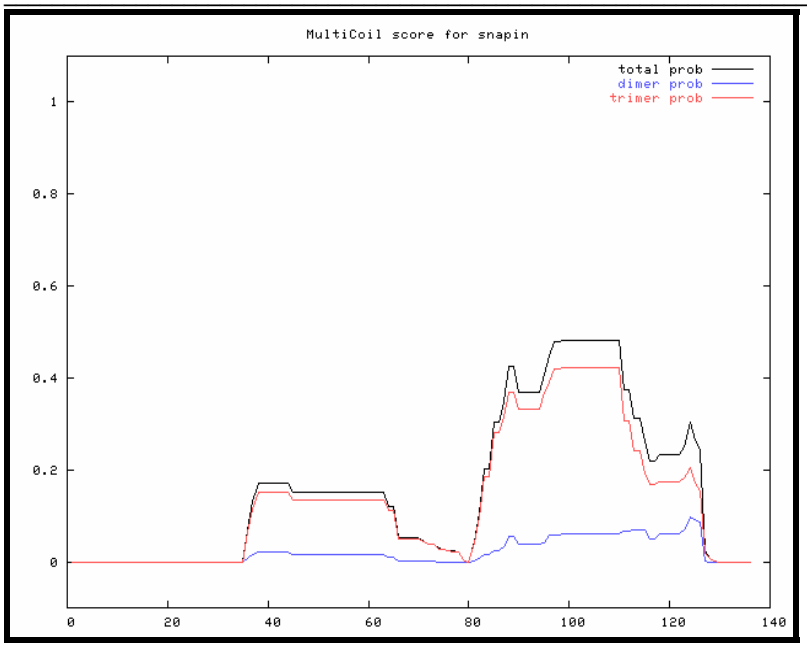

Figure 3: Graphical plot of coiled coil structure [MULTICOIL]

```
             1                                                    50
snapin__PSS  CCCCCCCCCC CCCCCCCCCC CH.HHHHHHH HHHHHHHHHH HHHHHHHHHH
snapin__Seq  MAGAGSAAVS GAGTPVAGPT GR.DLFAEGL LEFLRPAVQQ LDSHVHAVRE
-----------                    +T G- +-+-+-- +++L-+--++ --++----++
c1m5ia__Seq  .......... ........ST GYLEELEKER SLLLADLDKE EKEKDWYYAQ
c1m5ia__SS   .......... ........CC HHHHHHHHHH HHHHHHHHHH HHHHHHHHHH
CORE         .......... ........00 0050070000 0042004001 0001002100


              51                                                   100
snapin__PSS  HHHHHHHHHH HHHHHHHHHH HHCCCCCCHH HHHHHHHHHH HHHHHHHHHH
snapin__Seq  SQVELREQID NLATELCRIN EDQKVALDLD PYVKKLLNAR RRVVLVNNIL
-----------  +Q -L--+ID +L-+---+++ --Q     L+ --++++--A- ---+-+-+-+
c1m5ia__Seq  LQ.NLTKRID SLPSLQTDMT RRQ.....LE YEARQIRVAM EEQLGTCQDM
c1m5ia__SS   HH.HHHHHHH CCCCHHHHHH HHH.....HH HHHHHHHHHH HHHHCCHHHH
CORE         80.0900020 0300001001 000.....60 0190080015 0004100005


             101                                     138
snapin__PSS  .HHHHHHHHH HHHHHHHHHH HHHHHHHHCC CCCCCCCC
snapin__Seq  .QNAQERLRR LNHSVAKETA RRRAMLDSGI YPPGSPGK
-----------   +-AQ+R+-R ++ +++K+-- R-R-+L+S++
c1m5ia__Seq  EKRAQRRIAR IQ.QIEKDIL RIRQLLQSQA ........
c1m5ia__SS   HHHHHHHHHH HH.HHHHHHH HHHHHHHHHC ........
CORE         0003000200 50.0500050 0500020000 ........
```

Figure 4: 3DPSSM alignment

Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

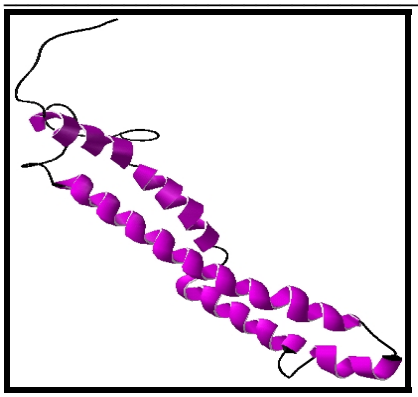

Figure 5: Structure of snapin

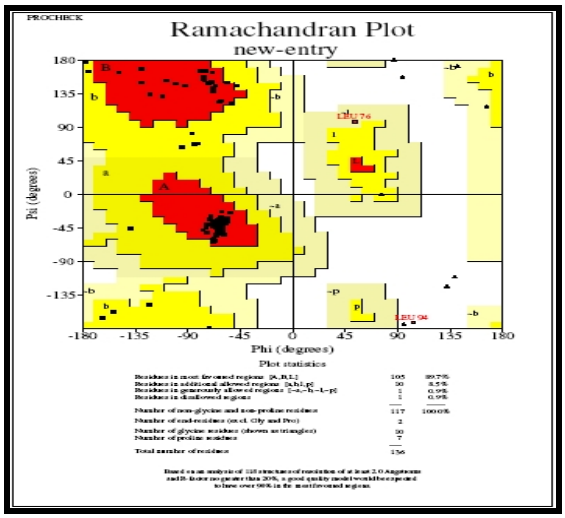

Figure 6: Ramachandran plot

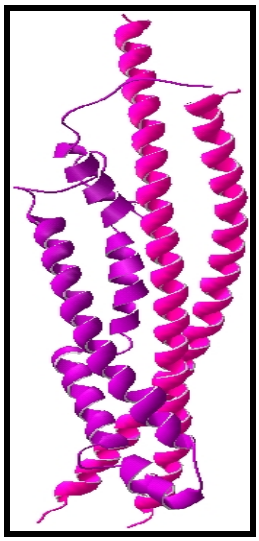

Figure 7: Model of snapin- SNARE complex interaction (ribbon representation)

Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

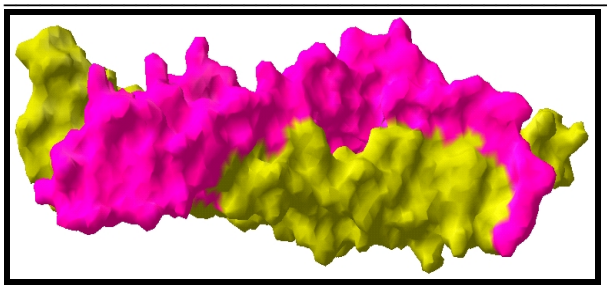

Figure 8: Model of snapin- SNARE complex interaction (surface representation)

Conclusion:

Snapin a 15-kDa protein of high therapeutic potential and biological interest was analyzed and its 3D structure was modeled. The analysis of the snapin sequence revealed that it had a predominantly alpha helical secondary structure. The protein has been reported to be highly enriched on purified synaptic vesicles. The coiled-coil structure of snapin was determined to have the ability to trimerize. Snapin has drawn in a wide interest because it increases binding to SNAREs and for the possibility of snapin to have a direct molecular link between the fusion apparatus and second messenger-dependent signaling cascades. Snapin has been found to bind to SNAP-25 of the SNARE complex indicating its regulatory role of SNAREs whose function is not confined to neuronal exocytosis. The interaction of snapin with SNARE complex is also of great importance in studying the neurotransmitter release process. We also reported a model of snapin-SNARE complex, which could be helpful to study the role of snapin as SNARE regulator in neuro transmitter release.

Acknowledgement:

The authors are grateful to Dr. K.N. Somasekharan, Dean - School of Chemical & Biotechnology and Dr. S. Swaminathan, Dean - Sponsored Research & Director CeNTAB, Shanmugha Arts Science Technology & Research Academy (SASTRA), for their support.

References:

[01] B. F. Akum, *et al., Nat Neurosci.*, 7:145 (2004) [PMID: 14730308]

[02] J. M. Ilardi, *et al., Nat Neurosci.*, 2:104 (1999) [PMID: 10195194]

[03] M. G. Chheda, *et al*., *Nat Cell Biol.,* 3:331 (2001) [PMID: 11283605]

[04] G. J. Evans & A. Morgan, *Biochem Soc Trans.*, 31:824 (2003) [PMID: 12887314]

[05] P. Thakur, *et al., J Neurosci.,* 24:6476 (2004) [PMID: 15269257]

[06] C. Ruder, *et al., Mol Biol Cell*, 16:1245 (2005) [PMID: 15635093]

[07] M. Chen, *et al*., *Molecular Biology of the Cel.,* 16:5103 (2005) [PMID: 16120643]

[08] O. Vites, *et al*., *J Biol Chem.,* 279:26251 (2004)[PMID: 15084593]

[09] D. A. Benson, *et al., Nucleic Acids Res.,* 34:D16 (2006)[PMID: 16381837]

[10] A. Kloczkowski, *et al., Proteins,* 49:154 (2002) [PMID: 12210997]

[11] T. Z. Sen, *et al., Bioinformatics,* 21:2787 (2005) [PMID: 15797907]

[12] A. Lupas, *et al., Science,* 252:1162 (1991) [PMID: 2031185]

[13] A. Lupas, *et al., Methods Enzymol.,* 266:513 (1996) [PMID: 8743703]

[14] A. V. McDonnell, *et al., Bioinformatics,* 22:356 (2006)[PMID: 16317077]

[15] D. T. Jones, *et al., Nature,* 358:86 (1992)[PMID: 1614539]

[16] L. A. Kelley, *et al., J. Mol. Biol.*, 299:499 (2000) [PMID: 10860755]

[17] L. Tickenbrock, *et al., J Mol Biol.,* 327:359 (2003) [PMID: 12628243]

[18] A. Sali, *et al., J Mol Biol.,* 234:779 (1993)[PMID: 8254673]

[19] R. A. Laskowski, *et al., J Biomol NMR.,* 8:477 (1996) [PMID: 9008363]

[20] J. Walshaw, *et al., J Mol Biol.,* 307:1427 (2001) [PMID: 11292353]

[21] D. Schneidman-Duhovny, *et al., Nucleic Acids Res.,* 33:W363 (2005) [PMID: 15980490]

[22] M. Shatsky, *et al., Nucleic Acids Res.,* 32:W503 (2004) [PMID: 15215437]

[23] G. N. Ramachandran, *et al., Adv Protein Chem.,* 23:283 (1968) [PMID: 4882249]

Edited by R. Sowdhamini

Citation: Gowthaman *et al.,* Bioinformation 1(7): 269-275 (2006)