BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 8(18)

Molecular Dynamics Simulation of Rap1 Myb-type domain in *Saccharomyces cerevisiae*

Koel Mukherjee, Dev Mani Pandey* & Ambarish Saran Vidyarthi

Department of Biotechnology, Birla Institute of Technology, Mesra, Ranchi-835 215, Jharkhand, India; Dev Mani Pandey – Email: dmpandey@bitmesra.ac.in; Phone: +91 651 2276223; Fax: +91 651 2276052; *Corresponding author

Received August 31, 2012; Accepted September 03, 2012; Published September 21, 2012

Abstract:
Telomere is a nucleoprotein complex that plays important role in stability and their maintenance and consists of random repeats of species specific motifs. In budding *Saccharomyces cerevisiae,* Repressor Activator Protein 1 (Rap1) is a sequence specific protein that involved in transcriptional regulation. Rap1 consist of three active domains like N-terminal BRCT-domain, DNA-binding domain and C-terminal RCT-domain. In this study the unknown 3D structure of Myb-type domain (having 61 residues) within DNA-binding domain was modeled by Modeller7, and verified using different online bioinformatics tools (ProCheck, WhatIf, Verify3D). Dynamics of Myb-type domain of Rap1was carried out through simulation studies using GROMACS software. Time dependent interactions among the molecules were analyzed by Root Mean Square Deviation (RMSD), Radius of Gyration (Rg) and Root Mean Square Fluctuation (RMSF) plots. Motional properties in reduced dimension were also performed by Principal Component Analysis (PCA). Result indicated that Rap1 interacts with DNA major groove through its Helix Turn Helix motifs. Helix 3 was rigid, less amount of fluctuation was found as it interacts with DNA major groove. Helix2 and N-terminal having considerable fluctuation in the time scale.

Keyword: Rap1, *Saccharomyces cerevisiae*, Homology Modelling, MD Simulation, PCA

Background:
Telomeric DNA consists of tandem array of repeated sequence and a 3' overhang. The replication of telomere and its length regulation is not solely achieved by telomerase, many complexes of telomere binding protein (Transcription Factors) helps in the synthesis. It is already reported that transcription factors bind to the telomere region of DNA for telomere length regulation and also involved in the aging of eukaryotes **[1].** These proteins bind the double stranded DNA with a very specific sequence motif. The overall sequence alignment of these proteins from different organism (Rap1 of *S. cerevisiae*, TRF1 and TRF2 of human, TAZ1 of *S. pombe*) may differ but all of them show very high structural similarity in Myb type DNA binding domain **[2].** *Saccharomyces cerevisiae is* one of the most well-known and commercially significant species of yeast. This is also a model of eukaryotic organism having 23% genome similarity with Human and shows substantial homology with telomeric transcription factors.

Rap1 (Repressor Activator Protein) is an essential site-specific DNA-binding protein in *S.cerevisiae* which plays a role in gene activation or repression. Rap1 is a modular protein and can be categorized as general Transcription factor which plays a multiple functional roles. It has been suggested that functionality of Rap1 is mainly associated with telomere length regulation and its maintenance **[3].** Rap1 associated with other protein like Sir3/Sir4 **[4]** and RIF1 **[5]** activates transcription regulation. In vitro study like immunofluorescence labeling characterizes that Rap1 molecules majorly found at the chromosomal end. It was also suggested that Rap1 binds DNA after every 18 bp on average along the telomeric DNA strand **[6].**

The full sequence length of Rap1 is 827 aminoacid **[7]** and it is divided into three main domains – (i) The N terminal BRCT domain **[8]** (ii) Central Myb-type domain and (iii) C terminal RCT domain. Above domains are considerably flexible in

nature and helps in interaction with DNA. The central Myb-type domain, a DNA binding domain (DBD), was defined in between residues 358 to 596 **[9].** The central DBD is homodimer in structure but binding to DNA is monomeric in nature. Both the monomeric domains contain three helix bundle and an N terminal arm which interacts with the major groove of DNA in the region of 360-445 (residues) and 446-578(residues), respectively **[10].** The orientation of helixes for both domains in contact with DNA shows similar pattern with other Myb-type DBD. The Myb-type DBD of Rap1 binds to a repetitive sequence of telomeric DNA within the binding site and helps in entire enclosement of DNA. In addition to DNA binding, Myb-type DBD of Rap1 can alter the conformation (bending, twisting, etc.) of DNA at its target site **[11].** The mode of contact of Rap1 appears to be related with human TRF2 which is known to protect the end of telomere region **[12].** Rap1 length wise is a long protein but it is reported that first 300 amino acid (BRCT domain) can be deleted without hampering any binding functionality of Rap1 **[13].**

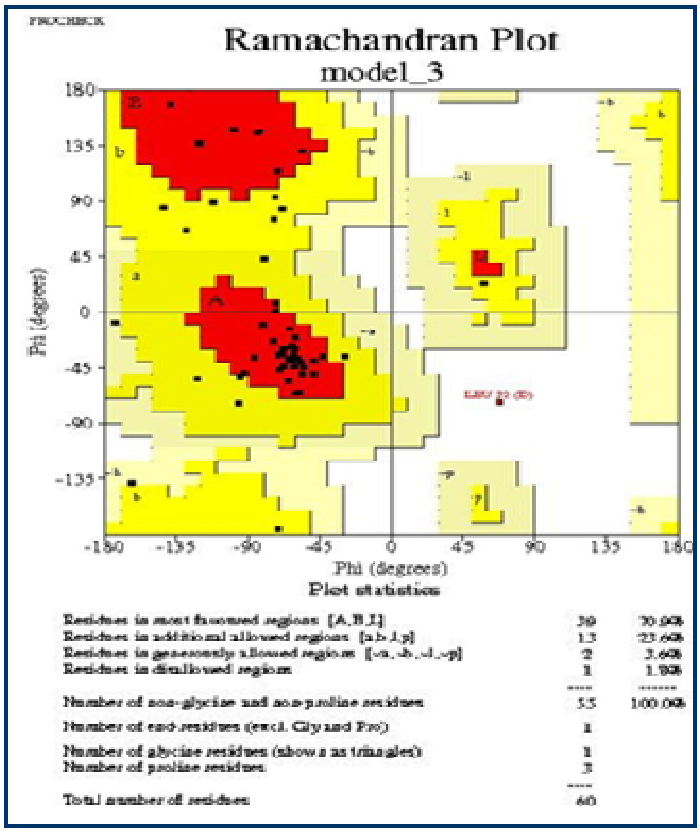

Figure 1: Ramachandran Plot of model3.pdb (rap1.pdb), the constructed structure of Myb-type domain of Rap1 DBD. The graph allows 70% residues are in core region and 23.6% allow region.

Molecular dynamics simulation is a very important tool to predict the small fluctuation in any biological molecule. To understand the motional properties over a time scale, simulation study can give the ultimate results. As biological molecules are not rigid in nature so for better understanding the functionality it is important to know the structural changes of proteins during interaction with DNA and other proteins **[14].** The Myb-type DBD of Rap1 helps in bending of the telomeric DNA during its interaction with nucleotide repeats. This information suggests that the DNA binding domain can alter its structure during the interaction. The motional behavior of Myb-type DBD of Rap1 can be only studied through simulation works. In this study we aim to model the 3D structure of Myb-type DBD of Rap1 (one monomer with 61 residues) and to study the dynamics of the structure by monitoring RMSD, RMSF and Radius of gyration **[15].** Principle Component Analysis (PCA) was also carried out to get information about the structural changes and its dependencies with the environment **[16].**

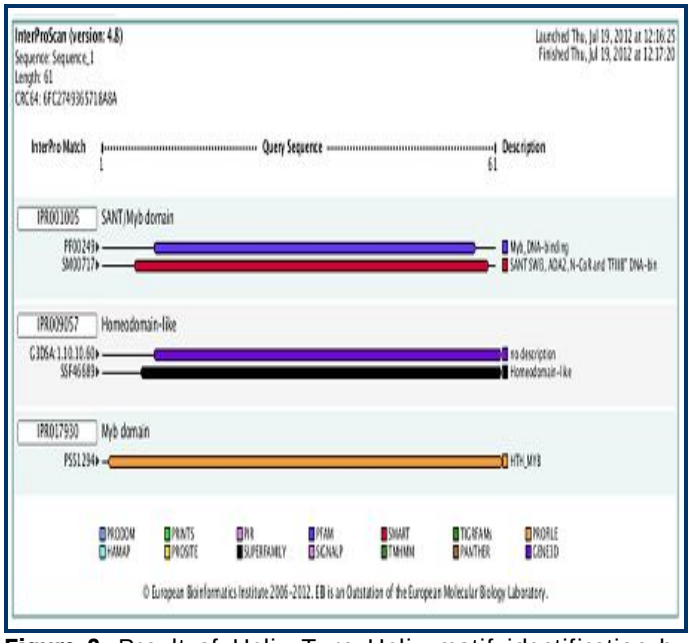

Figure 2: Result of Helix Turn Helix motif identification by InterProScan. The orange colour bar is showing the HTH-Myb type domain**.**

Methodology:

Homology Modeling and Model Evaluation

The amino acid sequence of Rap1 of *S. cerevisiae* (UniprotId: P11938) was retrieved in fasta format. The total length of Rap1 is 827 aminoacid but the HTH Myb-type domain constitute from 355 to 415 aminoacid. Sequence homologies were obtained for the required domain using BLASTp 2.2.24 **[17]** setting default parameters which was available at the National Centre for Biotechnology Information web server (http://www.ncbi.nlm.nih.gov/). The best hit selected based on query coverage, lowest E-value and good sequence similarity. We found the PDB entry 1IGN (Chain A) as template. The initial 3D model of HTH Myb-type domain of S. cerevisiae was constructed by MODELLER 9v4 **[18]** using the alignment between HTH Myb-type domain and the template protein 1IGN (Chain A). To crosscheck the model we used other online server Geno3D **[19]** (http://geno3d-pbil.ibcp.fr/cgi-bin/geno3d_automat.pl?page=/GENO3D/geno3d_home.html) and FOGUE (http://tardis.nibio.go.jp/fugue/prfsearch.html) with the sequence of 1IGN and other template sequences. Among all the results the best sequence similarity was shown for the 1IGN template sequence. Using the Structural Analysis and Verification Server (http://nihserver.mbi.ucla.edu/SAVES/) steriochemical quality and accuracy was evaluated of new modeled structure (model3.pdb) with PROCHECK **[20]** by Ramachandran plot analysis **(Figure 1)**.

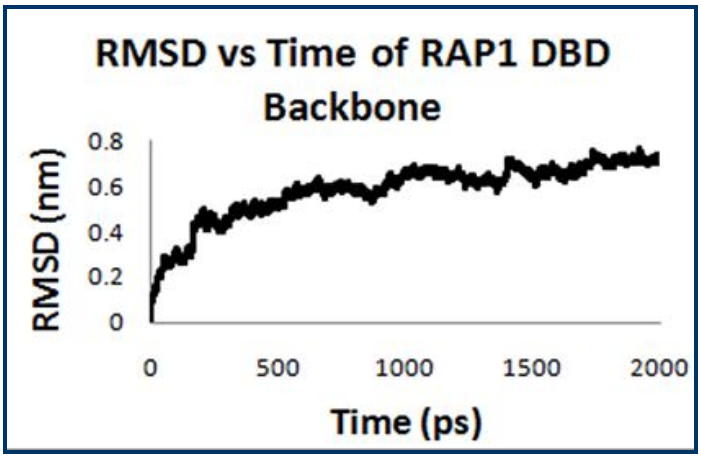

Figure 3: Time evolution of RMSD during whole simulation of time (2000ps)

Helix Turn Helix motif prediction

The secondary structure was also predicted through the online server JPred (v3) **[21]** (http://www.compbio.dundee.ac.uk/www-jpred/). It uses the Jnet algorithm in order to make more accurate 2-D structure predictions of proteins as well as prediction on Solvent Accessibility and Coiled – coil regions (Lupas method) basis. To predict the type of motif within the target sequence InterProScan (version 4.8) (http://www.ebi.ac.uk/Tools/pfa/iprscan/) **(Figure 2)** and MotifScan (http://myhits.isb-sib.ch/cgi-bin/motif_scan/) was done and both the result shows Helix Turn Helix motif (HTH).

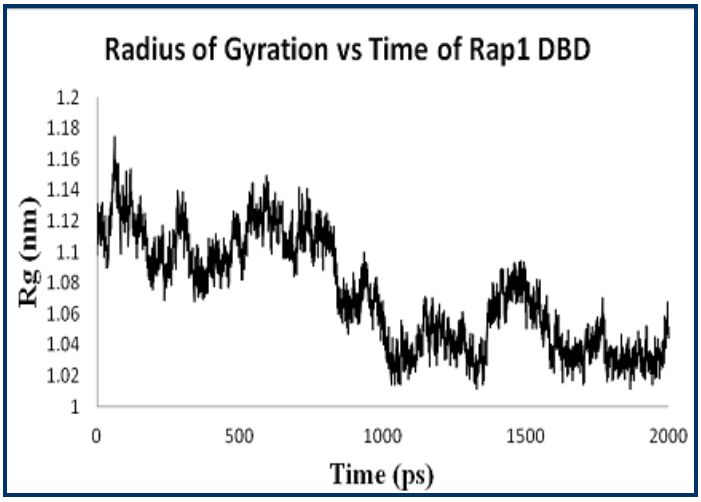

Figure 4: Plot of Radius of gyration (Rg) of backbone of Myb-type domain of Rap1.

Molecular Dynamics Simulation

All simulations were carried out using the GROMOS 96 Force Field within the GROMACS software package **[22].** The verified and evaluated homology model of Rap1 Myb type domain was used as the starting structure for MD simulation. The 3D structure of Rap1 DBD (model3.pdb) was taken in a cubic box with a 4.0 A° edge length. The simulation was conducted at a constant temperature of 300 K and a constant pressure of 1 atm and each component was coupled separately to an external bath using the Beredson Coupling method. An electrostatics interaction as Van dar Waal and coloumbic, cut-off was dealt with using a radius of 10 A°. MD simulation was performed for 2 ns. But before simulation the energy minimization was performed by steepest descent method (converged at 550 steps) and Conjugant Gradient method (converged at 4 steps). To solvate the condition the "SPC" water model (spc216.gro file) was used to fill up the box. The box contains 32108 no. of solvent molecules (water). The position restrain step was carried out for 1 n and finally the simulation was finished. After completion of simulation the trajectory files which were generated, analyzed with different tools of GROMACS.

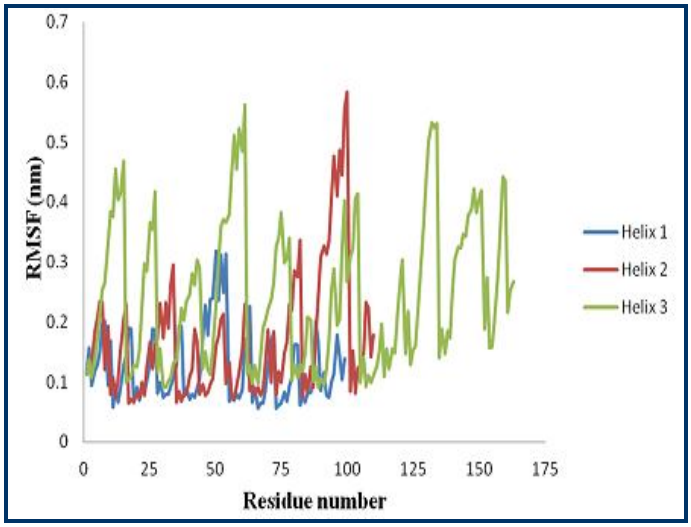

Figure 5: Comparison studies of RMSF fluctuation of Helix1, Helix2 and helix3 with the residue number. The maximum value of RMSF fluctuation shown in the graph for Helix 2(red line) is 0.3 nm**.**

Discussion:

The modeled structure was crosschecked and evaluated with different online servers. The chosen model was again analyzed by VERIFY 3D **[23]** and ERRAT (http://nihserver.mbi.ucla.edu/ERRATv2/). It was found that 63.93% of the residues had an averaged 3D-1D score > 0.2 in VERIFY 3D. Subsequently the amino acid environment was evaluated using ERRAT plots, ERRAT showed an overall quality factor of 98.077. RMSD value of the Rap1 Myb-type DBD demonstrated the fluctuation over time **(Figure 3).** It is evident from Figure3 that RMSD values increased upto 700ps, then it decreased a little bit and after 1000ps the fluctuation was in the range of 0.6-0.7 nm scale. Radius of Gyration (Rg) shows the distance of the atoms of the structure from either its centre of gravity or an axis. Our results show the much variation in the values of Rg during the simulation time **(Figure 4).** At the beginning of the simulation, increased fluctuation in Rg values was noticed but decreased as the time progresses. Our result of RMSD and Rg curve values revealed that Myb-type domain is flexible in solvent condition. Root Mean Square Fluctuation (RMSF) was performed to observe the flexibility of different segments (helix1, helix2 and helix3) as well as full structure of the Myb-type DBD of Rap1. It shows the fluctuation of each residue from its time averaged position. We noticed RMSF values for each residue of Helix1, Helix2 and Helix3 in time averaged position **(Figure 5).** In full structure of the Myb-type DBD of Rap1 RMSF analysis we noticed a high peak at the N terminal region while comparatively lower peak in the region of 25-30 residues only **(Figure 6).**

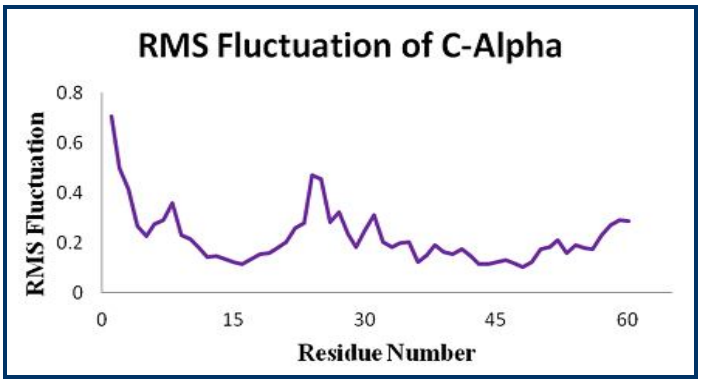

Figure 6: Plot of Root mean Square Fluctuation (RMSF) of C-alpha value in the solvent condition as a function of residue no. of Myb-type domain of Rap1.

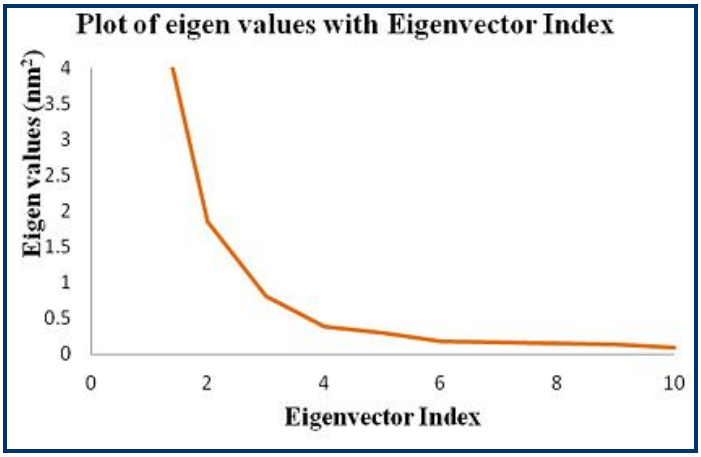

Figure 7: The plot describes the Eigen values with eigenvector index.

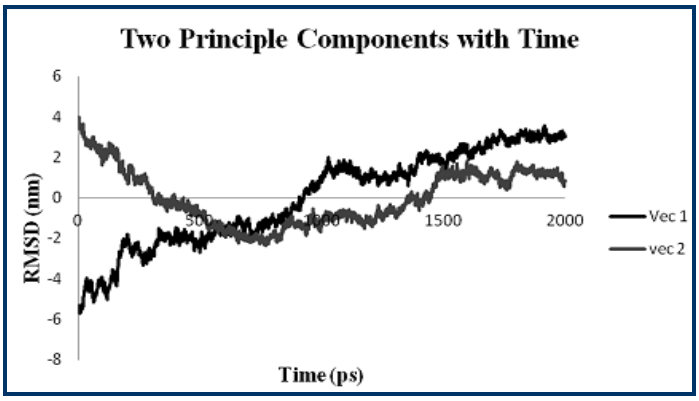

Figure 8: Plot of two principal components with simulation of time. PC1 as Vec1 and PC2 as Vec2 shown in the plot.

To analyze the major motions of protein, a small number of important modes by Principal Component Analysis (PCA) should be determined. PCA project the equations of motion on the low-dimensional vector space and facilitate the long time dynamics **[16].** PCA on Myb-type DBD of Rap1 reveals that initial 2 eigenvectors accounts for more than 85% of the global motion. They also show maximum eigenvalues and selected as the principal components 1 and 2 (PC1 and PC2) respectively **(Figure 7).** We also verified the RMSD values of PC1 and PC2 with the time evolution. Result shows the RMSD value of PC1 increased from negative value to positive value with time scale. On the other hand high positive RMSD value of PC2 decreased to negative value followed by its less positive value with time scale **(Figure 8).** This indicates that RMSD values of Myb-type DBD of Rap1lies <1.5 nm. To check the fluctuation of simulated structure (rap1_md1.pdb) with starting structure (rap1.pdb) in terms of RMSD, we superimposed these structures by Chimera 1.6.2 (http://www.cgl.ucsf.edu/chimera/).The result indicates RMSD of 13 atom pairs is 1.208 (**Figure 9).** The result of simulation at different time scale (250 ps, 500 ps, 1000 ps and 2000 ps) revealed that the helix 2 region and the N terminal showed the maximum movement **(Figure 10) [15].**

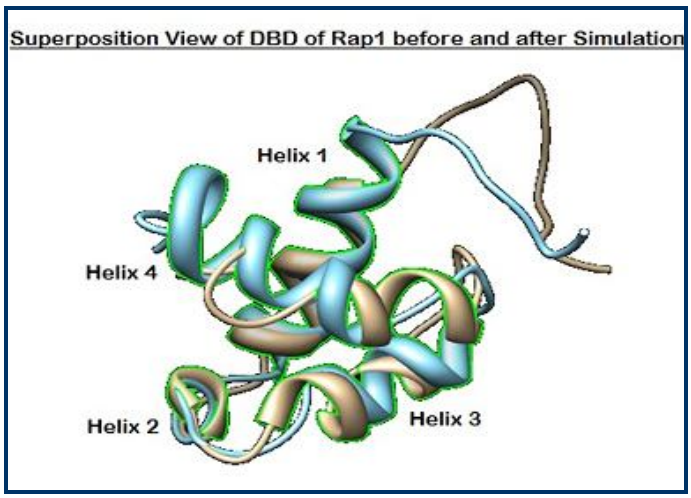

Figure 9: Graphical view of superimposed structure of DBD of Rap1 before and after simulation study by Chimera software. The blue and brown colour diagram respectively indicates the structure of after and before simulation respectively.

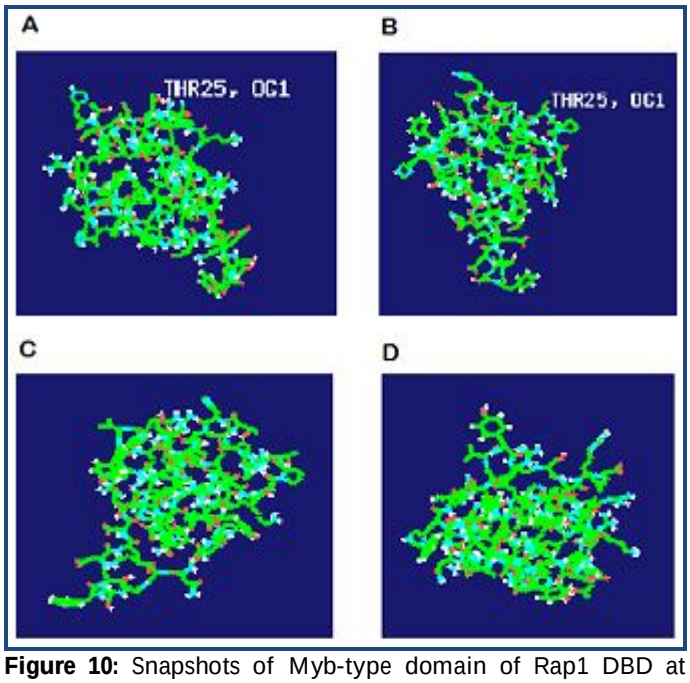

Figure 10: Snapshots of Myb-type domain of Rap1 DBD at different time points during simulation run at **(A)** 250ps; **(B)** 500ps; **(C)** 1000ps and **(D)** 2000ps**.**

Conclusion:

DNA binding domain (DBD) of Rap1 regulates the transcription process as an activator or repressor molecule. We modeled the

3D structure of DBD of Rap1 which interacts with DNA major groove through its HTH motif. Simulation study shows that helix2 and N terminal have ample fluctuation. The RMSD and RMSF values indicated that helix2 and N terminal of Rap1 is basically flexible in nature and attain conformational changes to facilitate the DNA protein interaction. RMSF curve of c-alpha displayed that helix3 is less fluctuating in comparison with other regions. As the helix3 region is known as recognition helix during the binding with DNA so its structure in nearly rigid in nature. Motional properties of different segments and full structure of Myb-type DBD of Rap1 will predict the orientation of protein and identify the amino acids which are directly interacting with the nucleotide. Further, study on the interaction of simulated structure of Myb-type DBD with DNA through docking process need to be elucidated.

Acknowledgement:
We thankfully acknowledge Mr. Uttam Kr. Mondol (Chemo bioinformatics Lab, Department of Chemistry, Raiganj University College,West Bengal, India) for suggesting in MD Simulation Studies.

References:

[1] Greider CW, *Anna Rev Biochent.* 1996 **6S**: 337 [PMID: 8811183]
[2] Wahlin J & Cohn M, *Nucleic Acids Res.* 2000 **28:** 2292 [PMID: 10871358]
[3] Conrad MN *et al*. *Cell.* 1999 **63:** 739 [PMID: 2225074]
[4] Moretti P *et al*. *Genes Dev.* 1994 **8:** 2257 [PMID: 7958893]
[5] Hardy CF *et al*. *Genes Dev.* 1992 **6:** 801 [PMID: 1577274]
[6] Gilson E *et al*. *J Mol Biol*. 1993 **231 :** 293
[7] Shore D *et al*. *EMBO J.* 1987 **6:** 461 [PMID: 15981337]
[8] Zhang W *et al. Biochem Biophys Res Commun.* 2011 **404:** 1055 [PMID: 21187076]
[9] Henry YA *et al*. *Nucleic Acids Res.* 1990 **18:** 2617 [PMID: 2187178]
[10] Konig P *et al*. *Cell.* 1996 **85:** 125 [PMID: 8620531]
[11] Hanaoka S *et al*. *Protein Sci.* 2005 **14:** 119 [PMCID: PMC2253311]
[12] Matot B *et al*. *Nucleic Acids Research Advance Access.* 2011 **1**: 11
[13] Muller T *et al*. *J Struct Biol.* 1994 **113:** 1 [PMID: 7880649]
[14] Karplus M *et al*. *Nature Structural Biology.* 2000 **9:** 9
[15] Mukherjee K *et al*. *J of Advanced Bioinformatics Applications and Research.* 2011 **3:** 324
[16] Bothra A *et al*. *The IUP Journal of Biotechnology.* 2011 **5: 3**
[17] Altschul S *et al*. *J Mol Biol*. 1990 **215:** 403
[18] Eswar N *et al*. *Curr Protoc Bioinformatics.* 2006 **5**: 6 [PMID:18428767]
[19] Combet C *et al. Bioinformatics.* 2002 **18:** 213 [PMID: 11836238]
[20] Laskowski RA *et al. J Appl Cryst.* 1993 **26:** 283
[21] Cole C *et al. Nucleic Acids Research.* 2008 **36:** 197
[22] Lindahl E *et al. J Mol Mod.* 2001 **7:** 306
[23] Eisenberg D *et al*. *Methods Enzymol*. 1997 **277:** 396 [PMID: 9379925]

Edited by P Kangueane
Citation: Mukherjee *et al*. Bioinformation 8(18): 881-885 (2012)