Temperature dependent conformation studies of Calmodulin Protein using Molecular Dynamics

Sahil Aneja[1], Vivek Kumar Bhartiya[2] and Sunita Negi[3]

[1] Cluster Innovation Centre (CIC), University of Delhi, Delhi, INDIA
[2] Centre for Photonics and Quantum Materials, Skolkovo Institute of Science and Technology, Moscow, Russia
[3]Amity University, Manesar, Gurgaon, Haryana, INDIA

Abstract. Calmodulin (CaM) protein plays a very crucial role in the calcium signaling inside the eukaryotic cell structure [1, 2]. It can also bind to other proteins/targets and facilitate various activities inside the cell [3, 4]. Temperature dependent conformation changes in the CaM protein are studied with extensive molecular dynamics simulations. The quantitative comparison of simulation data with various forms of experimental results probing different aspects of the folding process can facilitate robust assessment of the accuracy of the calculations. It can also provide a detailed structural interpretation for the experimental observations as well as physical interpretation for theory behind different aspects of the experiment. Earlier these kinds of studies have been performed experimentally using fluorescence measurements as in [5]. The calcium bound form of CaM is observed to undergo a reversible conformation change in the range 295-301 K at calcium ion concentration 150 mM. The transition temperature was observed to depend on the calcium ion concentration of the protein. Leap-dynamics approach was used earlier to study the temperature dependent conformation change of CaM [6]. At 290 K, both the N- and C-lobes were stable, at 325 K, the C-lobe unfolds whereas at 360 both the lobes unfold [6]. In this work, we perform molecular dynamics simulations of 100 ns each for the temperatures 325 K and 375 K on the apo form of CaM, 3CLN and 1CFD. A remarkable dependence of the temperature is observed on the overall dynamics of both the forms of the protein as reported in our earlier study [7, 8]. 1CFD shows a much flexible linker as compared to 3CLN whereas the overall dynamics of the lobes mainly N-lobe is observed to be more in later case. Salt bridge formation between the residues 2 (ASP) and 148 (LYS) leads to a more compact form of 1CFD at 325 K. The unfolding of the protein is observed to increase with the increase in the temperature similar to the earlier reported studies [8].

1. Introduction

Calmodulin (CaM) (an abbreviation for calcium-modulated protein) is a calcium binding messenger protein expressed in all eukaryotic cells. CaM is a multifunctional intermediate messenger protein that transduces calcium signals by binding calcium ions and then modifying its interaction with various target proteins. CaM mediates many crucial processes such as metabolism [9], apoptosis, smooth muscle contraction [10], inflammation [11], intracellular movement, short term and long term memory, and the immune response. CaM is expressed in many cell types and can have different subcellular locations, including the cytoplasm, within organelles, or associated with the plasma or


Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

organelle membranes. This protein is capable to bind a bulk of the other proteins or small organic compounds very well.

Many of the proteins that CaM binds are unable to bind calcium themselves, and use CaM as a calcium sensor and signal transducer. CaM can also make use of the calcium stores in the endoplasmic reticulum, and the sarcoplasmic reticulum. CaM can undergo post-translational modifications, such as phosphorylation, acetylation, methylation, and proteolytic cleavage, each of which has potential to modulate its actions. CaM is a small, highly conserved protein approximately 148 amino acids long (16706 Daltons). It contains four EF-hand motifs, each of which binds a Ca^{2+} ion. The protein has two approximately symmetrical globular domains (the N- and C- domain), separated by a flexible linker region. Calcium participates in an intracellular signaling system by acting as a diffusible second messenger to the initial stimuli**.**

Up to four calcium ions are bound by CaM via its four EF hand motifs. EF hands supply an electronegative environment for ion coordination. After calcium binding, hydrophobic methyl groups from methionine residues become exposed on the protein via conformational change. This presents hydrophobic surfaces, which can in turn bind to basic Amphiphilic Helices (BAA helices) on the target protein. These helices contain complementary hydrophobic regions. The flexibility of CaM's hinged region allows the molecule to "wrap around" its target. This property allows it to tightly bind to a wide range of different target proteins compared to the X-ray crystal structure; the C terminal domain solution structure is similar while the EF hands of the N terminal domain are considerably less open. The backbone flexibility within Calmodulin is key to its ability to bind a wide range of targets.

CaM undergoes large conformational changes that would depend upon its interacting partner and the environmental conditions would play a vital role in this interaction. Therefore our work is concerned with, conformational changes, in two different forms of CaM, CFD and 3CLN. CaM can be obtained from the protein data bank (PDB) [12], comprise of 148 amino acid and made of, two distinct domains, by convention, N-lobe (residue 1-64), central linker (residue 65-91) and C-lobe (residue 92-148). The coordinates of residues 1-4 and 148 are not included in the 3CLN PDB file as their coordinate are not reported in the X-ray data, but can be added by following recent studies [13].

Each domain consists of two EF-hands. Helix-loop-helix motifs are building blocks of these segments, with a Ca^{2+} binding site comprising a sequence of 12 residues, out of which the first 10 constitute the loop portion. 1CFD, PDB entry 1CFD, is free of calcium ions, in EF-hand motifs, two in both lobes.



(a) (b)

Figure 1 Initial structures of the simulations (a) 3CLN-Calmodulin protein (b) 1CFD -Calmodulin protein

2. Simulation Details

Various computational techniques are used to prepare, represent and analyze the system. Most importantly, we used NAMD [14] and VMD [15] to study structural time evolution of protein-water system. Simulations were performed for 100 ns each starting the PDB file (coordinates of atoms generated from techniques like X-r[a]y, NMR etc.) obtained from Protein Data Bank. Protein is immersed in a water box of length 10 Å using the VMD package. Water molecule is assigned as TIP3P model and CharmM27 force field [16] is used for protein and water system energetics. Initial

box dimensions are 75 × 76 × 55. These dimensions are not decided by the user, it is calculated once the system is ready by making use of tcl script in VMD. Later cutoff distance and system image ranges are decided by this data.

Particle Mesh Ewald method [17] is used to calculate long range electrostatic interactions. A cutoff distance 12 Å and a switching function at 10 Å is used to accomplish electrostatic calculations and to reduce computational load. Of course, one can increase cross-sectional area hence accuracy in energetics by increasing cutoff distance, but it requires a heavy computational power and time. As prepared protein with water box is not a neutral system. By making use of VMD tools, whole system is neutralized with Na^+ and Cl^- ions. Control in temperature is carried out by Langevin dynamics with a damping coefficient of 5 p/s. Langevin piston is used to control the pressure. Volumetric fluctuations are present to be isotropic. Spatial arrangement of ions is done automatically such that total energy of system remains minimum or ions go in a configuration of minimum energy. It is required to have a stable initial structure for studying the folding dynamics of the protein.

3. Results and Discussion

The changes that take place in the form of conformational changes of the protein are due to the change in the environmental conditions such as change in temperature, ion concentration and pH of the solution etc. Conformational change of CaM for different temperatures is an interesting phenomenon, with the increase of temperature helps the N and C-lobes to come close to each other. This is because the increase in the temperature unfolds the CaM. So unfolding in CaM is the result of some type of non-covalent interaction between lobes.

In this work, we discuss the results of our molecular dynamics simulations for different values of temperatures i.e. 325 K and 375 K for two different Apo forms of Calmodulin, 3CLN and 1CFD when rest of the environmental conditions are kept constant. The concentration of the Ca^{2+} ions is kept at 150 mM and the pH at physiological value of 7.0.

Root mean square deviation (RMSD) and the distance distribution between the residues on the two lobes are used to understand the dynamic conformational changes in the protein structure. We also calculate the energy fluctuations and salt bridge formations in order to satisfy the different conformational changes observed from RMSD and distance distribution, of the protein.

3.1. Root Mean Square Deviation (RMSD)

The root-mean-square deviation (RMSD) is the measure of the average distance between the atoms usually the backbone atoms) of superimposed proteins given as:

$$RMSD = \sqrt{\frac{\sum_{t=0}^{n}\left(x_{1,t} - x_{2,t}\right)^2}{n}}$$

where $x_{1,t}$ and $x_{2,t}$ are the two time series of the protein data.

The RMSD for the two forms of the CaM protein at various temperatures is compared as shown in Figure 2. The RMSD pattern shows a much more flexible linker for the 1CFD as compared to 3CLN. In 3CLN we observed more flexibility of the N-lobe. The protein RMSD of the 3CLN form is consistent with the RMSD of its N-lobe as it shows a similar pattern at 325 K.

The order of the RMSD is observed to remain between 3 Å to 5 Å where at 375 K it raises between 8 Å to 10 Å which shows a significant shift in the RMSD with an increase in temperature. Similar results are observed in the case of 1CFD when we compare our RMSD at 325 K and 375 K, at lower temperature the RMSD is between 5 Å to 10 Å whereas at higher temperature the value of RMSD lies between 10 Å to 15 Å. This increase in the RMSD value would be the result of kinetic energy imparted to the protein residues at higher temperature. The kinetic energy imparted to the protein would affect the stability of the protein which leads to conformation changes on changing the temperature.

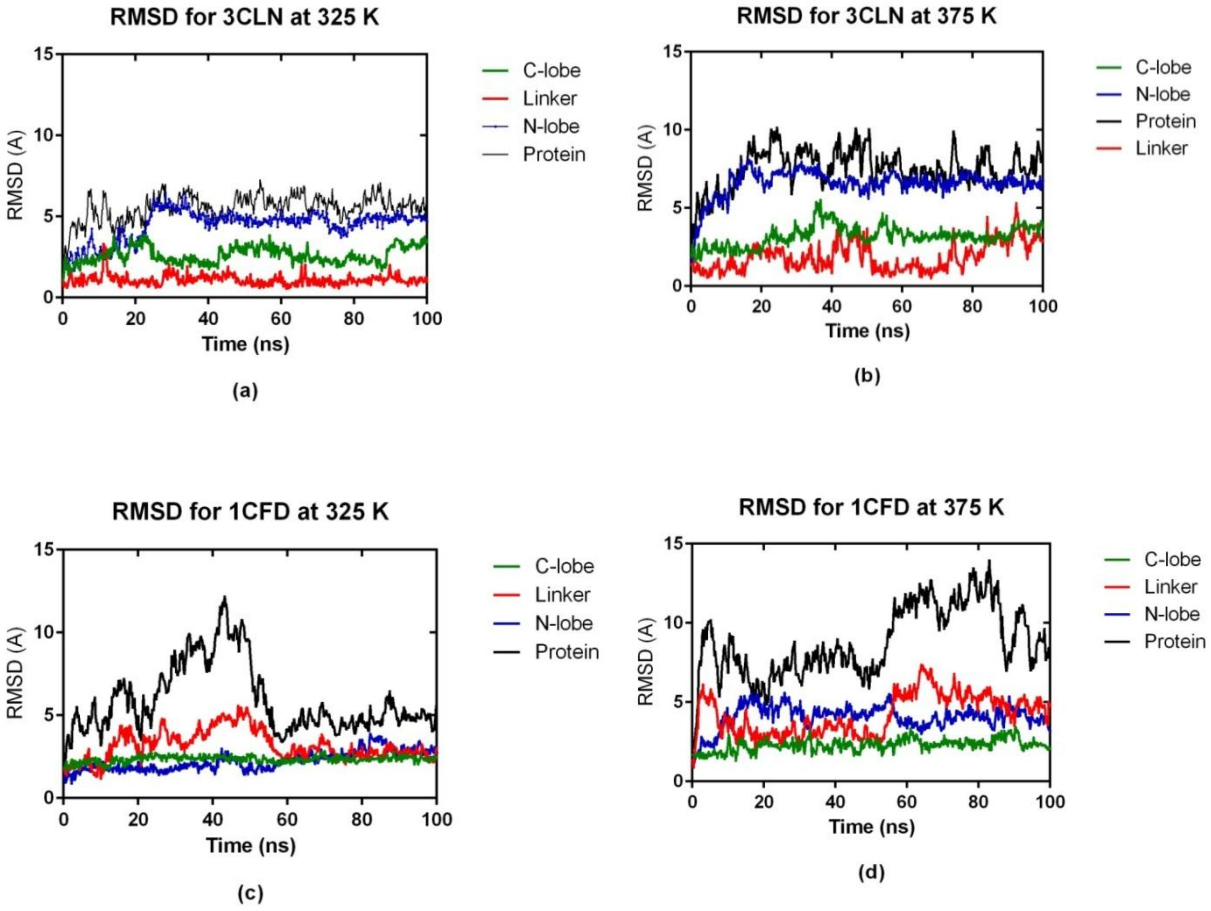


Figure 2 Temporal evolution of the RMSD for (a) 3CLN at 325 K, (b) 3CLN at 375 K, (c) 1CFD at 325 K and (d) 1CFD at 375 K.

3.2. Distance Distribution

The distance distribution is calculated to study the deviation of the protein structure from its mean position. Therefore, in this section we discuss the distance distribution calculated between residues 34 and 110 for the 3CLN and 1CFD proteins at various temperatures i.e. 325 K and 375 K. Figure 3(a) shows the distance distribution of the residues 34 and 110 of the N and C-lobes of the 3CLN protein at 325 K. The protein is observed to have a distribution around 30 Å in 3CLN at 325 K whereas on 375 K the distribution is observed around 45 Å which shows a compact form of protein at lower temperatures. Similarly we observed the compactness of protein in 1CFD but in this case order of compactness is much higher in comparison to $3CLN$, on higher temperature i.e s375 K the distribution lies around 37 Å where at 325 K the distribution is observed to be around 15 Å.

4. Another section of your paper

A thumb rule in physics is the free energy minimization as it becomes necessary to study energetics of a system before studying its properties or dynamics [18]. Because, if a system is not sitting in its lowest hoop on the energy surface, there might exist unbalanced forces, as a consequence these forces may contribute to the dynamics [19]. The initial structural energy minimization is done automatically by NAMD when commands are written in script file. Protein folding is a self-assembly process, guided by the structure of the minimum free energy. Therefore in order to understand the structure of

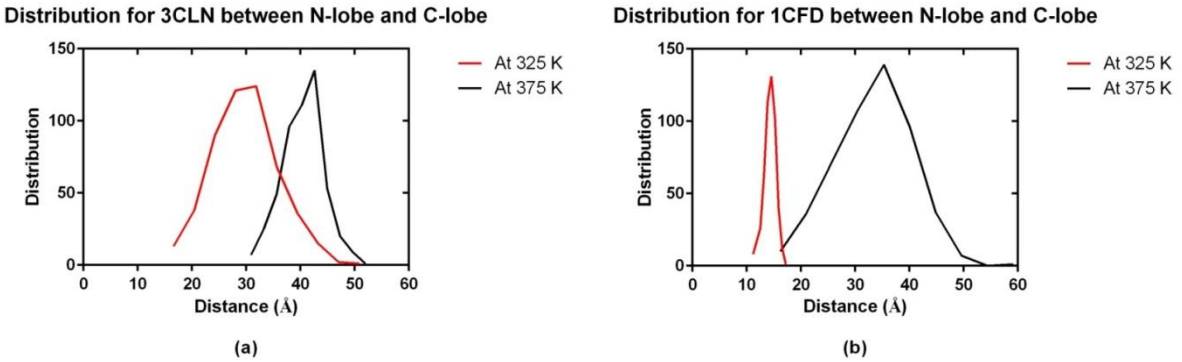


Figure 3: Distance distribution between the residues 34 and 110 of the N and C-lobes of (a) 3CLN at 325 and 375 K (b) 1CFD at 325 and 375 K.

the protein requires knowing of the folding energetics, which also governs the dynamics of protein or conformational changes in protein. The temperature dependent changes in conformation can be related to the energetics of the protein like the linker is more flexible in the case of 1CFD leading to bend compact form during the simulation at 325 K is due to change in temperature governs to break the rigidity between linker.

Here we are concerned with the energetics of various folding events during the time evolution of the system. Energy measurements are therefore a good tool to categorize folding and unfolding events. So in order to study the dynamics of the proteins at various temperatures in detail we perform an energy calculation using the "NAMD energy" plugin within the VMD program. The folding of proteins is due to presence of multiple conformations formed by various contacts between residues of the main and side chains. Figure 4 shows the linker energy term of the 1CFD and Figure 5 shows the N-lobe energy term of the 3CLN at 325 K and 375 K calculated using "namdenergy" plugin in VMD.

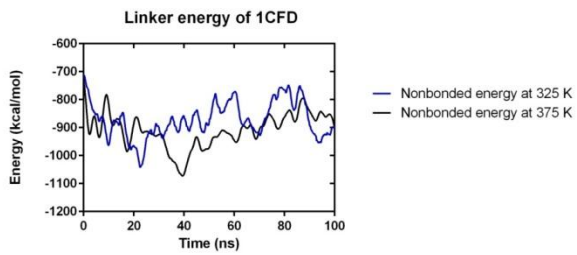


Figure 4: Linker energy calculated for 1CFD at 325 and 375 K using "NAMD Energy" plugin.

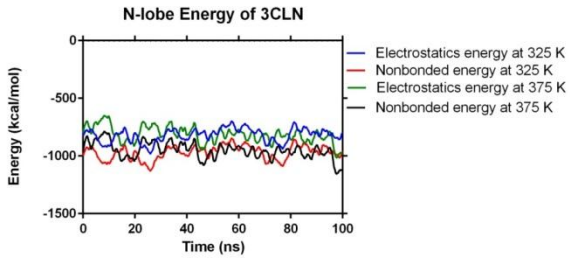


Figure 5: N-lobe energy calculated for 3CLN at 325 and 375 K using "NAMD Energy" plugin.

As discussed in Section 3.1, the flexibility of the linker is observed to be more in the case of 1CFD in comparison to 3CLN. This was confirmed from the RMSD of the two forms which showed a greater flexibility of the N-lobe in the case of 3CLN in comparison with 1CFD. We did the energy calculation using namd which also leads us to same result as shown as in Figure 4 and Figure 5. The Non-bonded energy term for the linker of 1CFD is as in Figure 4 at both the temperatures which represents the flexibility of the linker to be more in this case. On the other hand, the Electrostatics and Non-bonded energy terms for the N-lobe of the 3CLN form are as shown in Figure 5 which verifies the flexibility of N-lobe to be more in this case. The N-lobe energy calculated in the case of 1CFD is observed to change very less showing an insignificant variation in this case. Similarly the linker energy in the case of 3CLN does not vary much from its average value.

5. Salt Bridge Formations

In order to understand the dynamics and the conformational changes occurred with change in temperature, we want to analyze every possible interaction between the residues of the two lobes of the protein. We therefore calculate salt bridges formed between lobes of 1CFD at both temperatures 325 K as well as 375 K. A number of salt bridges are observed to be formed between the residues of the N- and C-lobes of the protein. Three salt bridges are observed to be formed between the residues 2 (ASP) and 148 (LYS), 7 (GLU) and 148 (LYS), and 11 (GLU) and 115 (LYS) of N-lobe and C-lobe of 1CFD at 325 K. Similarly, three salt bridges are observed to be formed between the residues 2 (ASP) and 115 (LYS), 6 (GLU) and 115 (LYS), and 7 (GLU) and 115 (LYS) of N-lobe and C-lobe of 1CFD at 375 K.

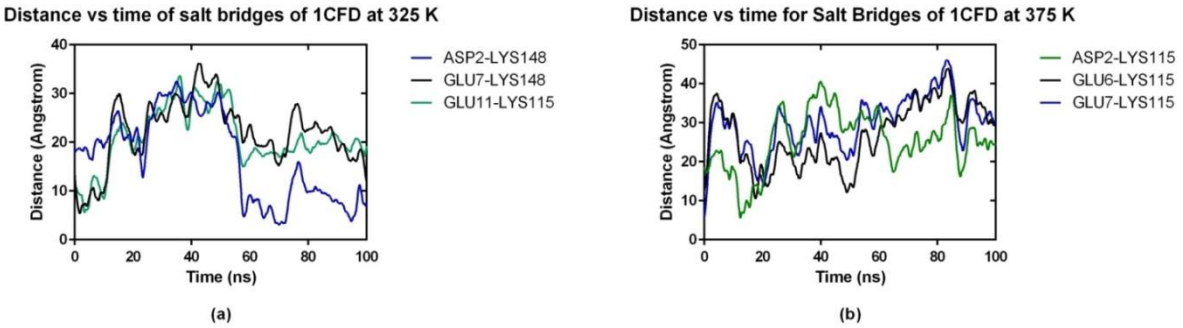


Figure 6: The variation of the salt bridge distance between the residues of N and C-lobes of 1CFD at (a) 325 K and (b) 375 K.

We analyze the distance between different salt bridges during the entire simulation as shown in Figure 6 and observed that in the case of 1CFD at 325 K one of the salt bridges between residues 2 (ASP) and 148 (LYS) distance reduces below 10 Å near 60 ns and remains such till 75 ns. Around 78 ns the residues move apart from each other but again gradually start coming close to each other around 80 ns making the salt bridge. We represent the snapshots of this particular salt bridge as shown in Figure 7, showing the interaction between residues 2 (ASP) and 148 (LYS) at different times of simulation. The salt bridge once formed leads to a more compact form of the protein at 325 K wherever at 375 K the salt bridges are formed but they do not get closer as shown in the Figure 6(b), so the decrease in the temperature shows compactness in protein 1CFD. In this case the linker is also observed to bend at the time of the salt bridge formations thus leading to a more compact structure. This is consistent with our earlier analysis which leads to a compactness of 1CFD at 325 K [8].

6. Conclusions

The functioning of so many proteins is dependent on environmental conditions like in CaM protein (3CLN and 1CFD) calcium ion removal, change in ionic strength and temperature variations leads to change in conformations or flexibility of protein. Extensive MD simulations for 100 ns performed on both Apo form of CaM (3CLN & 1CFD) at different temperatures 325 K and 375 K. All rest of the other environmental factors are maintained constant during these simulations. 1CFD shows a much flexible linker as compared to 3CLN whereas the overall dynamics of the lobes mainly N-lobe is observed to be more in later case. Salt bridge formation between the residues 2 (ASP) and 148 (LYS) leads to a more compact form of 1CFD at 325 K. The unfolding of the protein is observed to increase with the increase in the temperature similar to the earlier reported studies. Thus changing temperature comes out to be the key factor in the flexibility of CaM, RMSD calculations of CaM (3CLN & 1CFD) at different temperatures 325 K & 375 K shows the flexibility of N-lobe and Linker at these

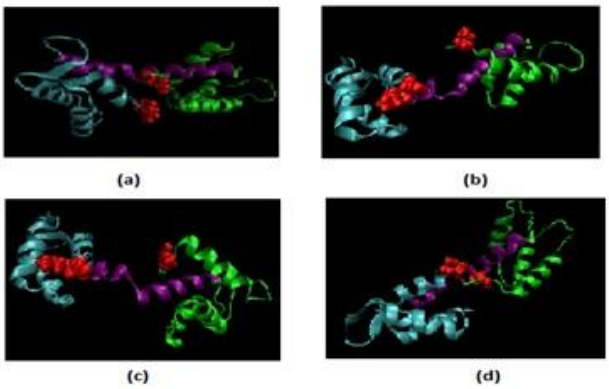


Figure 7: Snapshots showing the salt bridge formation between residue 2 (ASP) and residue 148 (GLU) of the N and C-lobes of 1CFD protein at 325 K at (a) 0 ns (b) 20 ns (c) 40 ns and (d) 58 ns.

temperatures. The flexibility of N-lobe (3CLN) and linker (1CFD) on both the temperatures is consistent with the experimental observations reported earlier. In order to understand this behavior in more detail, we calculate different energy terms corresponding to different parts of the protein. The flexibility of the linker is observed to be more in the case of 1CFD at both the temperatures i.e 325 K and 375 K in comparison to 3CLN which is shown in the form of nonbond energy terms of the linker. In the case of 3CLN, the N-lobe is observed to be more flexible at both the temperature values adding overall flexibility to the protein. The distance-distribution plots are shown at both the temperatures for 3CLN and 1CFD, where 1CFD shows more compactness at lower temperature 325 K. This compactness is observed as a result of salt bridge formations at this temperature in 1CFD whereas no salt bridge formation is observed in the case of 3CLN where the overall dynamics is governed by the N-lobe.

Acknowledgement
This work is funded by Department of Science & Technology (DST), Delhi. Mr. Sahil Aneja is receiving a fellowship from the DST project. We also acknowledge Inter University Accelerator Centre (IUAC), New Delhi for providing us computational resources for doing this research work.

References

[1] H. Frauenfelder, Proceedings of the National Academy of sciences of the USA 99, 2479-2480 (2002).
[2] M. Ikura, JB Ames, Proc. Natl. Acad. Sci. USA 103, 1159-1164 (2006).
[3] JL Fallon, Quiocho, FA Structure 11, 1303-1307 (2003).
[4] M Vandonselaar, RA Hickie, JW Quail, LTJ Delbaere, Nat. Struct. Biol 1, 795-801 (1994).
[5] Preeti Gangola and Harish Pant, Biochemical and Biophysical Research Communications, 111(1): 301-305 (1983).
[6] Jens Kleinjung, Franca Fraternali, Stephen Martin, R. and Peter Bayley, PROTEINS: Structure, Function and Genetics, 50:648–656 (2003).
[7] Sunita Negi, Biophysical Journal, 108, Issue 2, Supplement 1, 46, 108 (2015).
[8] Sunita Negi, The Biophysics Journal, 329703, 9 (2014).
[9] RM Anthony, RJ Dedman, Nature 285, 73 – 77 (1980).
[10] DP Wilson, Cindy Sutherland and MP Walsh, The Journal of Biological Chemistry 277, 2186-2192 (2002).
[11] JM Nakashima, DM Hyde and SN Giri, The American Journal of Pathology 124(3), 528-536 (1986).

[12] YS Babu, JS Sack, TJ Greenhough, CE Bugg, AR Means and WJ Cook, Nature 315, 37-40 (1985).
[13] JL Pascal-Ahuir, EL Mehler and H. Weinstein, Molecular Engineering 1, 231-247 (1991).
[14] L. Kale, R. Skeel, M. Bhandarkar, R Brunner, A Gursoy, N Krawetz, J Phillips, A Shinozaki, K Varadarajan and K Schulten, Journal of Computational Physics 151, 283–312 (1999).
[15] W Humphery, A Dalke and K Schulten, J Mol. Graphics 1, 27-33 (1996).
[16] BR Brooks, RE Bruccoleri, BD Olafson, DJ States, S. Swaminathan and M Karplus, J Comput. Chem 4, 183-187 (1983).
[17] TA Darden, L Perera, L Li and L Pedersen, Structure with Folding and Design 7, 55-60 (1999).
[18] S Negi, Biophysical Journal 108, 46 (2015).
[19] R. L. Baldwin, J. Mol. Biol., 371, (2007).