MOLECULAR DYNAMICS SIMULATION OF HORSE-HEART CYTOCHROME C IN WATER-METHANOL SOLVENT SYSTEMS

A Thesis by

DEVAKI NANDAN GAUTAM

Master of Science, Tribhuvan University, 2002

Submitted to the Department of Chemistry
and the faculty of the Graduate School of
Wichita State University
in partial fulfilment of
the requirements for the degree of
Master of Science

May 2014

MOLECULAR DYNAMICS SIMULATION OF HORSE HEART CYTOCHROME C IN WATER-METHANOL SOLVENT SYSTEMS

The following faculty members have examined the final copy of this thesis for form and content, and recommend that it be accepted in partial fulfilment of the requirement for the degree of Master of Science with a major in Chemistry.

……..

Katie Mitchell-Koch, Committee Chair

……..……

William C. Groutas, Committee Member

……..

Douglas S. English, Committee Member

……..

James G. Bann, Committee Member

……..

Elizabeth Behrman, Committee Member

ACKNOWLEDGEMENTS

I express my sincere gratitude to my supervisor, Dr. Katie Mitchell-Koch, Assistant Professor of Department of Chemistry, Wichita State University for her continuous support and guidance during my research and study in computational chemistry. Because of her inspiration, I learned the approach of computational science in chemistry and succeed to accomplish this research. Moreover, I am indebted to her in my lifetime for her financial support in my study.

I would like to thank Department of Chemistry, Wichita State University and its High Performance Computing Centre and Ablah Library for providing ample resources and valuable opportunities during my study. I dedicate my research to this university.

I am very thankful to Professor Dr. David Eichhorn, Chair of Department of Chemistry for his great help to pursue my MS in this university. I am also thankful to the professors with whom I completed chemistry courses for this Master's degree and to Dr. Kavin Langenwalter for his computer support.

Similarly, I would like to thank sincerely all professors of my graduate committee members for their valuable time to look at my research and their suggestions.

I appreciate my colleagues for their encouragement and support, specially, Lava Raj Kadel, Bushy Reddy and my lab-mates, Ja, Chandana and Rajni.

Finally, I would like to remember my family, my son Sudin and my dear Supriya, for their patience and time.

ABSTRACT

Molecular Dynamics simulations have been carried out to investigate the dynamics of horse heart Cytochrome C and associated crystallographic water molecules in different water-methanol systems. The 100 ns simulation predicts that hh-CytC undergoes different dynamical transitions with some common conformations in different solvents. With increase of methanol concentration in solvents, hh-CytC has increased flexibility, fluctuating its hydrophobic solvent accessible surface area (SASA), and number of persistent internal hydrogen bonds with long hydrogen-bond-lifetime. The protein became more liquid-like in mixed solvents compared to pure solvents; flexibility increases in the absence of the crystallographic water. Similarly, the number of hydrogen bonds between solvent molecules and hh-CytC decreased with increasing of methanol concentration. Water-protein and methanol-protein hydrogen bond lifetimes were computed 11.5 and 16.6 picoseconds, respectively, in pure solvents.However, in mixtures, solvent-protein hydrogen bond lifetime was higher in twenty percent methanol than in fourty percent in water. The surface crystallographic water molecules diffused easily in bulk solvents within 1 nanosecond and protein surface is stabilized by hydrogen bonds with a solvation layer. The two crystallographic water molecules which are buried internally in hh-CytC have 5 to more than 100 nanoseconds residence time in the conserved sites with 100's of picoseconds of hydrogen bond lifetime depending on the solvent compositions. The residence time might depend on the mechanism of conformational transition of protein in simulation. Solvent water molecules exchange these buried water molecules but exchange is less frequent than that in hydration layer. Even though methanol has succeeded to reside into these conserved sites in pure methanol solvent but its distance with hydrogen bonding partners more than 5 Å with labile hydrogen bonding state.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

Chapter Page

TABLE OF CONTENTS (continued)

LIST OF ABBREVIAITONS

ACF	Auto Correlation Function
Atm.	Atmostpare
Calctd.	Calculated from this simulation research
cryst.H_2O	Crystallographic Water molecules
CytC	Cytochrome C
D-A	Donor – Acceptor
Diff. Const.	Diffusion Coefficient (Constant)
Expt.	Experimentally determined value
H_2O	Water
H-bond	Hydrogen Bond
hh-CytC	Horse Heart Cytochrome C
K	Kelvin
kDa	kilo-Dalton
LINCS	A Linear Constraint Solver for molecular simulation
Max.	Maximum value of data, upper limit of range
MD	Molecular Dynamics Simulation
MeOH	Methanol
Min.	Minimum value of data, lower limit of range
MSD	Mean Square Displacement
NPT	Constant Composition (N), Pressure (P) and Temperaure (T)
ns	Nanosecond

LIST OF ABBREVIAITONS(continued)

NVT	Constant Composition (N), Volume (V) and Temperature (T)
OPLS-UA	United Atom Model for Optimized Potential for Liquid Simulation
ps	Picosecond
RDF	Radial Distribution Function
R_g	Radius of Gyration
RMSD	Root Mean Square Displacement
RMSF	Root Mean Square Fluctuation
SASA	Solvent Accessible Surface Area
SD	Standard Deviation or Error
SPC/E	Simple Point Charge / Extended: Model of Water
VHDF	van Hove Distribution Function

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES(continued)

Figure Page

LIST OF FIGURES(continued)

Figure Page

LIST OF FIGURES(continued)

Figure Page

LIST OF FIGURES(continued)

Figure Page

LIST OF FIGURES(continued)

Figure Page

LIST OF FIGURES(continued)

Figure Page

(this page intentionally left blank)

CHAPTER I

INTRODUCTION

1. Structure and Function of Cytochrome C.

Cytochrome C is a globular heme protein, which acts as an electron carrier in the electron transport chain for the production of ATP in mitochondrial respiration. Cytochrome C is considered as an electron donor co-factor for the electron acceptor oxidizing proteins with which it forms an electron transfer complex [1]. The buried heme prosthetic group in hydrophobic environment is covalently bound to the peptide chain by two thioether linkages resulting from addition of polypeptide cysteine-sulfhydryl group to vinyl group of heme. The heme iron is low spin hexa-coordinated with four heme-pyrole ring nitrogens in a plane and axially by nitrogen of histidine and sulphur of methionine. Heme is exposed to solvent only slightly and iron is involved in a reversible redox reaction and its proper orientation, conformation, axial ligand geometry; and its proximity to solvent is governed by the polypeptide chain. It possesses a relatively high redox potential in the range of 0.15 V to 0.35 V with saturated Hydrogen electrode [1, 2, 5].

Continuous efforts have been made to comprehend the molecular or structural factors that control redox potential in CytC, which includes: first coordination sphere effect on heme iron, pi-electron acceptor character of the axially coordinated thio-ether sulfur atom of methionine, the interaction of the heme group with surrounding polypeptide chain and solvent molecules, medium effects related to the nature of solvents and physical environments. In electron transfer reactions, complexes formed between CytC and other proteins are stabilized

by highly specific electrostatic interactions in which the positively charged domain formed by clustered lysine residues surrounding exposed heme edge of CytC plays a key role. Replacement of HIS-MET by HIS-HIS axial coordination decreases by about 0.16V of the redox potential indicating coordinative effect on redox potential of CytC. So the electrostatic interactions at heme-protein-solvent interface and axial coordination of heme-iron are key factors to be monitored at different environments [2, 5, 8, 19]. In our study, we took horse heart Cytochrome C (hh-CytC) as a model Cytochrome C (CytC) or protein.

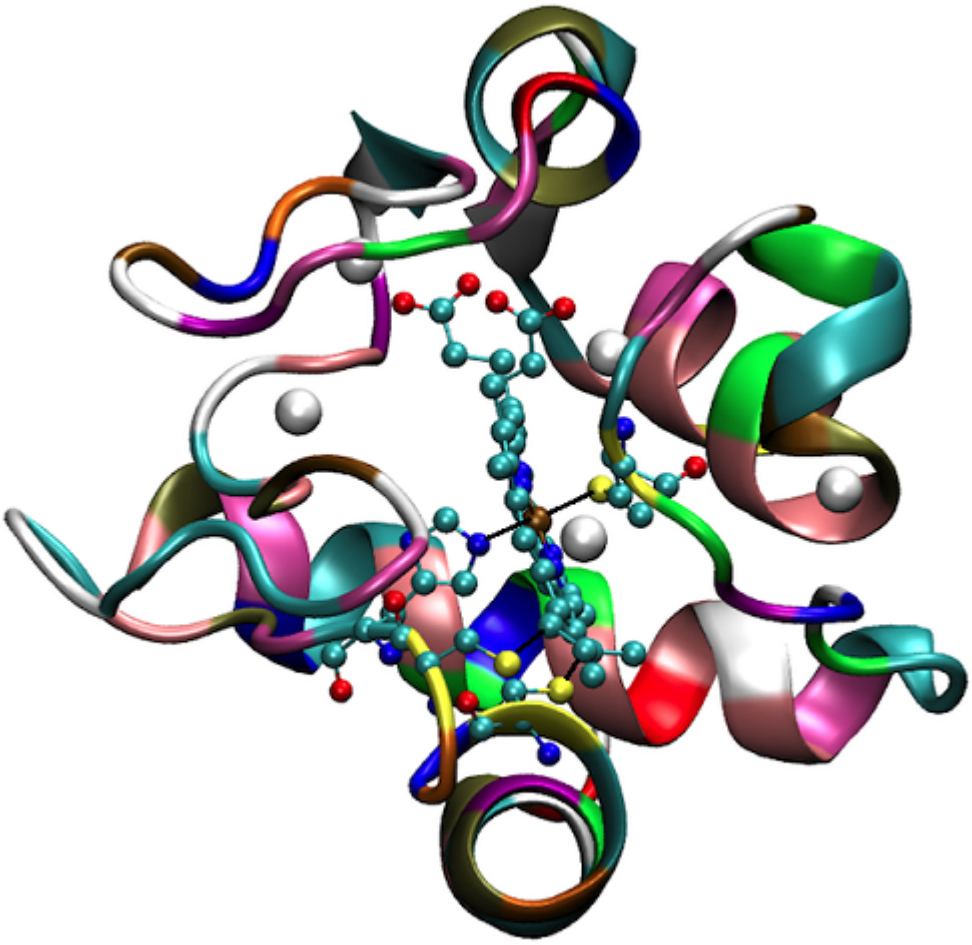


Figure-1: Structure of horse heart Cytochrome C (hh-CytC) (pdb code: *1HRC.pdb*) with five crystallographic water molecules [3].

2. Horse Heart Cytochrome C.

The horse heart Cytochrome C, hh-CytC (M_r = 12.4 *kDa*), is used as a model for the study of many biochemical phenomena *in silico* due to the availability of its high-resolution three-dimensional crystal structure, its relatively small size and its distinct structural properties [4, 8]. This protein consists of 104 amino acid residues in a single polypeptide chain with heme as 105th "residue". It consists of five α-helices with residues; 6-14(H_1), 49-54(H_2), 60-69(H_3), 70-75(H_4), 87-102(H_5) and two short double stranded anti-parallel β-sheets, 37-40(B1) and 57-59(B2) along with random coil structures and β-turns. Three Ω loops comprise residues: 20–35(Ω_1), 40–57(Ω_2), and 71–85(Ω_3). The two heme-protein thio-ether bonds are: CAB\(HEM)–SG(14CYS) and CAC(HEM)–SG(17CYS). The heme-Fe is axially coordinated with NE2(18HIS) and SD(80MET). It has 75 internal hydrogen bonds that comprise 45 H-bonds between main chain atoms, 20 H-bonds between main chain and side chain atoms, and 10 H-bond involving only side-chain atoms. Four salt-bridges, 5LYS-2ASP, 38ARG-105HEM, 53LYS-50ASP, 99LYS-61GLU are included in crystal structure. The heme is saddle shaped and overall only 7.5 % of the total heme surface is exposed to external solvents [3, 4, 6, 7].

Continuous research endeavors to harness nonaqueous enzymology and allied biochemistry is based on structural and conformational study of proteins and bio-molecules which are thermodynamically stable in non-water solvents [11, 12]. Aqueous mixtures of organic solvents are commonly used for such study. There are many factors governing the protein's flexibility, stability and their functions such as hydrogen bonding, electrostatic and van der Waals interaction, disulfide bonds, solvent polarity or hydrophobicity, nature of solubilization process, percentage of hydration and duration of exposure to solvents, and relative conformational entropy [8, 12, 13, 14]. From this perspective, hh-CytC in water-

methanol solvent systems, both in pure solvents and binary-mixture, were studied as a model protein to understand solvent mediated enzyme stability and activities in cosolvent/water mixture solutions [8, 13, 34, 61]. We have focused our research in understanding the sensitivity of water-methanol binary mixtures in hh-CytC internal structure and requirement of hydration in native structure of protein.

3. Structural and Hydration Water in Horse Heart Cytochrome C.

The protein's characteristics and functions are dependent on the degree of hydration. Water molecules interact with proteins on many length and time scales [12]. The optimum amount of water has been recognized as a controlling factor for nonaqueous protein activity. The water molecules used in hydration are associated with protein's structure and flexibility for proper functioning, and are supposed to act as lubricant to maintain flexibility for the protein in contact with organic solvent [13, 15]. On the other hand, at high level of hydration (or solvation), water and solvent molecules may bind to the active sites acting as inhibitors [13, 14]. So, the degree of intimacy or nature of interaction of solvent or water molecules in solvation or hydration shell with protein is one of the strategic research scopes to deal with solvent mediated protein biochemistry.

The high resolution X-ray crystal structure of hh-CytC includes a number of water molecules and some of these crystallographic water molecules occupied in apparently defined or conserved positions or hydration sites both exterior at the surface around polar and charged side chain as well as buried interior into both hydrophilic and hydrophobic cavities. These water molecules are regarded as essential or structural components for protein functioning [12, 13, 16]. The study of such structural water molecules is very important to reveal their role in

protein structure and function. Previous research has pointed out that the minimum number of water molecules may be associated with the proteins to maintain structural flexibility and functioning. These water molecules are usually named as ‘essential water’ or ‘biological water’ or ‘functional water’ of proteins [12, 13, 14, 16]. Simultaneously, on the other hand, we may think that these water molecules might be simply included as residual water molecules during the course of protein folding into native state.

To probe individual water molecules in hydration layer around protein or bound water molecules buried in the interior, and measure the lifetime of these bound water molecules, is very challenging experimentally. Many efforts have been devoted to elucidate the kinetic or mechanistic and structural characterization of these bound water molecules. The intermolecular H^1-H^1 NOEs between protein and water molecules and the chemical shift of exchange protons between water and amide groups of protein were used to identify protein-bound water molecules [15]. A diffusion controlled pulse field gradient NMR technique was attempted for the determination of lifetime of protein-bound water molecules but the limitation was that the exchange rate between bound and bulk water is relatively fast compared to interval of diffusion rate filter [16]. So the long molecular dynamics (MD) simulation serves as a very useful complement to experiments to study these protein-bound individual water molecules present in crystal structure which are inaccessible in NMR experiments [9, 14, 15, 16]. Simulation studies have been achieved to interpret the behavior of protein-solvent interface. Properties and lifetime of hydrogen bonding between protein and solvation layer, tendency of aggregation of solvent molecules around protein, and solvent hydrophobic interactions are considered important parameters to be analyzed. Most importantly, activity of water in enzyme hydration

for nonaqueous enzymology and the structure and energy of proteins [12, 14, 24] has been studied.

The hh-CytC has 124 crystal water molecules in its X-ray crystal structure [3]. An NMR study by Qi, P. X. *et al*, found that six water molecules in reduced hh-CytC and five water molecules in oxidized hh-CytC reside in these positions with more than a 100 picosecond residence time [6, 9]. Three water molecules in hh-CytC localized on the surface of protein are supposed to stabilize local polypeptide chain conformations. Two water molecules are located internally in the heme crevice. One water molecule mediates a charged interaction between the residue 38ARG and a heme propionate. The other water molecule is more centrally buried near the heme iron and is hydrogen bonded to the side chain of the conserved triad residues 52ASN, 67TYR and 78THR [3, 6, 9]. This water molecule is also within hydrogen bonding distance of 75ILE in reduced hh-CytC and it is found to undergo a large positional change consistently with change of oxidation state, and this is intimately linked to the value of the redox potential of heme [3]. Moreover, the detection and investigation of long lived, bound water molecules in hh-CytC by common NMR techniques has been cumbersome due to technical problems in resolving bound and bulk solvent water. The NOE experiment in aqueous solution of hh-CytC indicates that there must be a water molecule located near 57ILE, 63THR, 64LEU and 74TYR, which appears to be hydrogen bonded to the ring hydroxyl of 74TYR and the hydroxyl of 63THR. Another water molecule is located in a turn region containing 36PHE and 37GLY and is hydrogen bonded with the amide NH of 61GLU, 64LEU and 65MET. So, the aim of MD simulation study of these bound water molecules of hh-CytC is necessarily important to elucidate their role in understanding structure and function of biomolecules. In addition, the detailed study of lifetime of these water molecules in bound state and their dynamics at the

vicinity of specified conserved positions certainly might have profound insight in the mechanism of protein folding [3, 6, 9, 19].

4. Structure of Water-Methanol Solvent Mixture.

4.1 Structure of Water.

Water molecules have special ability to form hydrogen bonding networks. Since microscopic forces that define water structure are not exactly understood, many different potential functions for the water monomer and liquid water have been developed so that many anomalies and complex properties of water can be explained. One of the simple and popular models of water molecule is SPC/E (Simple Point Charge, Extended) model. We are using this model of water in our MD simulation because of its ability to reproduce dynamics and macroscopic properties of system. It is a three interactions site rigid model of water with constrained bonds and angles [10, 18].

4.2 Structure of Methanol.

Methanol has a hydrophobic methyl (CH_3) group in place of one hydrogen in structure of water due to which it exhibits many amphiphilic properties. In GROMACS, the methyl (CH_3) group is treated as a single particle in interaction with other atoms or molecules. So methanol is also represented by a three interactions site model with OPLS-UA force field [18, 21].

4.3 Water-Methanol Solvent Mixture.

The water-methanol solvent mixture has a long history of increasing theoretical and experimental interest to study its anomalous properties, because of the degree of mixing of

water and methanol in microscopic and macroscopic level. It is a good solvent for many amphiphilic solutes. For water-methanol mixture, thermodynamic properties such as entropy increase, while compressibility and mean molar volume are smaller than what would be expected for the ideal mixture of pure liquids. The main reason for the unusual behavior of water-methanol solvent mixture is not because of clustering of water due to hydrophobic effect of methyl group, but because of incomplete mixing of water and methanol in molecular level due to hydrophobic segregation [22]. In mixture, the ordering of methanol molecules bury the hydrophobic methyl group closer inside or away from water phase and push hydroxyl group apart outside resulting in the formation of micro micelles and retention of the hydrogen bonded network structure of the bulk water. Water molecules bridge the chain or ring of 6-8 methanol molecules in cluster and this clustering network is also observed in pure liquid methanol [21]. Since the binary mixtures of water and its cosolvents have always been applied as promising media in terms of their large change in physical and chemical behaviors that they exhibit compared with the individual components [8, 13, 14], water−methanol is one of such very important binary solvents, widely used in biology as experimental solvent because of its unique and non-ideal behaviour [23, 51, 62]. The water−methanol binary mixture is famous for exhibiting striking anomalies at various concentrations that essentially arise due to structural transformation of methanol through hydrophobic as well as hydrophilic interactions. The recent simulations have shown that both water and methanol molecules lose entropy in mixture where the rotational entropy is the more contributing factor but methanol molecules lose their entropy three times more than the water molecules do. Such nanoscale clustering of methanol molecules in water methanol mixture supports the concept of heterogeneous or incomplete

mixing of water and methanol at a molecular level [22, 23, 34, 56]. So water-methanol solvent mixture could be an appropriate environment for simulation study of hh-CytC.

Table-1: Characteristics of SPC/E H_2O and OPLS-UA MeOH Molecular Models [20, 21, 22].

Molecular Model Properties	**Water (H_2O)**	**Methanol (CH_3OH)**
Model	SPC/E	OPLS-UA
Dipole (Debye)	2.3900	1.690
r_{OH}(Å)	1.0000	0.945
r_{Me-O}(Å)	--------	1.430
σ_O(Å)	3.5533	3.070
σ_{Me}(Å)	---------	3.775
σ_H(Å)	0.000	0.000
ε_O(Kcal/mol)	0.1553	0.170
ε_{Me}(Kcal/mol)	---------	0.104
ε_H(Kcal/mol)	0.000	0.000
Angle, H-O-H(Θ°)	109.471	----------
Angle, Me-O-H(Θ°)	----------	108.50
$+q_H$(e)	+0.4238	+0.4350
$-q_O$(e)	-0.8476	-0.7000
$+q_{Me}$(e)	----------	+0.2650

5. Protein-Solvent Interaction.

Dynamics of protein-solvent interactions are fundamental in conformational fluctuations and concerted movements of proteins for the accomplishment of crucial physiological functions, but their investigation is still experimentally very demanding [24]. The function, specificity and efficiency of a protein can be tuned by changing the solvent properties. The major interactions include hydrogen bonding and van der Waals interaction, hydrophobic

interactions, and electrostatic interaction. So, molecular dynamic simulation study of protein in different solvent environments has important scope in theoretical understanding of structural and conformational landscape of protein. The enzyme – substrate accessibility also depends on the solvents in homogenous catalysis. On the other hand, protein influences the structure and dynamics of surrounding solvent molecules as observed in the ordered water around polar and charged side-chains in X-ray crystallography. These hydration waters make comparable contributions to the structure and energy of proteins. The coupling between fast hydration dynamics and protein dynamics is considered to have an important role in protein folding (12, 24). In addition, the recent approach to use organic solvents for nonqueous enzymology is still in intensive research both experimentally or *in silico* even though the organic solvents have different ability to stabilize the polar transition states and are found to bind the active site acting as inhibitor [13, 14, 24, 56]. So simulation study of water-methanol mixture around hh-CytC might be a further research in understanding of interactions between protein, organic solvent, and hydration layer.

6. Molecular Dynamic Simulation.

6.1 General Perspective.

With the continuing growth of computing power of super-computers, computation based on molecular models and computer simulation is playing a valuable role increasingly by providing essentially complementary results in idealized or extreme conditions that are inaccessible in real experiments as well as guiding course of research avoiding wasteful trial and error methods. It provides a direct route from the microscopic detail of a system, such as atomic properties, intermolecular interaction between them and molecular geometry, to

macroscopic properties of experimental interest, such as thermodynamic parameters, transport coefficients and other dynamic and functional properties [25, 26, 27]. The *in silico* experiments not only evaluate average properties but also provide structural and temporal resolution of any definable quantities, for example, conformational distributions or interactions between parts of systems [26, 29, 30].

One of the simulation methods is the *'classical' molecular dynamic simulation,* MD simulation, which describes the dynamics of atoms and molecules of system based on Newton's second law of motion. Molecules are treated as classical objects resembling very much the 'ball and spring' model. Atoms or certain groups correspond to soft balls and elastic springs correspond to bonds angles and torsions between them [29]. In MD Simulation, the time evolution of dynamics of interacting particles is followed via the numerical, step by step in femtoseconds, solution of the classical equation of motion (eq-1) as

$$F_i = m_i \frac{d^2}{dt^2} r_i(t) \qquad \text{[eq-1]}$$

where $r_i(t) = [x_i(t), y_i(t), z_i(t)]$ is the position vector of i^{th} particle and F_i is the force acting upon ith particle at time t, and m_i is the mass of the particle[25, 27]. The force required for i^{th} particle's motion is calculated from *its approximate interaction potential energy functions (U) defined as force field* (eq-2) in the system as follows

$$F_i = -\nabla_i U(r_1, \ldots\ldots, r_N) = -\left(\frac{\partial U}{\partial x_i}, \frac{\partial U}{\partial y_i}, \frac{\partial U}{\partial z_i}\right) \qquad \text{[eq-2]}$$

6.2 Molecular Dynamics Algorithm: Numerical Integration of the Equations of Motion.

The dynamics or time evolution of positions (coordinates) and velocities of interacting particles, which are called trajectories, are computed by integration with femtosecond time step

dt of above differential equation of motion (eq-1) successively over a long period of time starting from randomly assigned initial velocities from a Boltzmann distribution at the desired temperature. The propagation of positions and velocities of particles over a finite time or the nature of trajectory depends on the numerical integrators (method of integration) used [29, 30, 31].

The numerical integrators compute successive updated positions $\mathbf{r}_i(t + \Delta t)$ and velocites $\mathbf{v}_i(t + \Delta t)$ at $t + \Delta t$ of each particles in the system based on their initial positions and velocities. One of the symplectic integrator algorithms is *'leap-frog algorithm'*. It computes the updated positions and velocity at interleaved time points, staggered in such a way that they 'leapfrog' over each other, as follows

$$r_t = r_{t-\Delta t} + v_{(t-\frac{\Delta t}{2})}.\Delta t \qquad \text{[eq-3a]}$$

$$r_{t+\Delta t} = r_t + v_{(t+\frac{\Delta t}{2})}.\Delta t \qquad \text{[eq-3b]}$$

$$v_{(t+\frac{\Delta t}{2})} = v_{t-\frac{\Delta t}{2}} + a_t.\Delta t \qquad \text{[eq-3c]}$$

$$v_{t+\frac{2\Delta t}{3}} = v_{(t+\frac{\Delta t}{2})} + a_t.\Delta t \qquad \text{[eq-3d]}$$

$$r_{t+\Delta t} = r_t + \left(v_{t-\frac{\Delta t}{2}} + a_t.\Delta t\right)\Delta t \qquad \text{[eq-3e]}$$

$$r_{t+\Delta t} = r_t + v_{(t-\frac{\Delta t}{2})}.\Delta t + a_t\,\Delta t^2 \qquad \text{[eq-3f]}$$

Using Euler's velocity approximation at initial condition

$$v_{t\pm\frac{\Delta t}{2}} = v_t \pm \frac{1}{2}\,a_i\Delta t \qquad \text{[eq-4]}$$

It computes position and velocity as follows

$$r_{t+\Delta t} = r_t + \left\{v_t - \tfrac{1}{2}a_t\Delta t\right\}\Delta t + a_t\,\Delta t^2 \qquad \text{[eq-5a]}$$

$$r_{t+\Delta t} = r_t + v_t.\Delta t + \tfrac{1}{2}a_t\,\Delta t^2 \qquad \text{[eq-5b]}$$

and

$$v_{(t)} = v_{t+\frac{\Delta t}{2}} - \frac{1}{2}.a_t.\Delta t \qquad \text{[eq-5c]}$$

So, leapfrog integration being second order and symplectic in nature, conserves the energy of dynamical systems and minimizes the errors associated with global properties. Moreover, it has the property of time reversibility where integration in *n-forward* time steps and *n-reverse* integration gives the same positions.

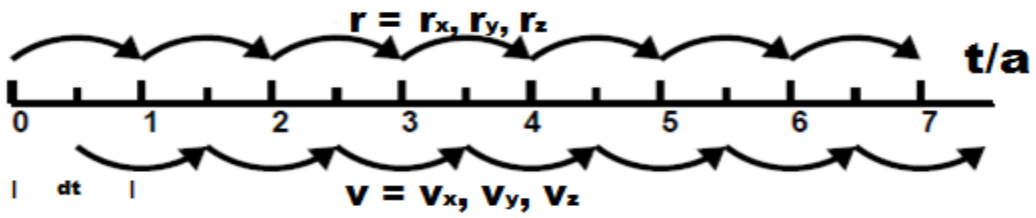


Figure-2: Illustration of Leap-frog Integration Model (edited google image).

In practice, *Δt* correspond to integration step and is determined by the fast motions in the system. Bonds involving light atoms such as O-H and N-H bond vibrate with periods of several femtosecond scales implying that *Δt* should be on a sub-femtosecond scale to ensure stability of the integration even though the fastest and not crucial vibrations can be eliminated by imposing constraints on the bond length in the integration algorithm. A long trajectory

constitutes a series of ensembles of molecular configurations (coordinates and velocities) saved at calculated time intervals over the entire simulation.

6.3 Force Field Model of Molecular System.

A model force field describes a collective sum of potential energy terms of all possible interactions among atoms and molecules as a function of their position ($r_i = x_i, y_i, z_i$) in a physical system including both bonded and non-bonded interactions along with their shape and geometry effects. The never ending attempt exist in finding a realistic force field or potential functions that would adequately mimic the true physical system yielding computed parameters in agreement with experimental results. A typical force field used in the molecular dynamic simulation takes the following form,

$$U(r_1, \ldots\ldots, r_N) = \sum_{bonds}^{N} \frac{1}{2} K_b \left(b_i - b_{0,N}\right)^2 + \sum_{angles}^{N} \frac{1}{2} K_\theta (\theta_i - \theta_{i,0})^2 +$$

$$\sum_{\substack{improper \\ dihedral}}^{N} \frac{1}{2} K_\omega \, (\omega_i - \omega_{i,0})^2 + \sum_{\substack{torsions, \\ proper\ dihedrals}}^{N} [1 + cos(n\, \varphi_i - \delta)\,] +$$

$$\sum_{\substack{non\ bonded \\ interactions}}^{atom\ pairs} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] + \sum_{\substack{electrostatic \\ interaction}}^{atom\ pairs} \frac{q_j q_i}{(4\pi\epsilon_o \epsilon_r\, r_{ij}}$$

[eq-6]

In this equation (eq-6), the first four terms give potential of bonded interactions defined by the covalent structure of the system, the second to last term gives non-bonded van der Waals interaction potentials between atom pairs separated by the distance, r_{ij}, and the last term gives electrostatic potential between same pairs.

The applicability of a force field depends on many factors. How accurately the potential-energy terms are formulated and parameterized for non-bonded interactions is very crucial to simulate a large system with sufficient accuracy and thermodynamic compatibility. Moreover, transferability of force field parameters over the varieties of chemical compounds without compromising accuracy, efficiency and reliability in simulation is the other side of the coin for the acceptance of a force field.

GROMOS is a widely used force field to simulate biomolecular systems for a long period of time. This force field is based primarily on reproducing the free energies of hydration and non-polar solvation for a range of compounds including amino acids and small peptides. The most recent force field, *gromos53a6*, is optimized by adjusting partial charges and fitting to reproduce the thermodynamic properties of pure liquids of a range of small polar molecules and the free energies of amino acid in water. The force field consists of following potential energy terms [18, 31].

(I) Bonding Potentials:

$$V_{bond}(r, K_b, b_0) = \sum_{n=1}^{N_b} \tfrac{1}{4} K_{b_n} (b_n^2 - b_{0_n}^2)^2 \qquad \text{[eq-7a]}$$

$$V_{angle}(r, K_\theta, \theta_0) = \sum_{n=1}^{N_\theta} \tfrac{1}{2} K_{\theta_n} [Cos\,(\theta_n) - Cos(\theta_{0_n})]^2 \qquad \text{[eq-7b]}$$

$$V_{\substack{improper \\ dihedrals}}^{harmonic}(r, K_\omega, \omega_0) = \sum_{n=1}^{N_\omega} \tfrac{1}{2} K_{\omega_n} [\omega_n - \omega_{0_n}]^2 \qquad \text{[eq-7c]}$$

$$V_{\substack{Trignometric \\ Dihedrals}}^{Proper\ torsional}(r, K_\varphi, \delta, m) = \sum_{n=1}^{N_\varphi} K_{\varphi_n} [1 + Cos(\delta_n)\, Cos(m_n \varphi_n)] \qquad \text{[eq-7d]}$$

(II) Non-bonding Potentials:

$$V_{pairs}^{Coulomb}(r, q\ ) = \sum_{pairs,i,j}\left[\frac{q_i q_j}{4\pi\epsilon_0\epsilon_m}\frac{1}{r_{ij}}\right] \quad \text{[eq-8a]}$$

$$V_{L-J}^{Non-bonded}(r, C_{12}, C_6) = \sum_{pairs,i,j}\left[\frac{C_{12}(i,j)}{r_{ij}^{12}} - \frac{C_6(i,j)}{r_{ij}^{6}}\right] \quad \text{[eq-8b]}$$

$$V_{Reaction\ Field,RF}^{Dipolar} = \sum_{pairs,i,j}\left[\frac{q_i q_j}{4\pi\epsilon_0\epsilon_m}\frac{\left(-\frac{1}{2}C_{rf}r_{ij}^2\right)}{R_{rf}^3}\right] \quad \text{[eq-8c]}$$

In the four body covalent terms, GROMOS has included proper and improper types of dihedrals separately. The non-bonded interaction due to reaction field is included separately. Only non-bonded interaction of solvent molecules is considered keeping the intra-molecular degrees of freedom frozen [18, 30, 31].

7. Scope of Research.

This research is a part of a **Master's Degree** in chemistry. Considering a stipulated time of completion, this research has the following objectives:

- Application of classical molecular dynamic simulation using *GROMACS* software [18] for large system simulation for nanosecond time scale.
- Study of the structure, stability and flexibility of horse heart Cytochrome C (using *gromos53a6* force field) in aqueous methanol solutions of varying compositions.
- Study of the solvation process at the atomic level by examining the solvation structure of water-methanol mixture around hh-CytC.

- Dynamic study of crystallographic water molecules present at the surface of hh-CytC or buried internally.
- Study of hydrogen bond dynamics in different solvent systems.

CHAPTER II

SYSTEM SETUP, METHODS OF SIMULATIONS AND ANALYSES

Simple and fundamental choices were made to perform molecular dynamic simulations in laboratory conditions: 298.15 Kelvin temperature and 1 atm pressure at constant NPT. Minimum sample sizes and cost of simulation were determined based on source of literature information and software [18].

1. Horse Heart Cytochrome C in Our System.

The X-ray crystallographic structure of horse heart cytochrome C having resolution 1.90 Å was used as initial structure for this study [3]. The structure's file code *1HRC.pdb* corresponds to ferric protein and was downloaded from PDB data bank website. Along with this crystal structure of hh-CytC, 124 water molecules were included as crystallographic structural water of hh-CytC in well-defined positions. These crystallographic structural water molecules were retained in every simulation. The N-terminus was deacylated and hydrogen was added to the N-terminus. With the histidine residues assumed to be neutral, the protein has a net +7 charge with +2 charge on heme. The hh-CytC was kept in the center inside a periodic cubic box of solvent. The systems were neutralized by adding seven chloride ions, which replace either water or MeOH molecules.

The topologies of both protein and heme are merged so that we can get both parts as single molecule. The missing bonds and angles were assigned very carefully based on '*gromos53a6*' force field and experimental results from literature so that the protein and heme

connectivity remained intact throughout simulation as contacts between protein and heme are essential for stabilizing the native structure [5, 15,18, 32].

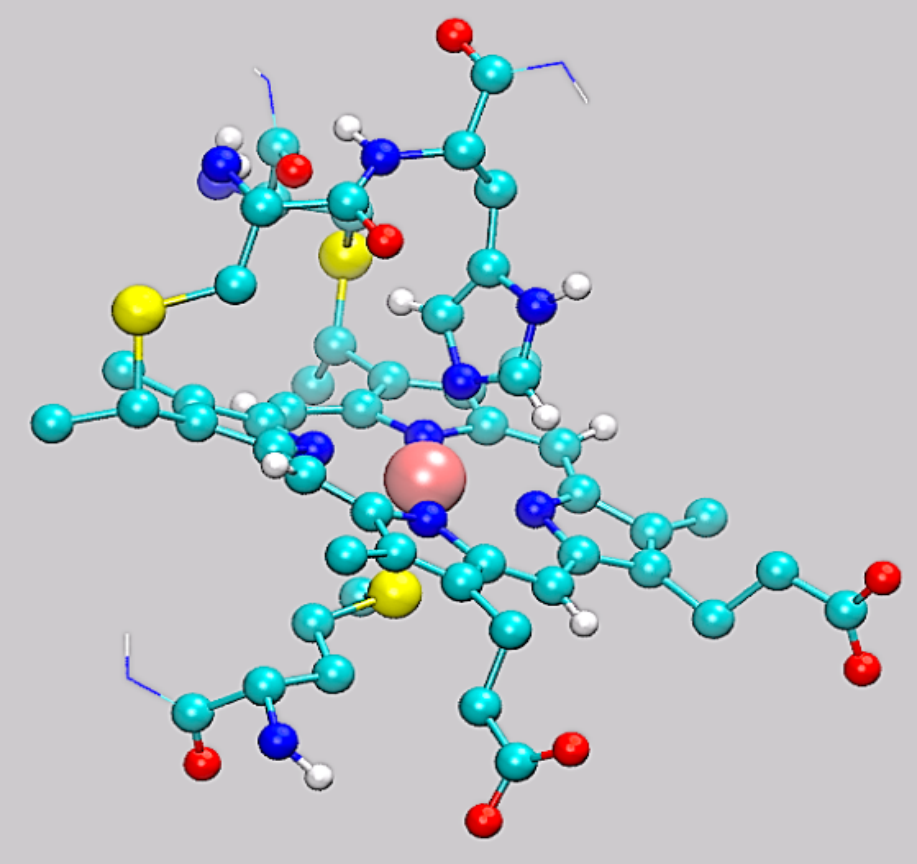


Figure-3: Protein-Heme linkages in hh-CytC; two thioether bonds between vinyl-carbon of heme and cysteine-sulphur, and axial coordination between Fe-S (80MET) and Fe-N (side chain ring N of 18HIS).

The energy of axial coordination bond between **Fe** (iron of heme) and **S** (80MET-SG of protein chain) was assigned 250 kcal/mol with reference to '*charmm*' force field included in GROMACS and based on the research of *Prabhu et al* [32]. The other axial coordination bond between **N** (side chain pyrole ring of 18HIS-NE2 of protein chain) and Fe was kept as assigned by '*gromos53a6*' for N – Fe, which is 249 kcal/mol (gb_34). The covalent bonds **S – C** between **S** (of 14CYS and 17CYS) and **vinyl carbon** (of heme prophyrin ring); viz.

(14CYS)SG—CAB(HEM) and (17CYS)SG—CAC(HEM); were assigned from GROMACS based on their bond lengths, 1.75 Å and 1.86 Å, which are gb_31 and gb_32 respectively. The bonds were found stable throughout our simulation. The angles and torsions associated with these bonds of protein and heme connections are observed using VMD and are also assigned according to '*gromos53a6*'.

2. Preparation of Simulation System: Boxes of hh-CytC in Different Solvent Compositions.

The SPC/E force field for water molecule and OPLS-UA force field for methanol molecule, were used in this molecular dynamics simulation work. Geometries of single molecules of H_2O and MeOH were obtained using MOLDEN software program and parameters of their respective model. Each H_2O and MeOH molecules are minimized separately to get the best optimum structures based on GROMACS simulation software. Then the cubic boxes containing hh-CytC in the center were filled with MeOH and H_2O molecules. Table-2 gives the details of compositions of solvents used in different simulation systems. The structure of solvent boxes were checked and observed using VMD. The number of molecules of H_2O and MeOH required to fill the spaces ($V_{solvent} = V_{Box} - V_{hh\text{-}cytc}$) were calculated as shown in Appendix-VI from experimentally determined density of solvent mixtures of required proportions and molar mass of water and methanol calculated in their pure form at 20 ^{0}C temperature (293.15 K) and 1 atm pressure [35].

Table-2: Structure of different systems with and without hh-CytC prepared for Simulation.

Solvent	hh-CytC	Cryst.H_2O Molecules	Added Water	Added MeOH	CL^- ions	System size (nm^3)
Solvent Only Simulation Boxes						
Water	0	0	6576	0	0	$(5.815)^3$ 196.661±0.197
20%MeOH in Water	0	0	5585	785	0	$(5.997)^3$ 215.71±0.421
40%MeOH in Water	0	0	4049	1518	0	$(6.005)^3$ 216.517±0.237
MeOH	0	0	0	3076	0	$(5.901)^3$ 205.522±0.211
hh-CytC with all crystallographic water molecules						
Water	1	124	7339	0	7	$(6.206)^3$ 239.098±0.547
20% MeOH in Water	1	124	5098	718	7	$(6.007)^3$ 216.73±0.534
40% MeOH in Water	1	124	3696	1388	7	$(6.015)^3$ 217.604±0.577
MeOH	1	124	0	2931	7	$(5.983)^3$ 214.112±0.821
hh-CytC without any crystallographic water molecules						
Water	1	0	7339	0	7	$(6.166)^3$ 234.429±0.512
20%MeOH in Water	1	0	5098	718	7	$(5.967)^3$ 212.508±0.873
40%MeOH in Water	1	0	3696	1388	7	$(5.980)^3$ 213.887±0.597
MeOH	1	0	0	2931	7	$(5.952)^3$ 210.889±0.597

3. Horse Heart Cytochrome C in a Solvent Box.

The minimum size of a cubic box required for simulation of hh-CytC in a solvent was calculated keeping hh-CytC at the center of the box with 0.78 nm distance from the wall of the box which should be at least less than half of the cut-off distance ($d_{protein\text{-}box\ wall} \geq$ half of $r_{cut\text{-}off}$, $r_{cut\text{-}off}$ = 1.4 nm) for successful periodic boundary condition. This gave about 5.776 nm^3 cubic boxes. The average diameter of the hh-CytC is 4.201 nm, which occupies volume 49.153 nm^3 (= 3.891 nm × 3.444 nm × 3.668 nm) at the center of the box. So the volume available for solvent around the hh-CytC is 143.547 nm^3. Then the box was filled with calculated numbers of SPC/E H_2O and OPLS-UA MeOH molecules. So our simulation system is a cubic box of solvent with single protein molecule, hh-CytC at the center of box as shown in Figure-4.

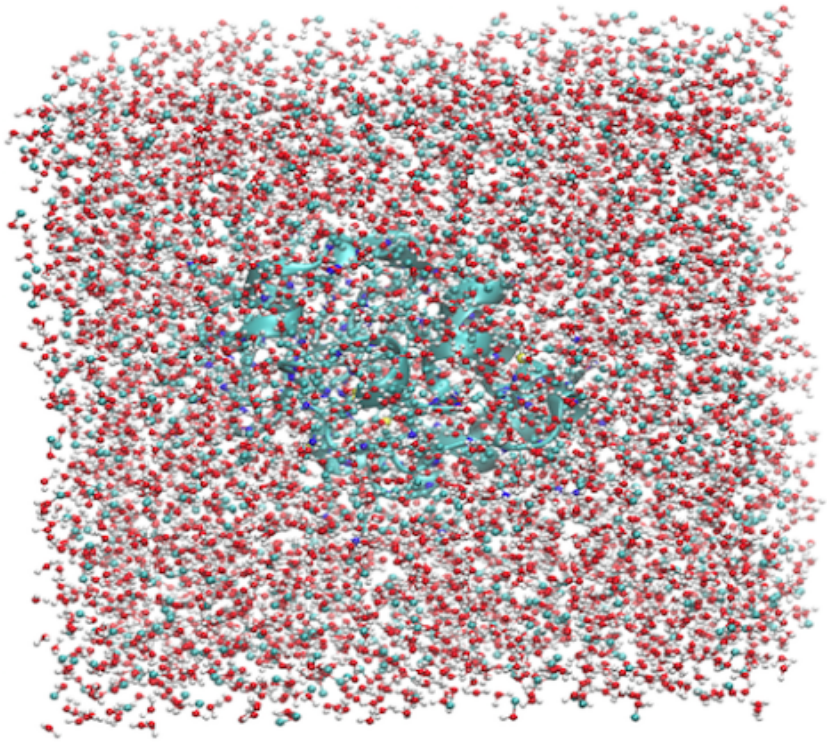


Figure-4: Horse heart Cytochrome C (hh-CytC) is centered in a cubic simulation box of water-methanol solvent system.

4. Molecular Dynamics Simulation Parameters.

In GROMACS-4.5.5 version with '*gromos53a6*' force field, the '*.mdp file*', containing all required simulation conditions or parameters, was prepared for different steps [18]. Appendix-V summarizes these simulation parameters in different steps. The parameters, which are not mentioned here, are all default parameters of GROMACS required for the simulation. The leapfrog integrator '***md'*** was set as the integration algorithm for successive solution of Newton's equation over 2 femtoseconds time steps. Neighbor group searching was set in grid for non-bonded energy calculations; and for the calculation of energy of electrostatic interaction, periodic boundary condition in three dimensions and Particle-Mesh-Ewald summation with grid spacing of 0.1 nm with an interpolation order of 4 were used. The '*cut-off* ' scheme was used for estimation of both short and long range interactions with 1.4 nm cut-off distance. A non-bonded pair-list was updated at every 10 time-steps. Different energy groups were made for different molecules in the system; viz. chloride ion, methanol, water, protein and heme groups. Berendsen thermostat and barostat was used to set up temperature and pressure respectively with 0.2 ps coupling time constant. Each group was coupled separately with the thermostat. Isotropic pressure coupling was set with reference pressure of 1 atm. The LINCS (Linear Constraint Solver) algorithm was used to keep bonds involving hydrogen atoms at their equilibrium length; other bonds were kept flexible [19]. The last structure of the system of the previous step of simulation was the initial or starting structure for the next step of simulation. All the simulation studies were performed in NPT condition at constant temperature 298.15 Kelvin and 1 atm pressure equivalent to laboratory condition. All the simulations more than 1 nanosecond were achieved in 'High Performance Computing Center (HiPCC) – Wichita State University using 32 processors.

5. Energy Minimization of a Simulation Box.

Energy minimization was performed using the steepest descent algorithm method to relax positions of solvent molecules inside the box, while keeping the position of hh-CytC constrained [18, 36]. It removes clashes between atoms that were too close. In the first energy minimization, a 2 femtosecond time step was used with 1000 kJ mol^{-1} nm^{-1} tolerance. After the first energy minimization, the interaction of solvent molecules with each other and with the protein molecules may open vacancies for additional solvent molecules. So the procedure of adding solvent molecules was attempted repeatedly by checking with VMD until there was no more empty space in the box [17, 18]. For the mix solvents, the size of box containing only hh-CytC was increased by one percent and all calculated numbers of solvent molecules were added inside the box and energy minimization was performed. This allows addition of required number of solvent molecules inside the box, as mentioned in Table-2, which maintain its correct size in the next step of equilibration. The stepwise energy minimization process was repeated with the same conditions by decreasing tolerance stepwise to 100 kJ mol^{-1} nm^{-1} until the system stopped in minimum energy convergence. Figures – 5a & 5b illustrate the potential energy profile in first and last equilibration steps.

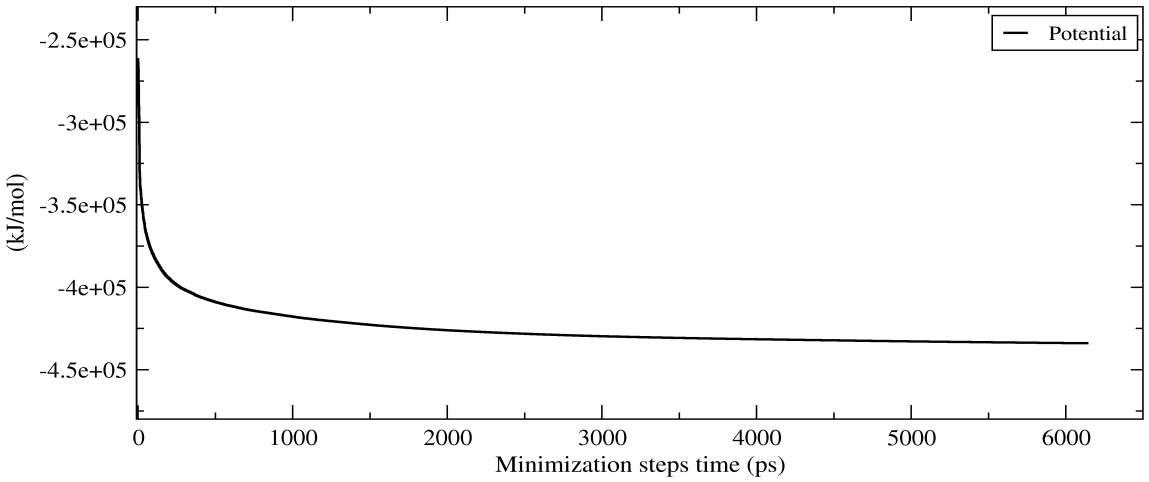


Figure-5a: An illustration of relaxation of Potential Energy of system in first energy minimization step by steepest descent method.

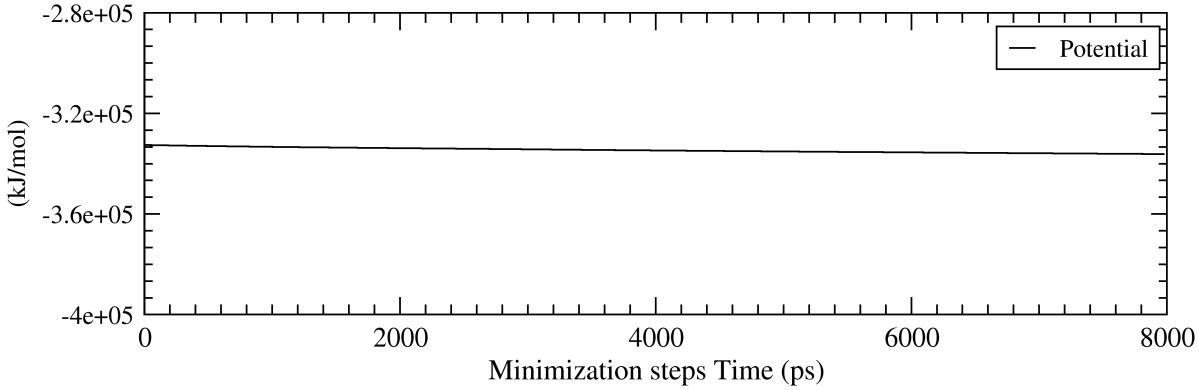


Figure-5b: An illustration of relaxation of Potential Energy of system in second energy minimization step by steepest descent method.

6. System Heating and Equilibration.

6.1 First Equilibration Step: Heating of Simulation System.

After energy minimization of solvent box, the system was heated slowly to thaw crystal protein to desired simulation temperature using programmed linear temperature ramp starting

from 20 K which is achieved by simulated annealing [18, 28]. Since there is no velocity in initial structure from energy minimization; for different ensembles of the same system, molecular velocity was generated randomly at 20 Kelvin and warmed slowly to 298.15 K in 31 steps keeping the position of hh-CytC restrained. The five different temperature groups were coupled to the thermostat separately. Since this step was NPT md-run, due to the isotropic coupling with the barostat and thermostat, the system achieves its size and reference pressure, 1 atm, with scaled reference coordinates along with programmed temperature in 200 picosecond time. This method of warming or equilibration of system is popularly known as position restrained molecular dynamics simulation. Figures - 6, 7 & 8 illustrate the temperature, potential energy and volume equilibration respectively in the first equilibration step.

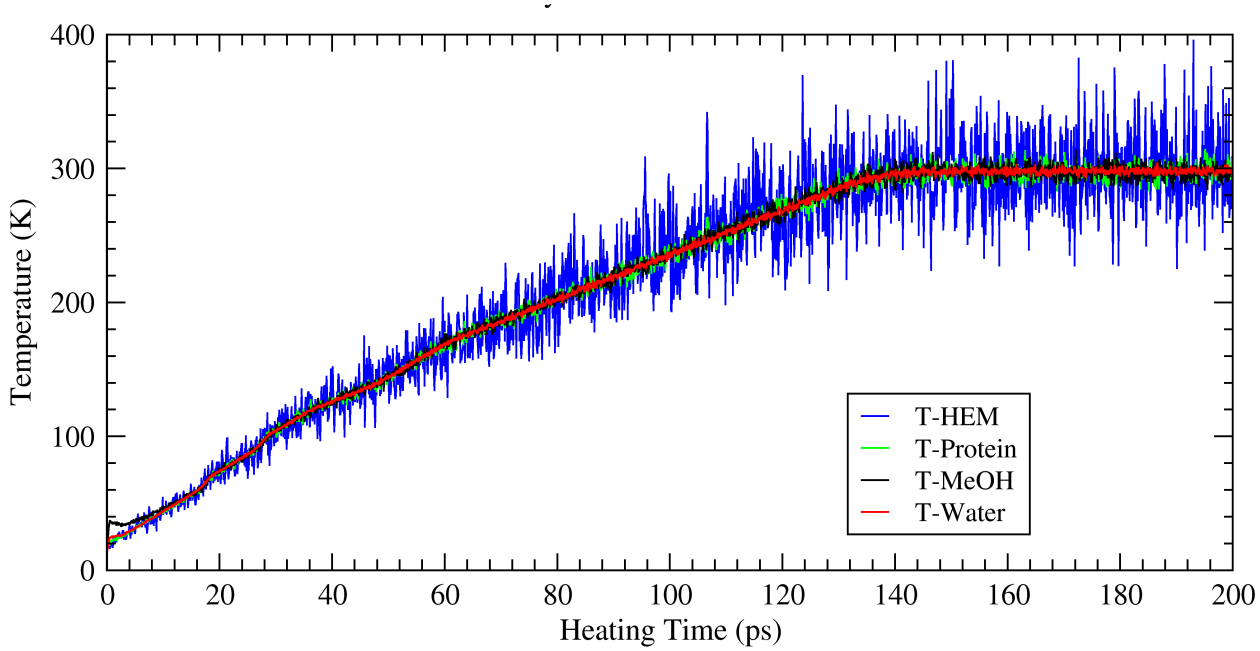


Figure-6: An illustration of programmed slow heating of a simulation system, hh-CytC in 20% aqueous MeOH solvent, in first equilibration or warm up and heating step from 20 K initial temperature to 298.15 K simulation temperature at 1 atm. pressure.

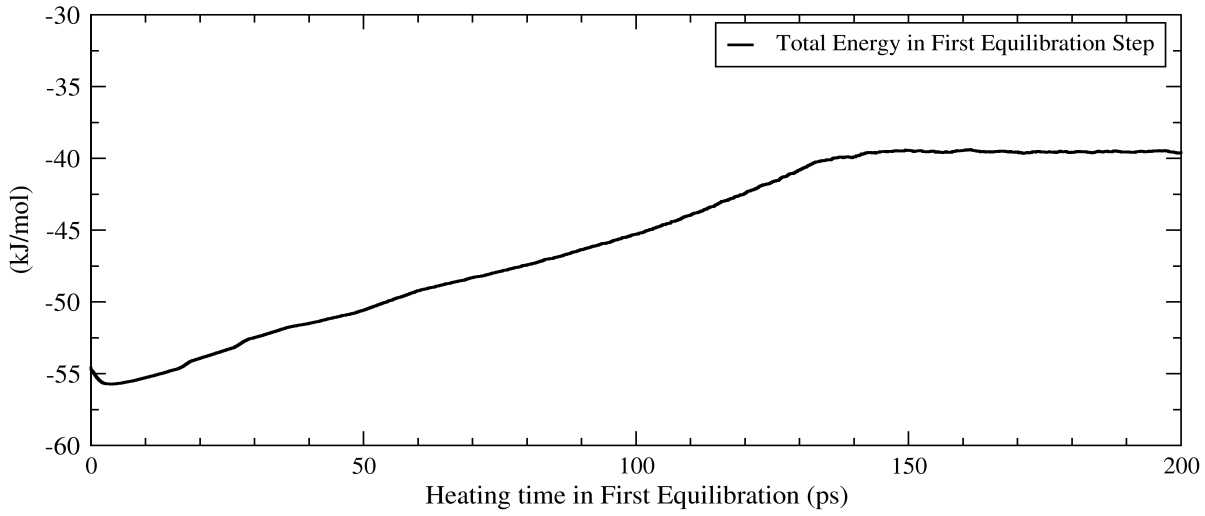


Figure-7: An illustration of equilibration of total energy of a simulation system, hh-CytC in 20% aqueous MeOH solvent, in first equilibration or warm up and heating step.

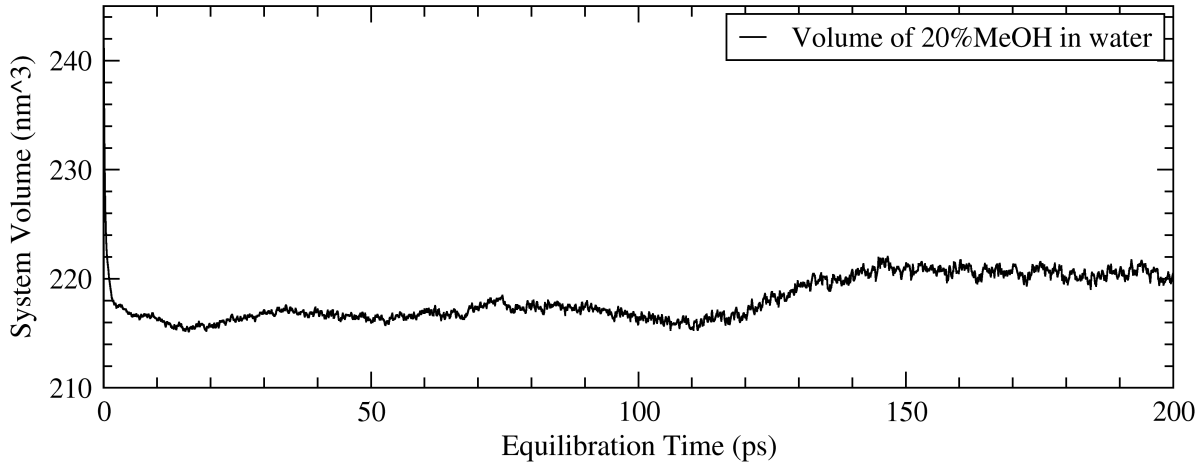


Figure-8: An illustration of equilibration of volume of a simulation system, hh-CytC in 20% aqueous MeOH, in first equilibration or warm up and heating step.

6.2 Second Equilibration Step.

In the second equilibration step, the warmed up system was equilibrated again at 1 atm pressure and 298.15 K for 200 ps. The X-ray crystal structure gives the average structure of

protein from the map of electron probability density. This provides an initial guess for a set of atomic position of protein but not an equilibrated structure in solution [37]. So, in this equilibration step, hh-CytC is let to move freely so that it can interact and adjust with solvent molecules and achieve an equilibrated conformation. This step gives the starting point of our exploration and makes the system ready for a data production run so that we achieve convergence of simulation results. It has all the same simulation conditions and md simulation parameters that are in the next long data production simulation step. The criteria of equilibration were set to have a plateau of RMSD (after least square fit removing rigid body rotational and translational motion) with constant fluctuations in protein conformations after ceasing initial non-equilibrium motion due to the initial structure, and constant PV and energy profiles which are independent of simulation time in NPT condition. Figures - 9, 10 & 11 illustrate the temperature, total energy and PV energy profiles respectively in second equilibration steps.

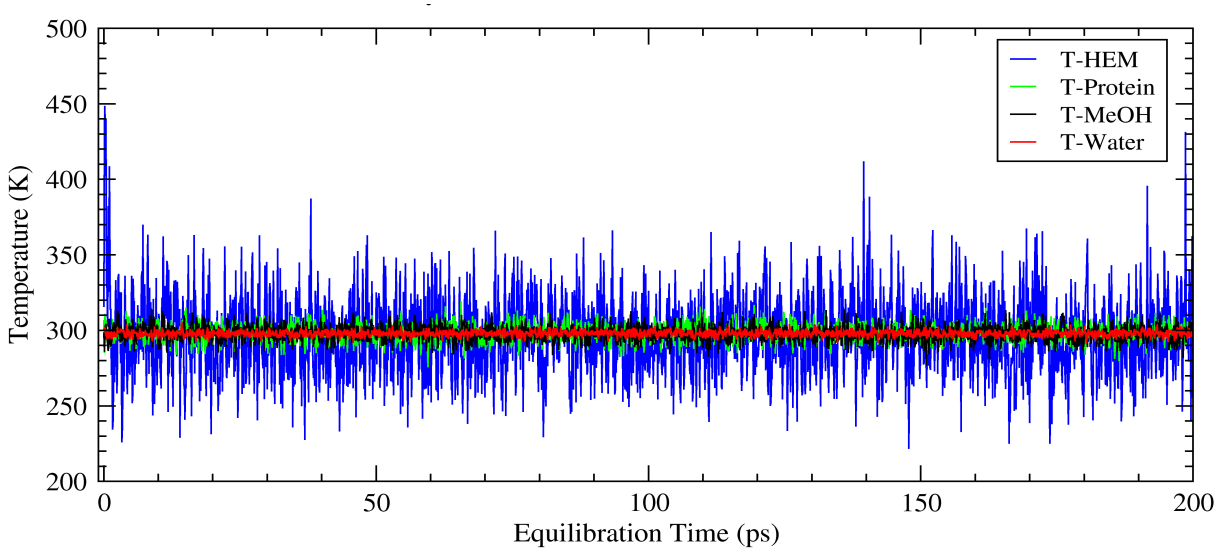


Figure-9: Illustration of equilibration of temperature after heating in a system of hh-CyC in 20% aqueous MeOH solvent for different assigned groups in second equilibration step at 298.15 K and 1 atm pressure.

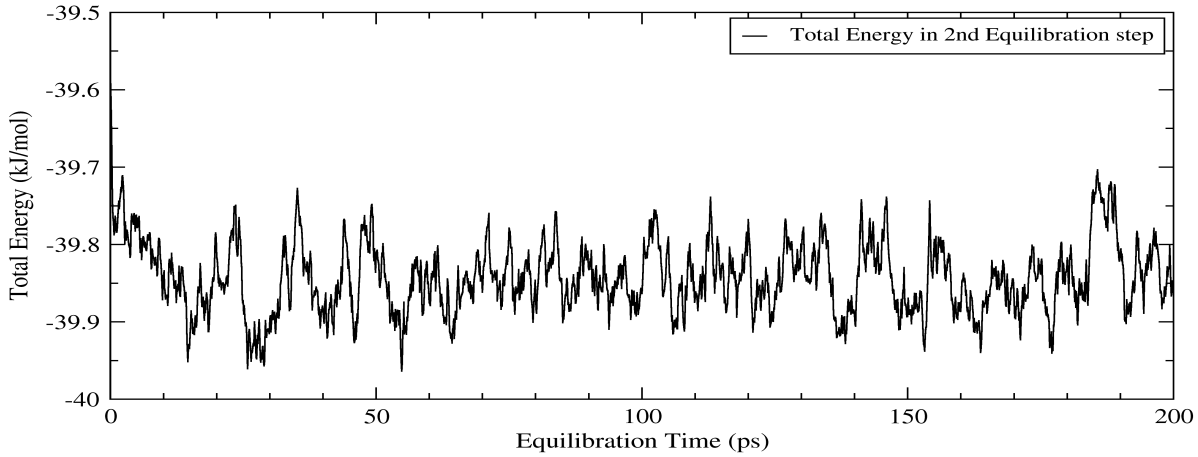


Figure-10: Illustration of equilibration of total energy of system of hh-CytC in 20% aqueous MeOH solvent in constant NPT in second equilibration step at 298.15 K and 1 atm pressure after heating the system.

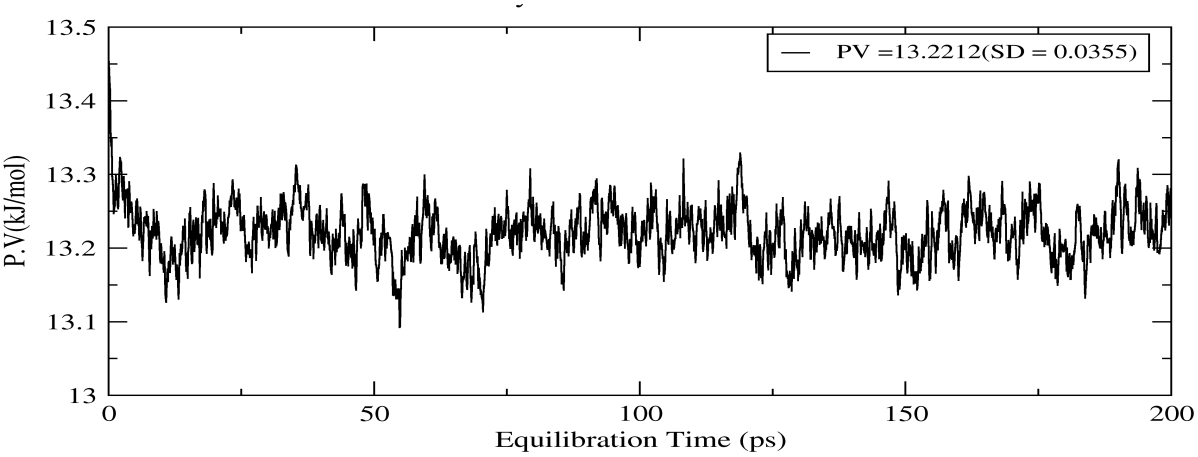


Figure-11: Illustration of the constant PV energy profile in a simulation of system of hh-CytC in 20% aqueous MeOH solvent in constant NPT in second equilibration step after heating the system.

Figure-12 illustrates RMSD of hh-CytC in three different simulation ensembles in second equilibration step indicating that protein structure is sufficiently equilibrated in solution and ready for the next data production simulation step.

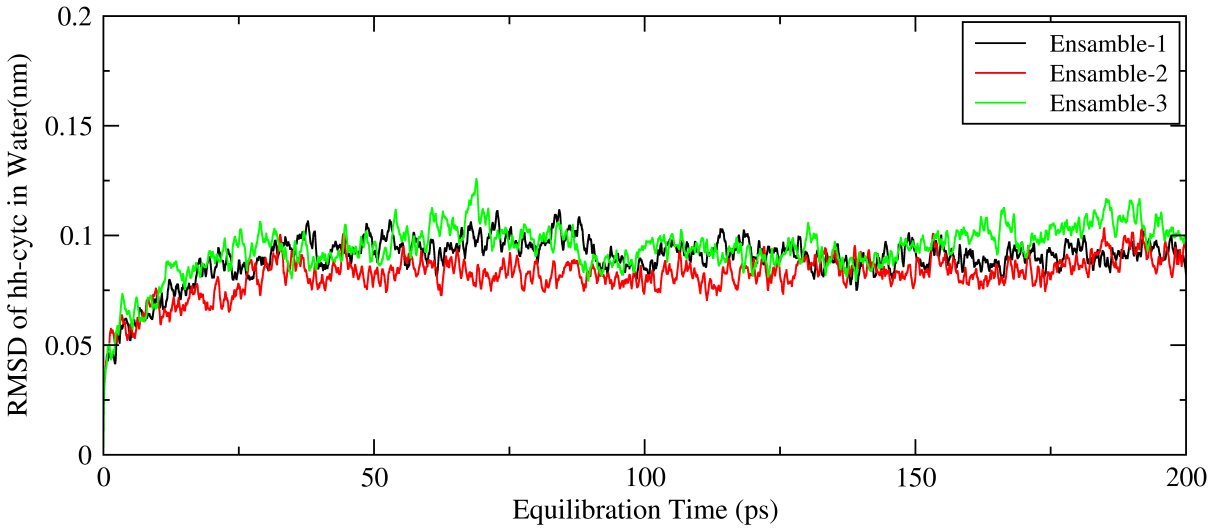


Figure-12: RMSD of hh-CytC in pure water solvents in three different ensembles in NPT second equilibration step at 298.15 K and 1 atm pressure indicating the plateau of equilibrium conformation.

7. Data Production Step: Molecular Dynamic Simulation of System.

After equilibration, a 100 nanoseconds NPT simulation was performed at 298.15 K and 1 atm pressure. Data were collected at every 1 picosecond intervals. Altogether twelve systems, four systems of solvent only, eight systems of hh-CytC in solvents including with and without crystallographic water were included in this study. For each system, three parallel simulations were performed as three ensembles of system where each ensemble was equilibrated by varying initial velocities to take ensemble average of simulation parameters to be studied [13]. Figure-13 illustrates the variation of C-α-backbone RMSD of same system, typically hh-CytC in water, in different ensembles evidencing the need for the use of replicas to extract meaningful conclusions from the results. We averaged the RMSD of the three ensembles to represent the more realistic condition of experiment.

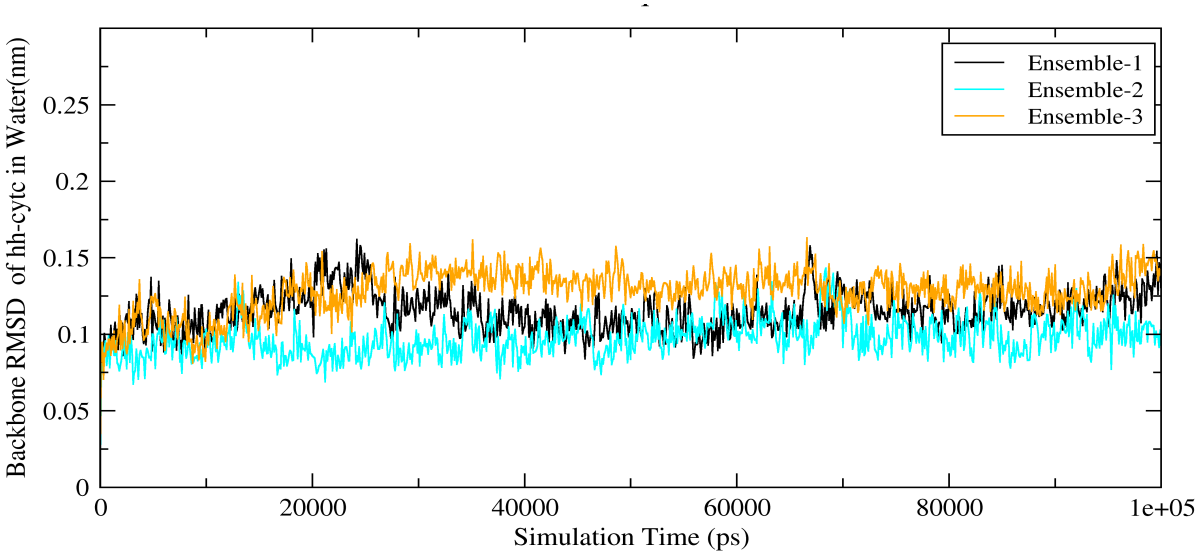


Figure-13: Variation of RMSD of backbone of hh-CytC in water solvent in different ensembles at 298.15 K and 1 atm pressure, indicating that our system of simulation is reproducible.

8. Methods of Analyses.

All the analysis is based on GROMACS software. The molecular graphic images were generated using the Visual Molecular Dynamics (VMD) software [17]. Data calculation and graphical presentation were performed using EXCEL and XMGRACE softwares [17, 18].

8.1 Solvent Macroscopic Properties: Some of the solvent macroscopic properties which represent the solvents' characteristics, such as dielectric constant (ε), diffusion coefficient, shear viscosity (η), molecular dipole moment, density (ρ), and hydrogen bond lifetime are evaluated using GROMACS calculation methods for comparison with experimental methods [18, 27, 28, 34, 36].

8.2 Solvent Viscosity from Transverse Auto Correlation Function: Based on the GROMACS computation methods, viscosity of solvents is determined from transverse-current correlation functions for plane waves using the Navier-Stokes equation,

$$u_x(z,t) = u_0 e^{-t/\tau_r} Cos(\mathbf{k}z) \qquad [9]$$

$$\tau_r = \frac{\rho}{\eta\, \mathbf{k}^2} \qquad [10]$$

where $u_x(z,t)$ is the transverse auto correlation function for the plane waves with amplitude **k** in box axis z (k-factor), $\boldsymbol{\tau_r}$ is the relaxation time(rotational) of solvent molecules and *t* is the simulation time, η is the viscosity of solvent, and ρ is the density of solvents. We used GROMACS command line for transverse autocorrelation function directly and the viscosity is estimated at **k** equal to zero [38, 39].

8.3 Root Mean Square Deviation: RMSD is a measure of how much a conformation of a molecule deviates in simulation from its initial X-ray crystal conformation. Mathematically it is defined as the root-mean-square deviation between simulated structure and the reference structure.

$$RMSD(t) = \sqrt{\frac{\sum_{i=1}^{Na}\left[r_i(t) - r_{ref(t=t_{ref}=0)}\right]^2}{Na}} \qquad [11]$$

where N_a is the number of atoms in protein, m_i is the mass of atom *i* , $r_i(t)$ is the position of atom *i* at simulation time *t*, and t_{ref} is the time is the time step in the simulation corresponding to the reference structure taken. With hh-CytC being a globular protein, RMSD is one of the better techniques to understand the structural changes that occur in protein as a function of simulation time from its original X-ray crystal structure. Even though there is a difference

between crystal and solution structure of proteins, the X-ray crystal structure of hh-CytC equilibrated at 298.15 K and 1 atm pressure while position restraint was the reference structure to calculate the RMSD in our study, after excluding rigid body translation and rotational changes. The small and stable RMSD (typically < 0.3 nm) for the protein backbone is a useful quality control for protein simulation [9, 13, 18, 28, 34, 40].

8.4. Root Mean Square Fluctuation: RMSF gives a time-averaged deviation of each conformational position of a molecule or residue from its average position over simulation time. It is defined as

$$RMSF(i) = \sqrt{\frac{\sum_1^{N_t}[r_i(t) - \langle r_i \rangle]^2}{N_t}} \qquad [12]$$

where $<r_i>$ is the time-averaged position of atom *i* and N_t is the number of configurations or time frames in the simulation trajectories. RMSF measures the standard deviation of position over time. It helps to judge the mobility of flexible segements of protein molecule [18, 40, 41].

8.5 Radius of Gyration: Radius of gyration, R_g, represents the size or dimension or compactness of the structure of protein. It measures the mass-weighted root mean square average distance of all atoms in a protein from its centre of mass.

$$Rg = \sqrt{\frac{\sum_{i=1}^{N_a} mi[r_i(t) - r_{com}(t)]^2}{M}} \qquad [13]$$

$$with \quad r_{com} = \frac{1}{M}\sum_{i=1}^{N_a} m_i\, r_i \quad and \quad M = \sum_{i=1}^{N_a} m_i \qquad [14]$$

where M is the molecular weight of protein with radius r_{com} as *centre of mass radius*, and r_i is the Cartesian position of atom i with mass m_i and N_a the number of atoms considered in protein molecules [18, 40, 41].

8.6 Dynamics and Structure of Hydrogen Bonding: The hydrogen bond is a special type of strong dipole-dipole electrostatic interaction that occurs between highly electronegative atoms like oxygen and nitrogen and hydrogen bonded to one of these atoms. It plays a very important role in protein structure and function as well as in solvent dynamics. We will measure characteristics of hydrogen bonding such as H-bond-distance, H-bond angle (H-D-A), H-bond lifetime and number of hydrogen bonds [18, 28, 34, 44]. The criteria of H-bond are 0.35 nm between Donor-Acceptor distance, H-bond distance, and H-bond angle is 30 degree.

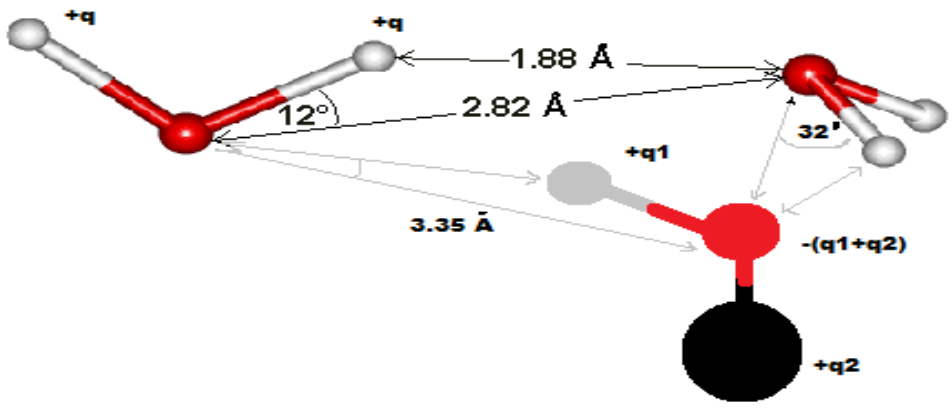


Figure-14: Illustration of hydrogen bond network between water and methanol molecules with H-bond Donor-Acceptor distances, $d_{D\text{-}A}$, and H-bond angles, $\theta_{H\text{-}D\text{-}A}$, mentioned in the criteria of hydrogen bonding in GROMACS.

8.7 Salt-bridge Distance in Protein: The salt-bridges in protein are special names of two types of interactions, hydrogen bonding and electrostatic interactions, between oppositely charged residues that are sufficiently close to each other. They are regarded to contribute in

achieving thermodynamically more favorable conformation of protein. Usually, the salt bridges arise between the anionic carboxylate ($RCOO^-$) of either aspartic acid or glutamic acid or carboxyl-terminal and the cationic ammonium (RNH_3^+) from lysine or the guanidinium ($RNHC(NH_2)_2^+$) of arginine or amino-terminal. Depending on solvent properties, other amino acid residues with ionizable side chains, such as histidine, tyrosine, and serine can also be involved in salt bridges. The salt bridges may exist between two opposite charges or as a complex network of three or more charges. The distance (≤ 4.0 Å) and geometry (H-bonding criteria) between amino acid residues involved in salt bridges are very crucial to exhibit their favorable as well as unfavorable contribution in particular protein conformation [34, 42, 43].

8.8 Hydrogen Bond Auto Correlation Function (ACF) and H-bond Lifetime (τ_{HB}): The H-bond ACF gives the hydrogen bond kinetics among the possible pairs of H-bonding partners. The probability binary function, the H-bond operator (H), for H-bond between a donor-acceptor pair, *i* , at time t = 0 is $H_i(0)$ and at any simulation time t is $H_i(t)$, with condition that $H_i(t) = 1$ if the H-bond exists (or criterion of H-bond is valid) and $H_i(t) = 0$ if the H-bond ruptures or does not exist. Then the auto correlation function for H-bond between all possible pairs is given by

$$C_i^{HB}(t) = \frac{\langle H_i(0)\, H_i(t)\rangle}{\langle H\rangle} \qquad [15]$$

The $C_i^{HB}(t)$ gives the conditional probability of existence for the H-bond between donor–acceptor pairs in first coordination shell at time *t*, $H_i(t)$, provided the H-bond was intact at $t = 0$, $H_i(0) = 1$.

GROMACS estimates H-bond lifetime (τ_{HB}) from ACF based on Luzar and Chandler's [44, 45, 46] proposal where they considered the formation and breaking of H-bond exist as equilibrium as

$$D—H \quad + \quad A \rightleftharpoons D—H\text{----}A$$

Since H-bonds break and re-form continuously over simulation time *t*, and at equilibrium, the measured ACF, $C_i^{HB}(t)$ represent "Intermittent H-bond Correlation Function" which is independent of previous condition of H-bond or possible H-bond breaking events.

$$\frac{C_i^{HB}(t) - C_i^{HB}(\infty)}{C_i^{HB}(0) - C_i^{HB}(\infty)} = C_i^{IHB}(t) \qquad [16]$$

and the rate of decay of IACF, $C_i^{IHB}(t)$ is expressed as as first-order chemical kinetic equation as

$$\frac{-dC_i^{IHB}(t)}{dt} = k(t) = kC_i^{IHB}(t) - k'n^{IHB}(t) \qquad [17]$$

where $\boldsymbol{k}$ as H-bond formation rate constant and $\boldsymbol{k'}$ is the H-bond breaking rate constant. $n^{HB}(t)$ is the probability function for the number of accessible D & A pairs within H-bonding distance cut-off (3-4 Å) in the second coordination shell.

The average H-bond lifetime is given by the inverse of forward rate constant of H-bond formation only [34, 44, 45, 46].

$$\tau_{HB} = {}^{1}/_{k} \qquad [18]$$

The geometric criteria or conditions of H-bond being formed between all possible donor and acceptors in our calculation are: $r_{(D\text{-}A)} \leq 3.5$ Å and Angle H-D-A $\leq 30^0$.

8.9 Radial Distribution Function: The RDF is a pair correlation function and measures the probability of finding a particle at a distance r with respect to other particle taken as reference. It is calculated by determining how many particles are within distance of r and r+Δr around away from reference particle.

$$N(dr) = g(r).4\pi r^2.\rho.dr \qquad [19]$$

where ρ is the number of particles per unit volume (N/V). We measure $g(r)$ as a function of r. RDF provides information about the density of particles at radius r and hence characterizes the structure of the system [27, 28].

8.10 Mean Square Displacement and Diffusion Coefficient: The translational diffusion coefficient of hh-CytC or solvent molecules was calculated as an average from three independent molecular dynamic simulations of 100ns. From each ensembles, the translational diffusion coefficient was computed using gromacs command which use the Einstein formula

$$D_{trans} = \frac{\langle |r(t)-r(0)|^2 \rangle}{6t} = MSD/6t \qquad [20]$$

where numerator term is Mean Square Displacement, MSD, attained by molecule's centre of mass in a time interval of t [27, 28, 34, 47].

8.11 Lindemann Parameter: The Lindeman's disorder index is a the useful parameter in determining the malleability and stability of protein. Since the internal motions of proteins play an essential role in their biophysical activities, the Lindemann's parameter utilizes the characteristics of internal motions with its atomic distributions inside protein in determining the flexibility and stability of protein. The formula to estimate Lindemann's disorder index ($\boldsymbol{\Delta_L}$) is

$$\Delta_L = \frac{\sqrt{\frac{\sum_{i=1}^{Na} \langle \Delta r_i^2 \rangle}{Na}}}{a'} \qquad [21]$$

where N_a is the number of atoms considered to calculate mean square fluctuation, $\langle \Delta r_i^2 \rangle$ is the mean square fluctuation of atom i over all trajectories, a′ is the most probable non-bonded near-neighbour distance and it is estimated as the distance corresponding to peak position in atomic radial distribution function of hh-CytC as shown in Appendix-VII [48]. The critical value of $\boldsymbol{\Delta_L}$ is 0.15 which is relatively independent of the types of substance or protein, the nature of the interaction potential, and the crystal structure. If the value of $\boldsymbol{\Delta_L}$ is less than 0.15, the protein has a solid-like nature in rigidity (low flexibility), whereas for the values higher than 0.15, proteins exhibit high flexibility behaving, like a liquid [11, 48, 49].

8.12 van Hove Distribution Function: VHDF is also called a dynamical radial distribution function. The van Hove distribution function, *G(r,t)*, is a real space dynamical correlation function for characterising the spatial and temporal distributions of pairs of particles in a fluid. It gives the probability of finding particles at distance ***r*** at time *t*, where |r| = ***r***, given that one of the particles was located at the origin at time *t = 0*. In other word, it measures the distribution of distance moved of one particle relative to other at time *t*. We used VHDF to find the time dependent diffusion or movement of crystallographic molecules around the hh-CytC.

$$G(r,t) = \frac{1}{N} \langle \sum_{i=1}^{N} \sum_{j=1}^{1} \delta(r_i(0) + r_j(0) - r_i(t)) \rangle \qquad [22]$$

Here, i = 124, the number of crystallographic water molecules and j = 1, the single protein molecule in our system. Since, protein move very slowly relative to water molecules, $r_j(0) = 0$, and the distribution gives the average relative movement of cryst.H_2O molecules from protein

surface at time t [18, 63]. In other words, the van Hove distribution function in our system gives radial distribution of crystallographic water molecules from protein surface as a function of time.

CHAPTER III

RESULTS AND DISCUSSION

From simulation data, the structure and dynamics of hh-CytC was analyzed and compared in different mixtures of water and methanol. Dynamics of crystallographic water were studied and the interaction of solvent and hh-CytC is analyzed. Similarly, some properties of solvents are also studied in these systems.

1. Analysis of Solvent Properties.

Before simulation of hh-CytC in solvents, solvent only boxes of similar sizes were simulated for 20 ns at the same NPT conditions, 298.15 K and 1 atm. The important solvent liquid properties were computed and compared with available experimental data as shown in Table-3. The computed data for pure solvent properties are quite comparable with experimental properties. In mixture, both solvents lose their identity even though they have similar computed molecular charges in the mixtures and pure solvents with identical force fields. The effect of mixing of two solvents was seen in dielectric constant where ε_{MeOH} has decreased more compared to ε_{H2O} in mix-solvents even though computed dielectric constant was lower than the experimental value. We opined this behavior might be due to shielding of MeOH dipole by H-bond network of H_2O primarily, and microscopic augmentation of hydrophobicity resulting from methyl group [23]. The diffusion constant of the binary mixture was found to be lower in 40% aqueous MeOH solvent, which may be due to higher mixing effect in each component. The shear viscosities were calculated from transverse-current correlation function for plane waves in NVT simulation [38].

Table-3: Computed macroscopic properties of solvents of different compositions of water-methanol binary system from NPT – MD simulation at 298.15 K temp. and 1 atm pressure.

Solvent Parameters		Solvents, Water-Methanol Mixture			
		Water	**20%MeOH**	**40%MeOH**	**MeOH**
Dielectric Constant (ε)	**Expt.**[35]	78.86	64.9	54.1	33.30
	Calctd.	73.04±0.11	62.48±0.11$^{\gamma}$ 52.80±0.76$^{\alpha}$ 3.33±0.23$^{\beta}$	53.21±0.11$^{\gamma}$ 35.43±0.43$^{\alpha}$ 6.32±0.18$^{\beta}$	25.70±0.58
Diffusion Constant (× 10^{-9} m^2s^{-1})	**Expt.**	2.60±0.20[23]	1.12[51]	1.05[51]	2.42±0.05 [52]
	Calctd.	2.67±0.49	1.95±0.08$^{\gamma}$ 2.07±0.03$^{\alpha}$ 1.74±0.19$^{\beta}$	1.75±0.14$^{\gamma}$ 1.76±0.03$^{\alpha}$ 1.58±0.17$^{\beta}$	2.30±0.38
Shear Viscosity (mPa.s) [38, 53]	**Expt.** [35,53]	0.893	1.604$^{20\cdot C}$	1.837$^{20\cdot C}$	0.586$^{20\cdot C}$
	Calctd.^{NVT}	1.05 ± 0.22	1.50±0.53	1.41±0.37	0.61±0.08
Dipole Moment (Debye)	**Expt.**[20]	2.95^{a} 2.10^{b}	----------	----------	2.54^{a}[21] 1.69^{b}
	Calctd.	2.3505	2.3449$^{\gamma}$ 2.3505$^{\alpha}$ 2.3055$^{\beta}$	2.3380$^{\gamma}$ 2.3505$^{\alpha}$ 2.3046$^{\beta}$	2.30±07
Density (g.cm^{-3})	**Expt.**$^{20\cdot C}$ [35]	0.9992	0.9666	0.9347	0.7917
	Calctd.	0.9991	0.9662	0.9296	0.7938
H-bond Lifetime (Pico-second) [18, 40, 45]	**Expt.**	--------	-------	--------	-------
	Calctd.	2.205	2.998$^{\theta}$ 2.251$^{\zeta}$ 2.543$^{\Omega}$	3.883$^{\theta}$ 3.217$^{\zeta}$ 3.298$^{\Omega}$	5.79

^{a}Liquid state, ^{b}Single molecule in gaseous state | $^{\alpha}H_2O$, $^{\beta}$MeOH , $^{\gamma}$Binary Mixture | $^{\theta}H_2O$-H_2O, $^{\zeta}$MeOH-MeOH, $^{\Omega}H_2O$-MeOH

The calculated viscosities of pure solvents were higher than the experimental values; but for mix-solvents, calculated values of viscosities were lower than the experimental values. Since the X-H bond was constrained by LINCS algorithm, both MeOH and H_2O molecules

have same dipole moment in mixture and pure liquid in spite of Me-O unconstrained bond. Moreover, density of each solvents were the best reproducible as experimental values.

Since both H_2O and MeOH are hydrogen bonding liquids, the characteristics of H-bond were computed in different solvent compositions. The intermittent H-bond lifetime between donor-acceptor pairs was calculated from GROMACS over 20 ns simulation [18, 44, 45]. In pure liquids, the H-bond lifetime between water molecules, $\tau_{H2O\text{-}H2O}$, was found shorter than the H-bond lifetime between MeOH molecules, $\tau_{MeOH\text{-}}$MeOH , which may be simply the bulky methyl group lags MeOH molecules to find new H-bonding partners vis a vis in liquid water. The $\tau_{H2O\text{-}H2O}$ has increased with increase of MeOH percent in the mixture but $\tau_{MeOH\text{-}MeOH}$ has decreased with decrease of MeOH percent in the mixture. The cross H-bond lifetime, $\tau_{MeOH\text{-}H2O}$, has increased with increase of MeOH percent. To explain this behavior, we can exploit hypothesis of 'microscopic segregation' of water and methanol in their mixture [8]. The hydrophobic methyl group try to become far from polar end resulting micro-micelle interlocked in a network of H_2O-H_2O H-bond so that existence of H-bond between MeOH is meager due to orientation constraint, and H-bond between H_2O and MeOH becomes highly probable.

2. Analysis of Structure of Horse Heart Cytochrome C in Different Solvents.

The protein C-α backbone RMSD of hh-CytC, measured with respect to its X-ray structure after first equilibration, was calculated and compared in different solvents of water and methanol. As mentioned earlier, the RMSD value provides the information related to how much the protein structure deviates from the X-ray crystalline structure in the different environments. Table-4 gives the average values of RMSD over three ensembles of 100 ns

simulation and figure -15 shows the dynamical change of RMSD as a function of simulation time of one simulation in four different solvents at 298.15 K and 1 atm. All simulations show reasonable values of RMSD, typically less than 0.30 nm, from the X-ray structure indicating that the hh-CytC was quite stable in our simulation time of 100 ns and the simulation conditions were reproducing the correct physics of the system.

Table-4: Average C-α backbone Root Mean Square Displacement of hh-CytC in different solvent at 298.15 K and 1atm, with and without including crystallographic water.

Solvent	**RMSD (nm) ± SD of backbone of hh-CytC**		**Relative Difference (%)**
	With cryst.H_2O	**Without cryst.H_2O**	
Water	0.116±0.006	0.158±0.033	36.21 ↑
20%MeOH in Water	0.131±0.010	0.159±0.047	21.37 ↑
40%MeOH in Water	0.137±0.008	0.140±0.015	2.19 ↑
MeOH	0.166±0.031	0.127±0.014	23.49 ↓

As the protein can fold via multiple parallel path ways [54, 62], as shown in Figure-13 and Appendix-IV, different simulation ensembles in same solvent system have shown slightly different variation of RMSD with overlapped intermediate conformations in our simulation time even though we started with same initial conformation. This information also implies that protein folding mechanism does not follow a single conformational path. The higher value of RMSD may be due to differences in structure of hh-CytC between crystal and in the solution, even though the NMR solution structure of reduced horse heart Cytochrome C has showed the backbone RMSD of 0.67 ± 0.10 Å [6, 33]. Compared to the RMSD of hh-CytC in water, the presence of MeOH in mix-solvents led to an increase in the RMSD of hh-CytC, and it reached the highest value in pure MeOH solvent indicating the dependence of RMSD values on the

number of water molecules in hydration to maintain protein structure in simulations more akin to its crystal structure. But RMSD attains a somewhat consistent fluctuation, due to interplay of motional constraints from the hydrophobicity of MeOH, intramolecular protein cross-linking effect of methanol, and stabilizing nature of MeOH to α-helices which results in conformational entropy loss of in the protein [34, 56, 62].

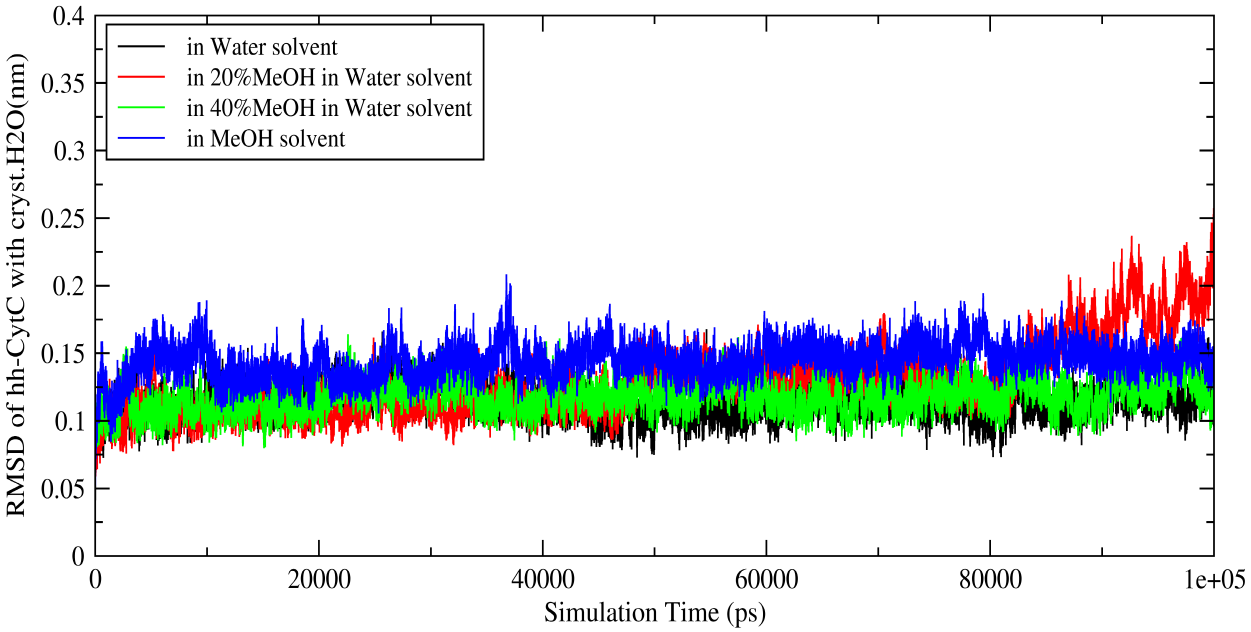


Figure-15: Time evolution of C-α backbone RMSD of Protein keeping crystallographic water with hh-CytC in different solvents at 298.15 K and 1 atm at constant NPT condition.

After 50 ns, the RMSD seems stable except in 20% aqueous MeOH solvent. In 20% aqueous MeOH, RMSD of protein shows higher value with higher fluctuation (±SD) over simulation time compared to 40% aqueous MeOH solvent. Typically, after 80 ns, the R_g (Figure-23) and RMSD were increasing and reached values higher than those in MeOH solvent. So, in spite of having C-α backbone RMSD below 3 Å and being limited within our simulation time of 100 ns, based on fluctuations we might think that hh-CytC might tend to unfold in mix-

solvents earlier than in pure solvents, as observed protein folding behavior in ethanol-water mixture [62]. But the final structures of hh-CytC in all four solvents as shown in Figure-25 did not show any visible unfolded secondary and tertiary structures in our simulation except minor positional changes.

We also performed the MD simulation experiments of hh-CytC, removing all crystallographic water molecules mentioned in X-ray crystal structure of hh-CytC [3]. Figure-16 to 19 displays the C-α backbone RMSD of hh-CytC with and without crystallographic water molecules in different solvents separately and figure-20 gives the RMSD of hh-CytC without cryst.H_2O at 298.15 K and 1atm in different solvents. Even though, there is not any drastic change in data, as shown in Table-4, implying any protein unfolding or any tertiary structural changes. Also, the C-α backbone RMSD of hh-CytC without crystallographic water molecules has increased surprisingly more in solvents with high water content. In water, even though the protein has similar conformations with and without cryst.H_2O in first 5 ns, hh-CytC reaches higher meta-stable states along a 100 ns simulation which implies that protein conformational change may follow different mechanisms within our simulation time with late success to regain the solvent water molecules in the empty sites of cryst.H_2O inside protein so that structure becomes comparatively more labile with high energy meta-stable state. On the other hand, compared to 20% aqueous MeOH solvent, the change in RMSD in 40% aqueous MeOH solvent is lower and RMSD decreases in MeOH solvent. Here we may explain this behavior as an interplay of hydrophobic effect of tiny methyl group where MeOH has α-helix stabilizing property [8, 56, 62], and lubrication properties of water, where lack of cryst.H_2O hh-CytC loses internal flexibility and gains surface rigidity. So, these results clearly reveal the importance of

these crystallographic water molecules in buffering structural flexibility and rigidity of hh-CytC in different environments.

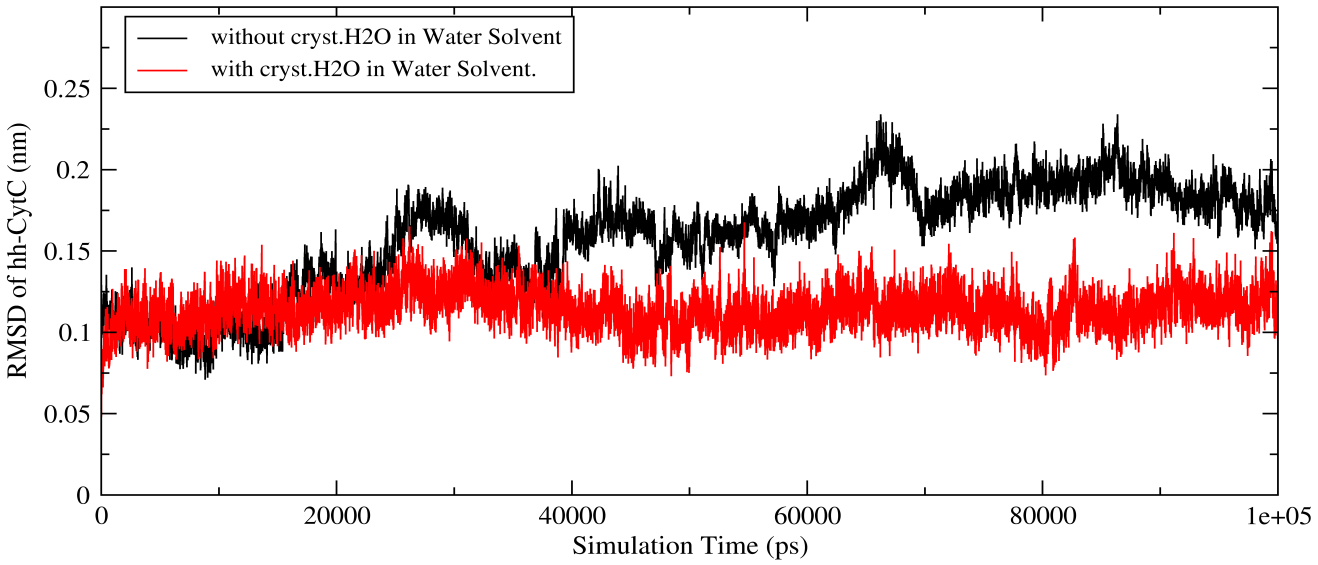


Figure-16: RMSD of C-α backbone of hh-CytC with and without crystallographic water molecules in pure water solvent at 298.15 K and 1 atm.

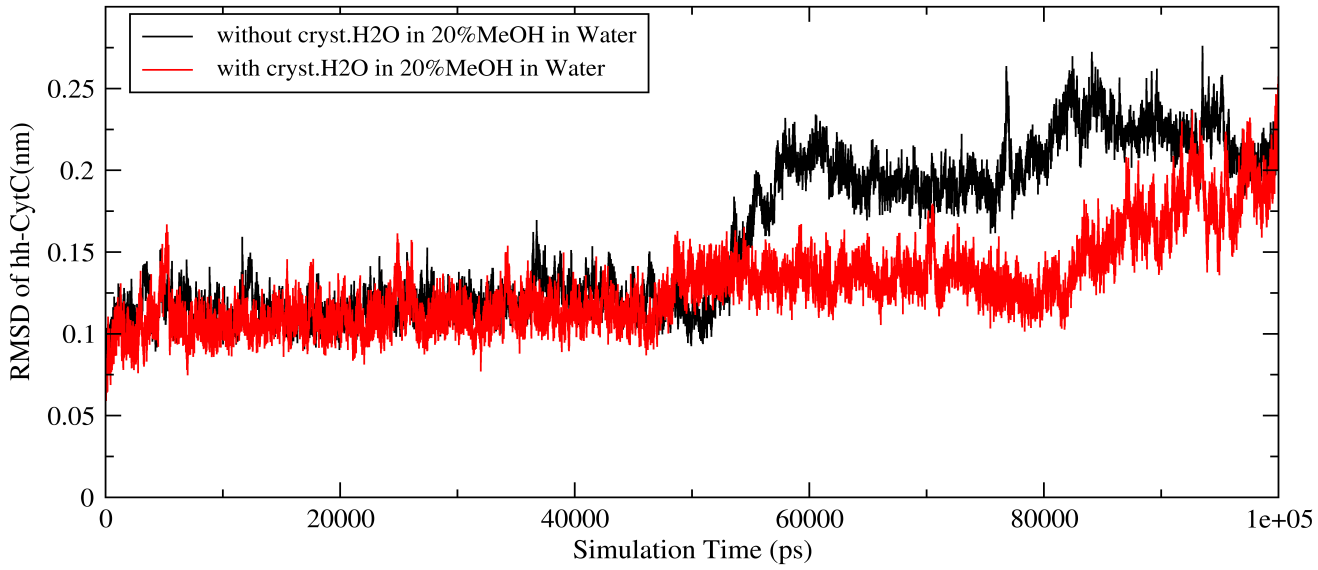


Figure-17: RMSD of C-α backbone of hh-CytC with and without crystallographic water molecules in 20% aqueous MeOH solvent at 298.15 K and 1 atm.

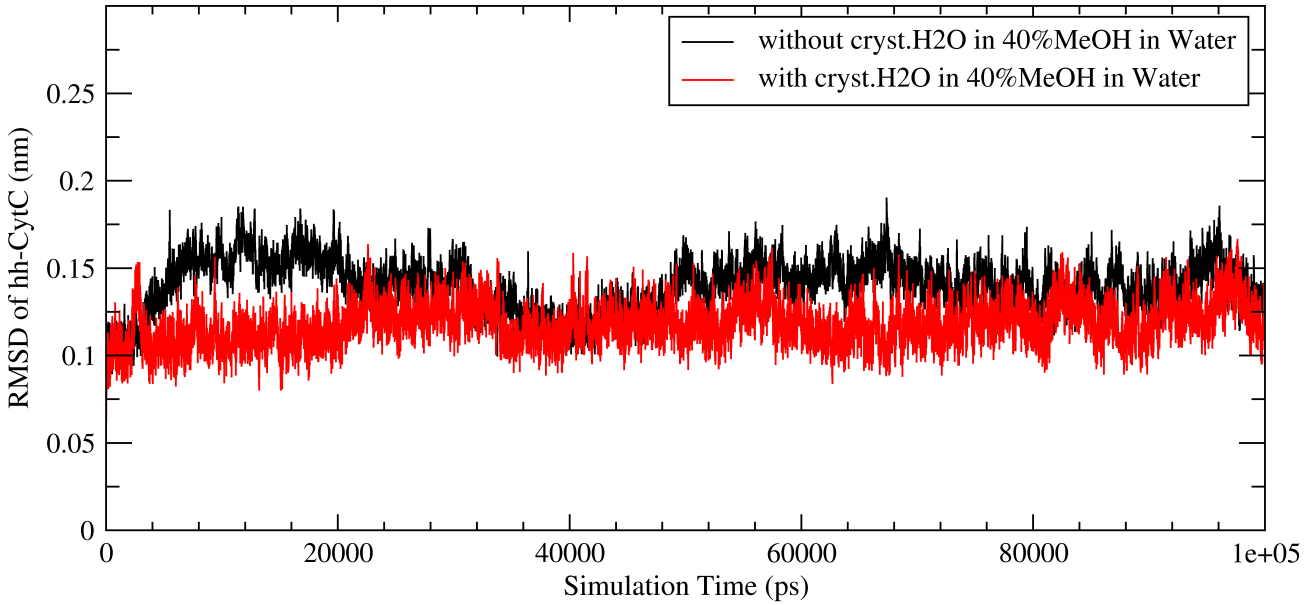


Figure-18: RMSD of C-α backbone of hh-CytC with and without crystallographic water molecules in 40% aqueous MeOH solvent at 298.15 K and 1 atm.

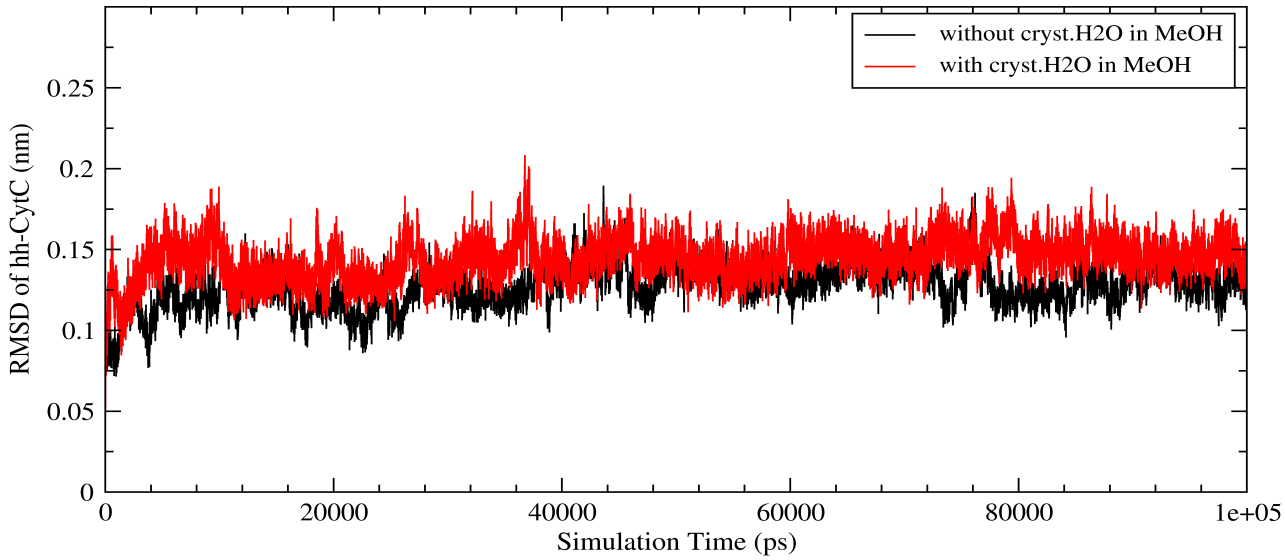


Figure-19: RMSD of C-α backbone of hh-CytC with and without crystallographic water molecules in MeOH solvent at 298.15 K and 1 atm.

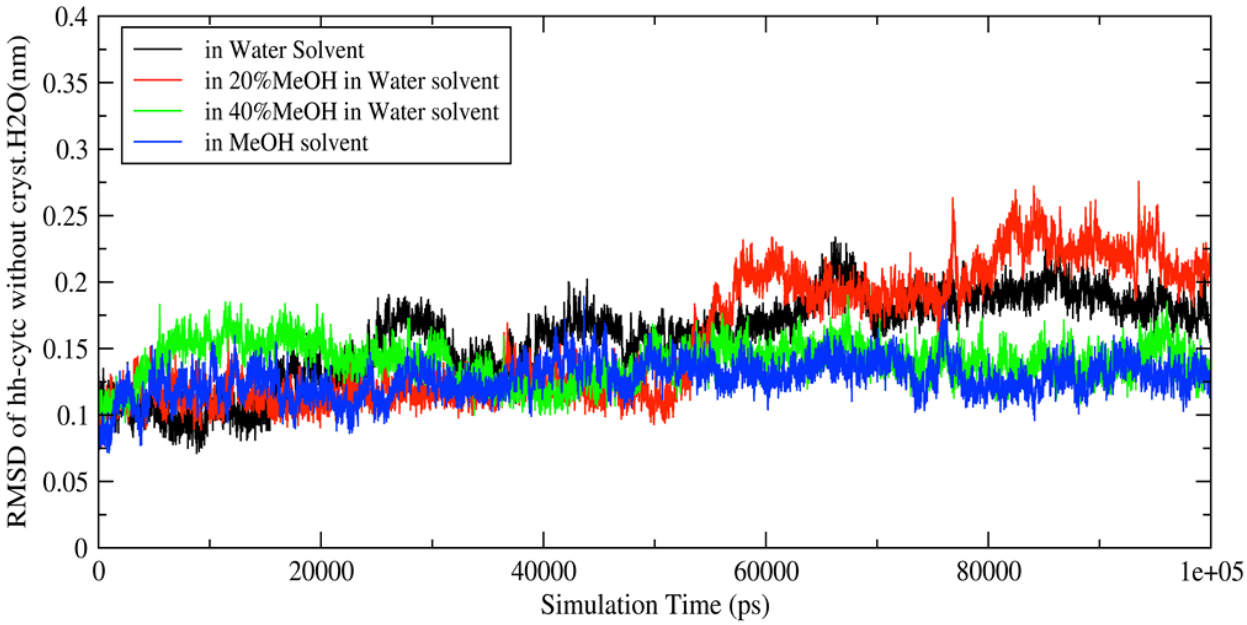


Figure-20: RMSD of C-α backbone of hh-CytC without crystallographic water in different solvents at 298.15 K and 1 atm.

The buried HEM group showed very low RMSD as shown in RMSD distribution in figure-21, but it is still influenced by outer solvent environment, with 0.01 – 0.1 ranges of values in all systems with and without crystallographic water. The RMSD distributions demonstrate that HEM spends more diverse conformations without cryst.H_2O than with cryst.H_2O. Even though the heme group is bound to protein at two sites by covalent bonds and at **Fe** by two axial coordinate bonds, we did not observe any correlation between RMSD of HEM and C-α backbone RMSD of protein in hh-CytC.

Similarly, as shown in table-5 and figure-22 and 23, we estimated the radius of gyration, R_g, and Lindemann's disorder index, Δ_L of hh-CytC in different solvents at 298.15 K and 1 atm to analyze the overall variations of structural flexibility or rigidity. These parameters provide

further insight supporting RMSD analysis in the study of the degree of conservation of a protein structure.

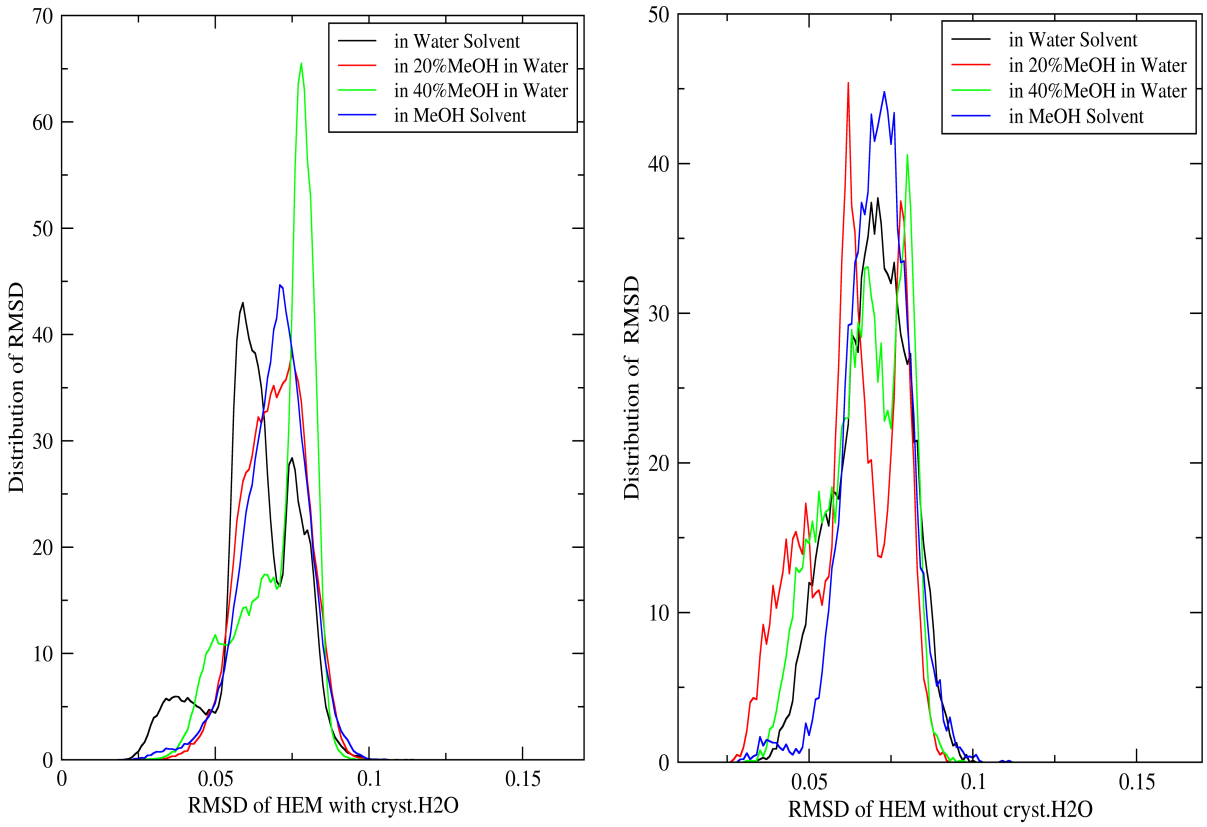


Figure-21: RMSD Distribution of heme group of hh-CytC with (left) and without (right) crystallographic water molecules in different solvents at 298.15 K and 1atm.

Like the RMSD, the R_g increases very slightly with increasing MeOH concentration. This may be due to adopting a slightly more open conformation that exposes the hydrophobic side-chains outside to the solvents, which are stabilized mainly by hydrophobic interactions from methanol. We did not observe any MeOH molecules bound to any specific protein sites like crystallographic water in our long simulation.

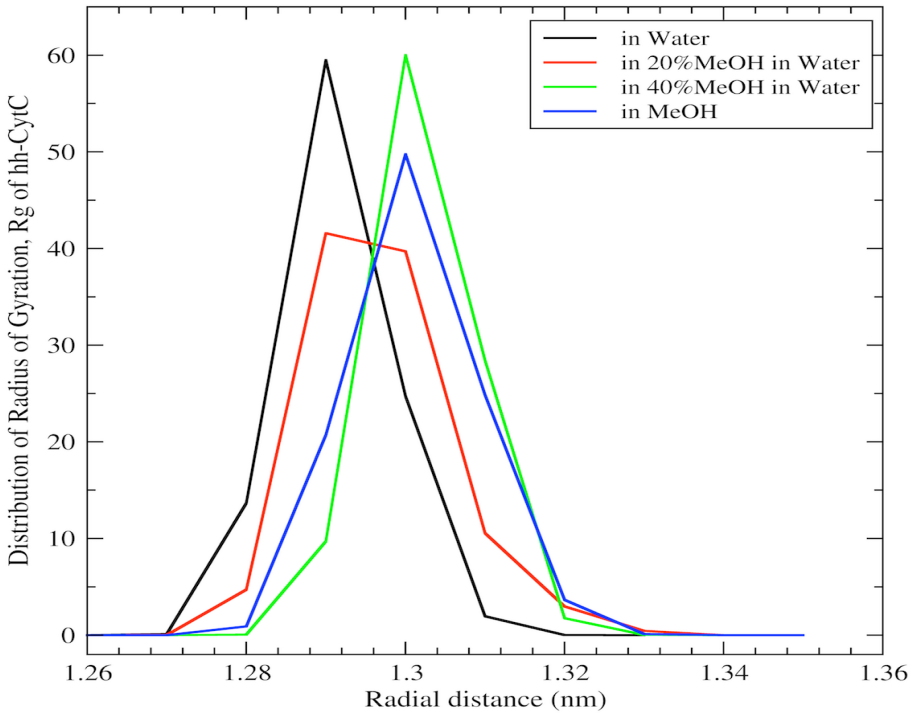


Figure-22: Distance Distribution of Radius of Gyration (R_g) of hh-CytC in different solvents in 100 ns simulation at 298.15 K and 1 atm pressure.

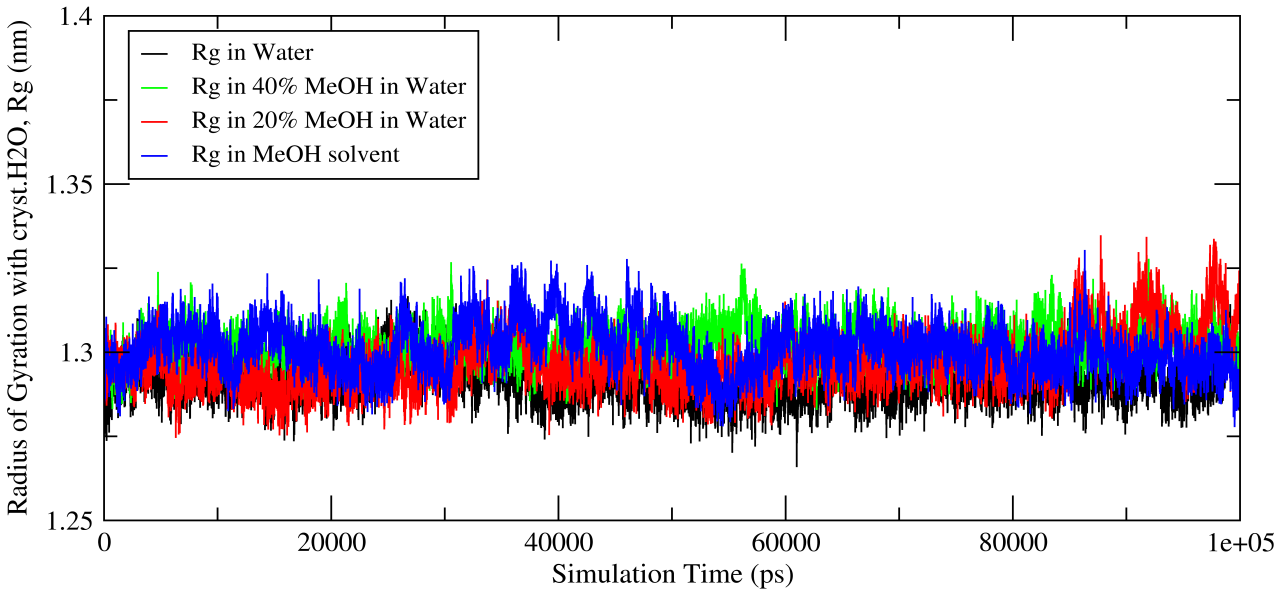


Figure-23: Time trajectories of Radius of Gyration (R_g) of hh-CytC in different solvents at 298.15 K and 1 atm pressure.

Accompanying the structural fluctuations, the dipole moment of hh-CytC has also fluctuated as shown in Table-5 in different solvents. In 20% aqueous MeOH solvent, the dipole moment of hh-CytC is low with high fluctuation. But in pure methanol, the dipole moment of hh-CytC increases.

For the determination of solid-like or liquid-like behavior of hh-CytC, we calculated the Lindemann's disorder index, Δ_L. A lower value of Δ_L means more solid like behavior of protein [11, 48]. For the typical borderline criteria for solid and liquid transitions, if Δ_L is less than 0.15, the structure is considered to be solid-like.

Table-5: Average Molecular properties (Radius of Gyration, Lindemann Parameters and Dipole Moment) of hh-CytC in different solvent at 298.15 K and 1atm.

Solvent	**Rg(nm) ± SD**	**Lindemann's disorder Index (Δ_L)**				**Dipole Moment (Debye)**
		With Cryst. H_2O		**Without Cryst.H_2O**		
		All Atoms	**Backbone atoms**	**All Atoms**	**Backbone Atoms**	
Crystal	1.264[21]	---------	---------	----------	----------	255[21]
Water Solvent	1.290±0.001	0.2728	0.1539	0.3366	0.2176	237.18±12.75
20%MeOH in Water	1.295±0.004	0.3235	0.2026	0.3569	0.2377	240.21±14.08
40%MeOH in Water	1.299±0.008	0.2697	0.1458	0.2923	0.1636	225.67±27.70
MeOH Solvent	1.305±0.005	0.3160	0.1912	0.2733	0.1611	241.21±35.29

The present results reveal that the interior of the protein is more solid-like, while its surface is more liquid like in all solvent system. However, the surface molten solid of proteins is likely to be essential for protein stability and function [11, 48]; high surface fluidity may not be the requirement for protein's proper functioning without preserving protein's stability. In

water, protein backbone that is around heme has borderline Δ_L value (Δ_L = 0.1539) indicating the need of constant backbone flexibility for hh-CytC functioning. The hh-CytC is highly liquid-like in 20% aqueous MeOH solvent (Δ_L = 0.2026) and in methanol (Δ_L = 0.1912); while in 40% aqueous MeOH solvent (Δ_L = 0.1458), the interior is unexpectedly solid-like even though protein surface is enough liquid-like. These behaviors may have ensued from the non-ideality of water-methanol mixture.

In a nutshell, the Lindemann parameter might serve as a good measure of the degree of internal motion inside proteins, which should have correlation with structural entropy. The cores of protein and heme prosthetic group are comparatively rigid compared to the surface of protein in different solvents irrespective to their fluctuations. But the change in internal motion of protein depending on properties of external solvent may be entropically more crucial in concluding solvent specific enzymology.

When we calculated the solvent accessible surface area of hh-CytC using a surface probe sphere of radius 0.14 nm and averaged over whole simulation trajectories as shown in Figure-24 and Appendix-VIII, the total surface area did change by 2-3 percentage as the R_g increased slightly with increasing MeOH concentration, but the fluctuations in hydrophobic surface area increases with increase of MeOH and it is high in MeOH solvent, indicating hydrophobic interaction of methyl group with protruding hydrophobic side chains of the protein. But, for the heme group, the solvent accessible surface area is almost same in all solvents indicating the active site structure varies less than the overall protein structure as solvent changes. The Figure-25 gives the pictorial view of SASA of hh-CytC shown in last conformation of 100 ns simulation in different solvents at 298.15 K and 1 atm. These changes in size and surface area may be attributed as a result of high surface flexibility, hydrophobic

effects, excluded volume at the protein surface by the methyl group of MeOH rather than a change in the tertiary structure of the protein.

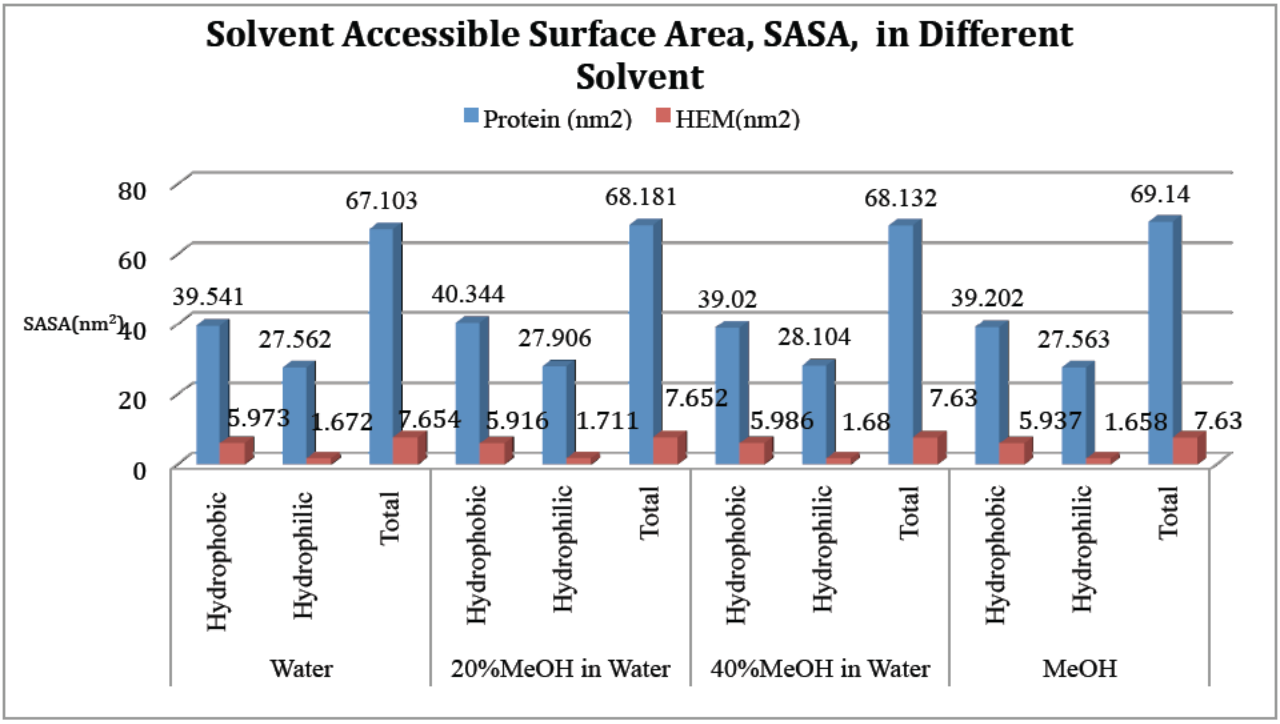
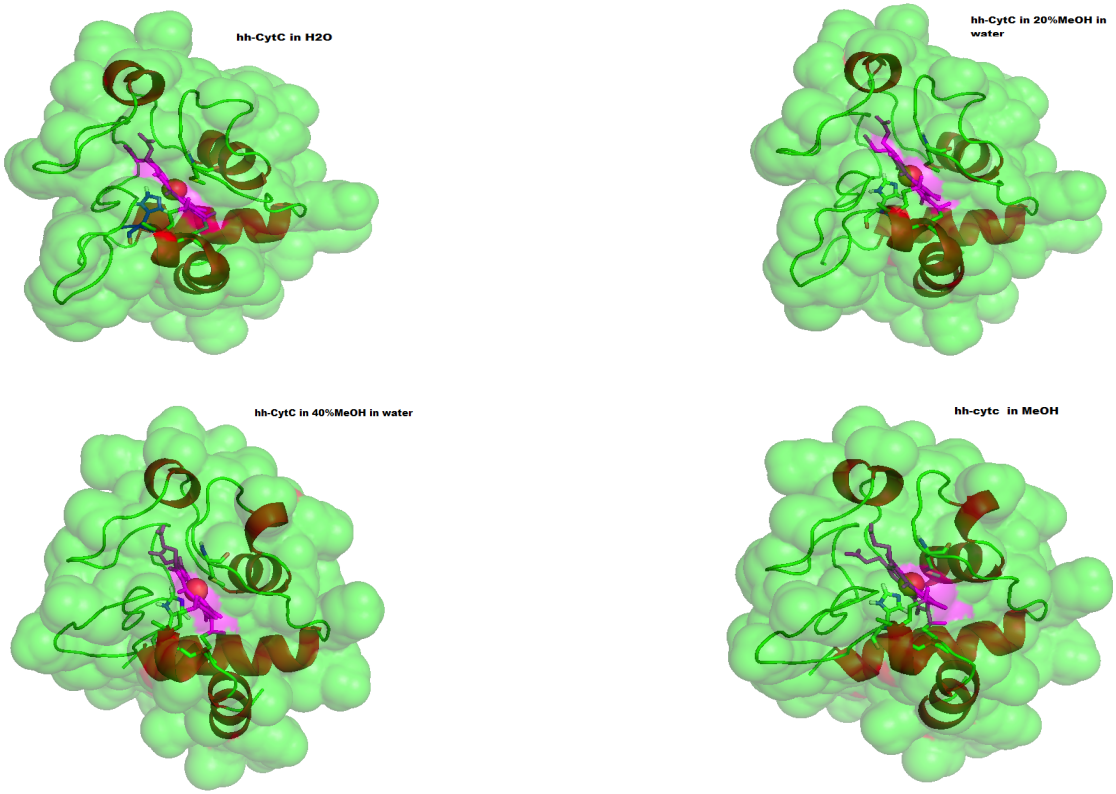


Figure-24: Bar Diagram of calculated values of Solvent Accessible Surface Area, SASA, (in y-axis) of Protein and Heme group using probe 0.14 nm in hh-CytC in different solvents at 298.15 K and 1 atm.



Figures-25: Pictorial view of Solvent Accessible Surface Area, SASA of Protein (green) and Heme group (pink) in the last conformations of hh-CytC after 100 ns simulation in different solvents at 298.15 K and 1 atm pressure at constant NPT.

The Root Mean Square Fluctuation, C-α backbone RMSF of each amino acid residues averaged over whole 100 nanoseconds simulation time as shown in Figure-26 indicates the structural fluctuations are restricted to certain segments with some considerable changes in the flexible part of hh-CytC across the solvents studied. Generally, five α-helices [H1(6-14), H2(49-54), H3(60-68), H4(70-75, H5(87-102)] provide rigidity to the protein with coils and turns in between as flexible segments[4,8]. The flexible segments in hh-CytC which exhibit high mobility are the residues 1GLY-2ASP-3VAL-4GLU of amino-terminal coil, residues 21GLU-22LYS-23GLY-24GLYof the type-II β-turn and 25LYS-26HIS of the coil in Ω_1 loop, the hydrogen-bonded turn comprising residues 41GLY-42GLN-43ALA-44PRO-45GLY-

46PHE, residues 49THR-50ASP-51ALA-53LYS-54ASN of the α-helix and residues 55LYS-56GLY of the coil in Ω_2 loop, the α-helix residues 72LYS-73LYS-74TYR-75ILE, residues 76PRO-77GLY-78THR of the type-II β-turn, the coil of residues 83ALA-84GLY-86LYS and 87LYS-88LYS-89THR of the α-helix in Ω_3 loop, and the carboxyl-terminal coil of residues from 103ASN-104GLU. The α-helix residues 12GLN-13LYS-14CYS, which is a comparatively rigid segment, have shown more mobility in mixed solvents rather than in pure solvents. In some foldons especially the amino-terminal coil, the coil residue 60LYS and the α-helix residues 61GLU-62GLU-63THR and 72LYS-73LYS-74TYR-75ILE, the type-II β-turn of residues 76PRO-77GLY-78THR and the carboxyl-terminal coil of 103ASN-104GLU exhibit more mobility in 20% aqueous MeOH solvent compared to other solvents.

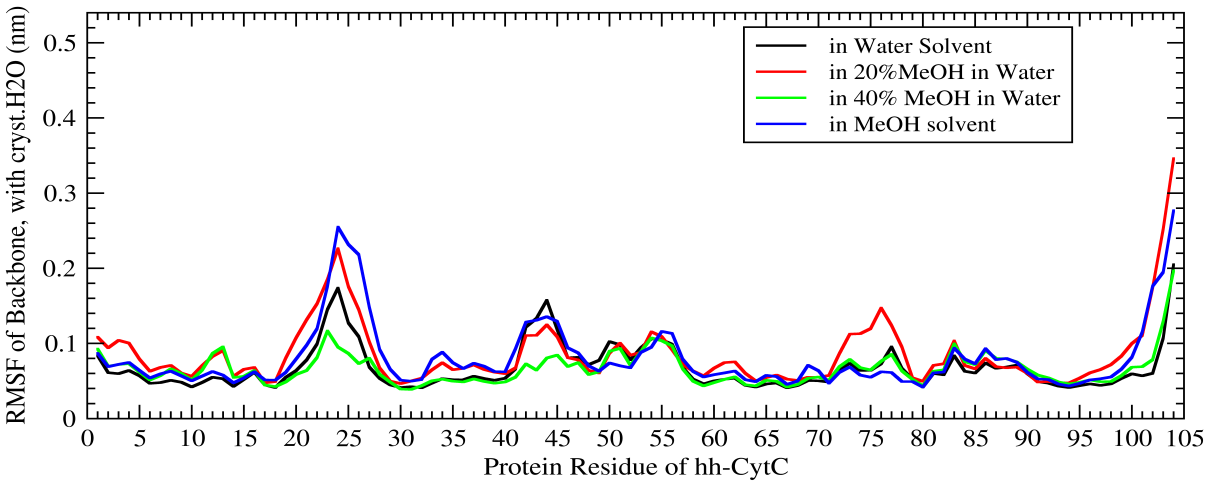


Figure-26: Time average C-α backbone RMSF of all residues of hh-CytC with keeping crystallographic water molecules in different solvents in 100 ns simulation at 298.15 K and 1 atm.

When we compare the C-α backbone-RMSF of all residues of hh-CytC with cryst.H_2O (Figure-26) and without cryst.H_2O (Figure-27), the pattern of fluctuations is the same but fluctuations are higher without cryst.H_2O. The carboxyl-terminal showed high fluctuation in water and 20% aqueous MeOH solvent compared to other solvents. Moreover, among the

highly conserved residues in Cytochrome C [8], the major fluctuations were observed in 52ASN, 76PRO and 78THR in all solvents, both with and without crystallographic water molecules. These results showed that the hh-CytC in our systems has slightly different structural and dynamical properties due to cryst.H_2O validating the significance of presence of crystallographic water which should be more important in internal motion of a protein. This result also indicates structural nuances are more critical in protein functioning along with the major tertiary structures.

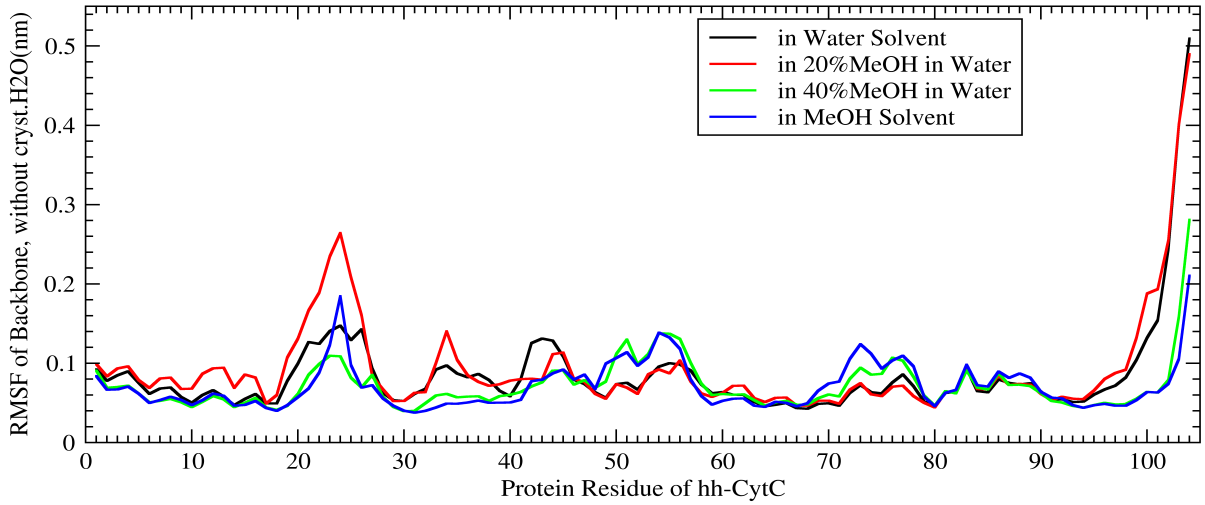


Figure-27: Time average C-α backbone RMSF of all residues of hh-CytC without keeping crystallographic water molecules in different solvents in 100 ns simulation at 298.15 K and 1 atm.

On the other hand, in our simulation condition, the RMSD of hexa-coordinated complex on Heme-Fe is quite consistent with average value 0.06 ± 0.005 Å and there is no comparative change in the axial coordination distance of protein on heme-Fe in different solvents. Moreover, the thioether linkage is observed to be quite stable in our simulation system with 0.1- 0.2 Å of fluctuations (Table-6).

Table-6: Data of average distance measured between salt-bridge atoms mentioned in crystal structure [3] and axial coordination in hh-CytC in different solvents at 298.15 K and 1 atm.

Atom Pairs	**Average distance (Å) over 100 ns md-simulation time (Max. - Min.)**				**Bond Dist. in X-ray structure from VMD (Å)**
	hh-CytC in water	**hh-CytC in 20%MeOH in water**	**hh-CytC in 40%MeOH in water**	**hh-CytC in MeOH in water**	
Salt Bridges in X-ray crystal structure					
5LYS NZ - 2 ASP OD2	6.47±0.13 11.70 – 2.89	6.23±0.12 10.80 - 2.83	6.19±0.11 11.50 – 2.93	6.27±0.12 10.06 - 2.92	6.29
38 ARG NH1 - 105 HEM O1A	4.17±0.08 8.15 – 2.78	4.29±0.08 7.85 – 2.78	4.32±0.07 7.99 – 2.82	9.91±0.27 14.30 – 2.81	4.95
53LYS NZ - 50ASP OD2	7.58±0.21 15.20 – 1.52	8.40±0.23 13.50– 2.90	6.50±0.78 13.80 – 2.90	9.30±0.27 14.70 - 2.94	7.02
99LYS NZ – 61 GLU OE2	4.62±0.33 13.10 – 2.74	4.73±0.15 12.60 – 2.75	4.16±0.11 11.20 - 2.80	4.00±0.09 9.88 - 2.78	2.76
HEM-FE and Protein axial coordination distances					
18HIS NE2 - 105 HEM FE	2.04±0.01 (2.25 -1.82)	2.05±0.01 (2.24 -1.86)	2.04±0.01 (2.30 -1.85)	2.05±0.01 (2.29 -1.82)	2.04
80MET SD - 105 HEM FE	2.32±0.01 (2.52 -2.10)	2.32±0.01 (2.56 - 2.10)	2.32±0.01 (2.52 - 2.12)	2.31±0.01 (2.55 -2.06)	2.32
Thio-ether bond distance between HEM and Protein					
14CYS SG – 105HEM CAB	1.78±0.002 (1.85 - 1.71)	1.78±0.002 (1.87 - 1.68)	1.79±0.002 (1.86 - 1.71)	1.78±0.002 (1.87 - 1.70)	1.75
17CYS SG – 105HEM CAC	1.83±0.002 (1.91 - 1.76)	1.83±0.002 (1.92 - 1.75)	1.83±0.002 (1.91 - 1.76)	1.83±0.002 (1.92 - 1.75)	1.86

The formation or rupture of salt-bridges or any specific contacts between residues inside hh-CytC may serve as an important parameter in the study of protein dynamics, its internal motion and intra-molecular contact order. The correctly folded native structure of a protein must have precise contacts among the residues creating salt-bridges, H-bonding and

hydrophobic interactions internally. Observing the changes in these contacts and interactions in hh-CytC in different solvents, we could predict the nuances of the protein folding or unfolding and hence protein functioning [42, 43, 62]. So, we analyzed the distance between salt-bridge atoms and characteristics of intra-molecular hydrogen bonding of hh-CytC. The average distance between atoms in salt bridges mentioned in x-ray crystal structure (Table-6) obviously predicts considerable solvent effect in internal structure of hh-CytC vis a vis its structure in different solvents. Methanol has significant effect on the salt-bridges of 38ARG-NH1 - 105HEM-O1A and 53LYS NZ – 50ASP-OD2. There are new other salt-bridges as shown in Table-7 were observed in our simulation based on our criteria of salt-bridges (4.0 Å between atoms involved in contacts in any trajectories in 100 ns simulation). Some of the contacts have survived in all four solvents. Those contacts which were at the distance of the most-probable non-bonded near-neighbor distance ($a' = 4.8$ Å, Appendix-VII) in all solvents may be crucial in maintaining folded conformation of hh-CytC. The amino-terminal 1GLY-NH_3^+ is consistently near the 92GLU-COO^- and 93ASP of 5th α-helix. The NH_3^+ of 5LYS has maintained electrostatic proximity to COO^- of both 92ASP and 2ASP. The 13LYS-NH_3^+ : 90GLU-COO^- pairs were closer forming H-bonding in pure solvents rather than mixed solvents. These contacts between N-terminal and C-terminal helices are important in stabilizing interactions in folded hh-CytC [8, 15, 42, 61, 62]. We did not observe 38ARG-GD^+ : 104GLU-COO^- contact in other solvent except in pure methanol. Since 38ARG-GD^+ is involved in hydrogen bonding with buried crystallographic water, 112H_2O, this water must have lost from its site without replacing by other water molecules. Similarly, hh-CytC has lost the 69GLU-COO^- : 91ARG-GD^+ contact in presence of methanol but 66GLU-COO^- : 91ARG-GD^+ contact has observed in 40% aqueous MeOH and in pure methanol solvents.

Table-7: New Salt-bridge in hh-CytC observed and distance calculated between atoms of respective salt-bridges from GROMACS in our systems of simulation in different solvents at 298.15 K and 1atm pressure at constant NPT.

Salt-bridge AA-residue Atoms in hh-CytC	**Nitrogen–Carbon distance (Å) between AA-residues that possess salt bridge groups in different solvent**				**Distance in X-ray Crystal Structure (Å)**
	Water	**20% MeOH in Water**	**40% MeOH in Water**	**MeOH**	
1GLY-NH_3^+: 92GLU-COO^-	4.27±0.87	4.28±1.01	3.6±1.0	3.94±0.97	3.77
1GLY-NH_3^+: 93ASP-COO^-	3.35±0.25	3.29±0.22	3.37±0.25	3.37±0.02	4.68
4GLU-COO^-: 7LYS-NH_3^+	6.04±1.80	6.30±1.87	6.35±1.89	4.30±0.40	8.27
4GLU-COO^-: 8LYS-NH_3^+	5.97±2.19	5.41±1.95	5.04±1.90	5.21±4.5	9.64
5LYS-NH_3^+: 93ASP-COO^-	3.38±1.04	3.84±1.16	3.42±1.0	3.23±0.53	5.09
13LYS-NH_3^+: 90GLU-COO^-	5.47±2.00	12.72±3.11	13.62±3.44	5.25±4.94	2.83
21GLU-COO^-: 25LYS-NH_3^+	8.18±2.4	8.39±2.27	4.75±1.8	8.66±3.6	7.90
22LYS-NH_3^+-: 104 GLU-COO^-	6.37±2.96	6.49±3.95	4.38±1.90	7.22±5.53	5.62
38ARG-GD^+: 104GLU-COO^-	----------	----------	----------	5.33±2.20	10.04
60LYS-NH_3^+: 62GLU-COO^-	5.24±2.30	5.20±2.07	4.17±1.64	2.53±0.17	8.79
66GLU- COO^-: 91ARG-GD^+	----------	----------	3.03±0.60	6.95±1.27	8.76
69GLU COO^-: 91ARG-GD^+	3.37±0.73	----------	----------	----------	4.22
69GLU-COO^-: 73LYS-NH_3^+	6.24±2.35	4.89±2.41	5.48±1.79	----------	10.20
69GLU-COO^-: 86LYS-NH_3^+	5.67±2.07	5.74±1.97	6.23±2.37	5.21±1.18	4.84
79LYS-NH_3^+: 105HEM-COO^-	4.36±145	3.67±0.96	3.83±1.52	3.04±0.2	4.97
87LYS-NH_3^+: 90GLU-COO^-	4.12±1.23	3.88±1.10	3.79±1.05	3.85±1.82	4.11
88LYS-NH_3^+: 92GLU-COO^-	4.91±1.63	5.19±2.06	6.14±2.19	4.46±3.84	4.56

Moreover, the oscillations (±SD) in all these contacts also infer the internal flexibility of protein structures. So, the study of these contacts or salt-bridges has also indicated that the folded of hh-CytC may exist in different solvents which may or may not be functionally important, or the rupture of some contacts may be the starting point of protein unfolding and loss of its activity.

3. Analysis of Hydrogen Bonding Characteristics within Horse Heart Cytochrome C.

Solvent effects on internal H-bonding characteristics were computed in this simulation study. In water, 17 more H-bonds were observed for hh-CytC, relative to the crystal structure, using the criteria of H-bonding of $d_{D\text{-}A} \leq 3.5$ Å and $\boldsymbol{\theta}_{D\text{-}H\text{-}A} \leq 30^0$. The calculated data of H-bond characteristics within hh-CytC are tabulated in Table-8 and Figure-28 shows the fluctuation in the number of internal hydrogen bonds along the 100 ns simulation.

Table-8: Hydrogen Bonding Characteristics within hh-CytC at different solvents at 298.15 K and 1 atm pressure at constant NPT.

Solvent	$N_{H\text{-}Bond}$	**H-bond angle, $\theta_{H\text{-}D\text{-}A}$ (^{0})**	**H-Bond distance, d_{D-A} (Å)**	**H-bond lifetime, $\tau_{H\text{-}bond}$ (ps)**	**Diff.-Const. of hh-CytC (×1e-7 cm^2s^{-1})**
Crystal hh-CytC [3]	75*[3]	15.30▲	2.96±0.24	-------	----------
Water	88.9±2.44	15.56±1.14	3.02±0.13	326.8	0.340 ± 0.223
20%MeOH in Water	91.8±3.03	15.96±2.76	3.02±0.17	662.6	0.354 ± 0.166
40% MeOH in Water	96.3±2.19	15.95±2.11	3.03±0.11	315.7	0.512 ± 0.076
MeOH	99.3±1.19	16.23±1.56	3.02±0.27	553.8	0.508 ± 0.086

*with the criteria of H-bonds: $d_{H\text{-}A} = 2.6$ Å and a $\theta_{D\text{-}H\text{-}A} = 120^0$ | ▲ standard deviation of $\theta_{D\text{-}H\text{-}A}$. [3]

With increasing MeOH concentration, the number of intra-protein H-bonds has also increased with a slight increase of H-bond angles ($\boldsymbol{\theta}_{H\text{-}D\text{-}A}$) as shown in Figure-29 and Table-8, even though the H-bond distance ($d_{D\text{-}A}$) distance did not vary significantly (Figure-30, Table-8).

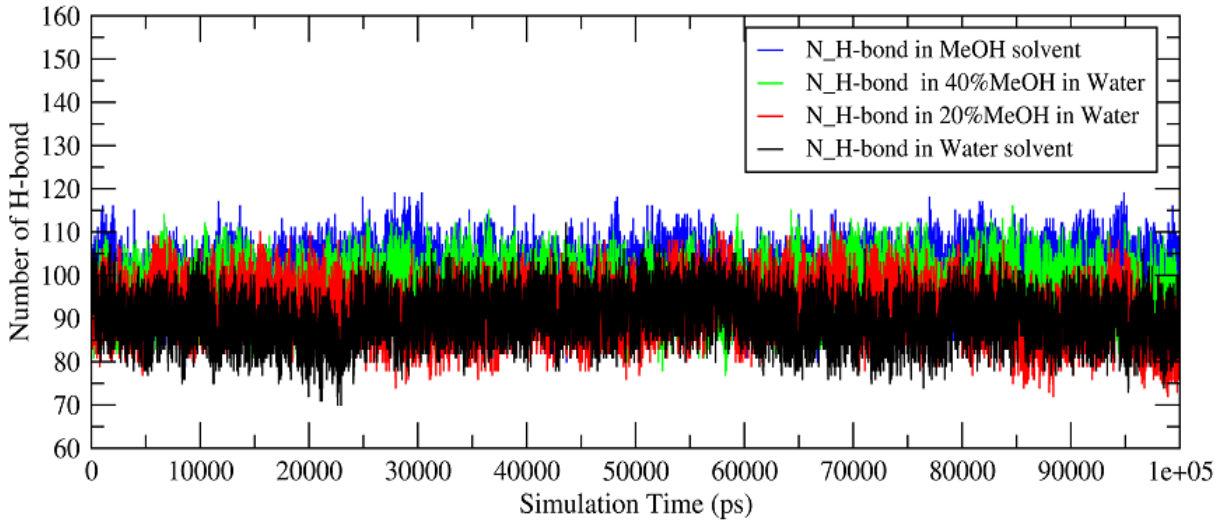


Figure-28: Time evolution of numbers of intra-protein Hydrogen Bonds within hh-CytC in different solvents at 298.15 K temp. and 1 atm pressure in 100 ns in constant NPT simulation.

In general, at low hydration, the protein surface is less rigid due to hydrophobic methyl group which replaces water from protein surface and protein will hydrogen bond with itself when not enough H_2O molecules are available at solvation layers; whereas at higher hydration, the H_2O molecules in solvation layers compete successfully for hydrogen bonding with donors and acceptors of hh-CytC and MeOH and succeeds in breaking the surface H-bonds of hh-CytC, leading to the overall reduction of intra-protein H-bonds at surface, apparently increasing surface rigidity of protein, and favoring solvent exposure of polar residues. These characteristics of internal H-bond along with the R_g values in different solvents gives the idea of swelling up of hh-CytC having molten protein surface with increase of MeOH concentration.

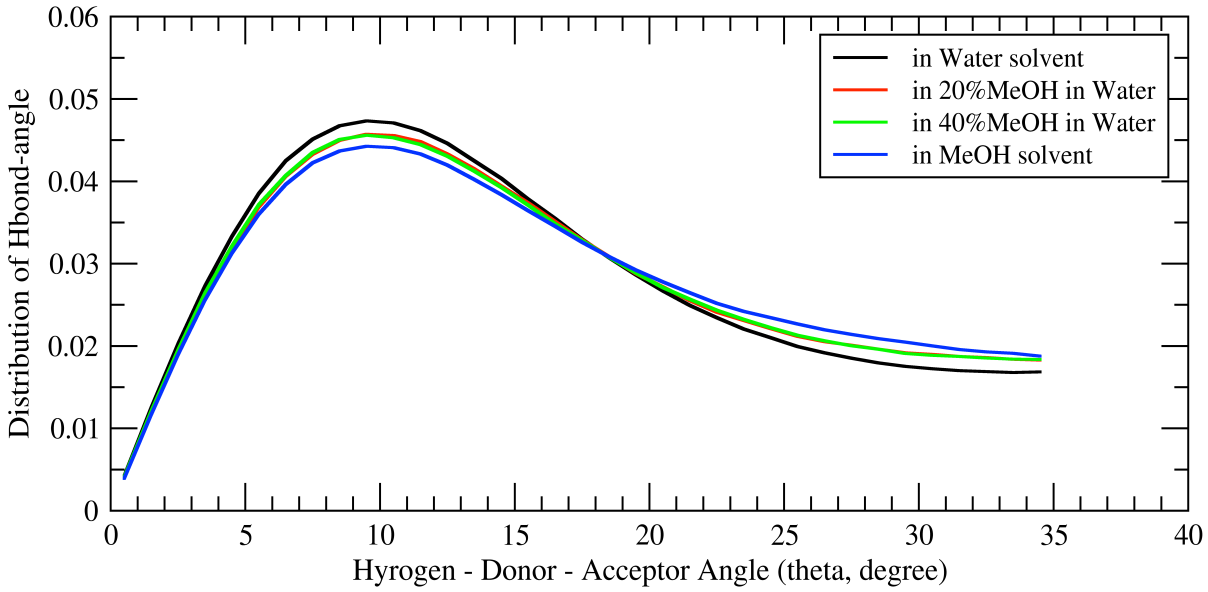


Figure-29: Distribution of intra-protein Hydrogen Bond Angle ($\theta_{H\text{-}D\text{-}A}$) within hh-CytC in different solvents at 298.15 K and 1atm.

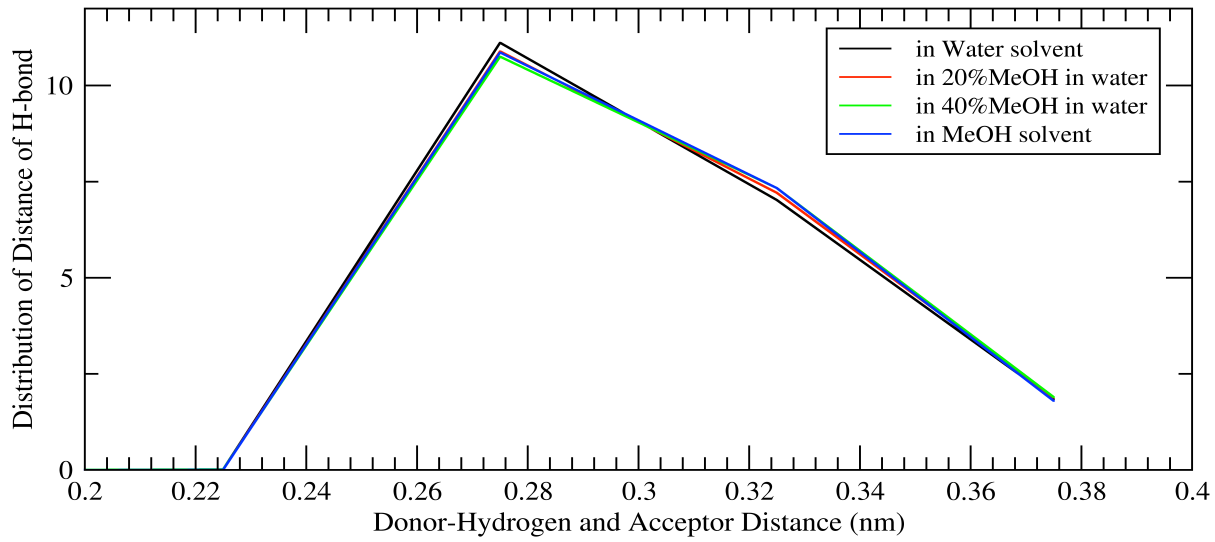


Figure-30: Distribution of intra-protein Hydrogen Bond Distance ($d_{D\text{-}A}$) within hh-CytC in different solvents at 298.15 K and 1atm.

The average lifetime for intra-protein H-bonds was calculated using intermittent H-bond autocorrelation function from GROMACS as stated by A. Luzar, and van der Sopel *et al* [44, 45]. The hydrogen bond time autocorrelation function (ACF), $C_{HB}(t)$ of intra-protein H-bond of hh-CytC as shown in Figure-31(using geometric criteria: $r_{(D-A)} \leq 3.5$ Å and Angle H-D-A $\leq 30^0$) at 298.15 K and 1 atm pressure, gives the probability that a H-bond between a pair of donor (D) and acceptor (A) exists at $t = 0$ and still exists at time t even if the bond breaks at some intermediate time. Even though the decay of ACF is sharper in pure solvents that in mixed solvents, a the ACF remains above 0.65 in all solvents, which corroborates that H-bonds inside the protein are highly stable, more stable in mix-solvents than in pure solvents, and these H-bonds are highly correlated with long average H-bond lifetime, as these H-bonds always exist within a protein. Even though the number of H-bonds within hh-CytC increases with MeOH concentration, the H-bond lifetime is relatively higher in mix-solvents.

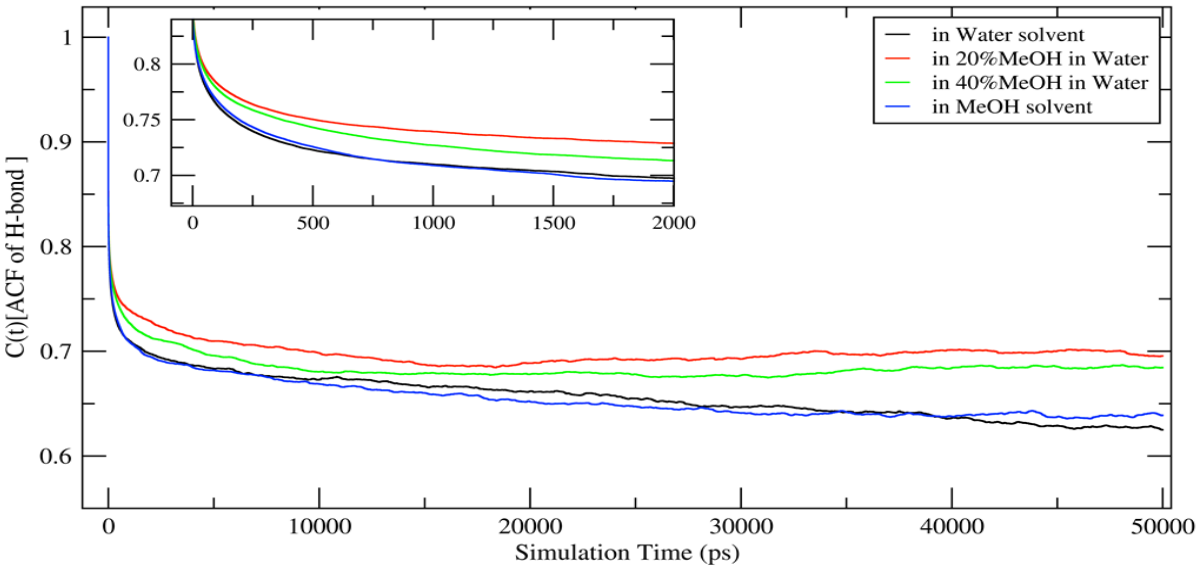


Figure-31: Time Auto-Correlation Function (ACF) of intra-protein Hydrogen Bond within hh-CytC in different solvents at 298.15 K and 1 atm at constant NPT condition.

The hundreds of picoseconds of H-bond lifetime with increase of number of H-bond within hh-CytC implies lower compatibility of hh-CytC with methanol solvent, even though other structural analysis parameters like RMSD, RMSF, and R_g look good. The ACF of H-bond displayed in Figure-31 shows that H-bonds within hh-CytC in mix-solvents are more stable than in pure solvents, which may be due to effect of non-ideal behavior of water-methanol solvent mixture; the ACF never cascades down below 0.65 even in water solvent in 100 ns simulation which should be due to persistent α-helix backbone H-bonding in protein to maintain secondary structures.

4. Dynamics of hh-CytC in Different Solvents and Protein-Solvent Interface.

4.1 Mean Square Displacement of hh-CytC.

Figure-32 display the mean-square displacement, MSD, of hh-CytC in different solvents at 298.15 K and 1 atm pressure in different ensembles. The protein moves very slowly relative to solvent molecules. In solution, the MSD of a molecules following Brownian motion grows linearly with time [27], but the MSD of hh-CytC in different solvents and in different ensembles reveals that the motion of hh-CytC is restricted and it moves asymptotically away from its initial position where its translational motion may be governed by the internal protein anharmonic motions or conformational fluctuations of hh-CytC along with solvent drag. In spite of having high viscosity differences in all four solvents (Table-3), the protein's movement in different ensembles was not found to be correlated with solvent viscosity or density, as expected, in whole simulation time. Rather it was found that the protein moved relatively slowly with linear increase of MSD in initial 20-40 ns, suggesting that solvent viscosity may be regarded as the most important source of friction in initial protein motion.

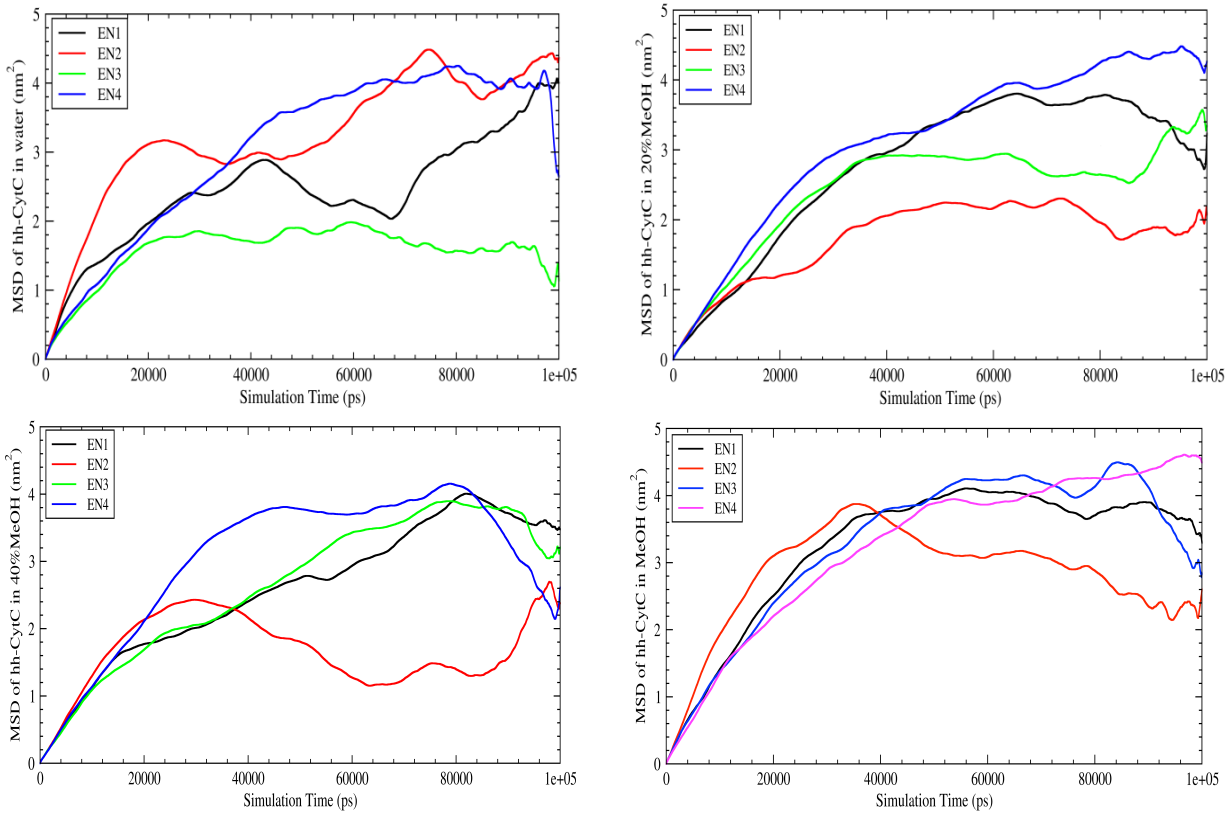


Figure-32: Mean Square Displacement (MSD) of hh-CytC in different solvents over 100 ns simulation time in 298.15 K temperature and 1 atm pressure in different ensembles.

The simulation time to reach maximum value of MSD of hh-CytC is different in different solvents and in different ensembles as depicted by Figure-32, which might depend on mechanism of conformational transition of a protein in different environment in different ensembles. After linear increase of MSD, hh-CytC has started to decelerate or move asymptotically. The MSD decreases slowly and is controlled by inernal motion of protein rather than diffusive motion in different solvents even though we did not observe any significant variation of RMSD in the second half of the simulation except in 20% aqueous MeOH solvent. Since the movement of hh-CytC is found to be asymptotic and the MSD has

not increased linearly in whole 100 ns simulation time, it is relatively difficult to compute the physically meaningful average diffusion coefficient of hh-CytC from a single molecule simulation method in different ensembles. Still, the diffusion coefficient of protein calculated from GROMACS in different solvent compositions in our simulation (Table-8) from first 40 ns. Based on these observations, we may guess the protein's internal motion and solvent properties for frictional drag are equally responsible for MSD of hh-CytC and hence the diffusion constant. Since the MSD of hh-CytC is more sensitive to its slow translational motions [58], it demands much longer simulations time than 100 ns in order to obtain convergence for MSD that should reach a plateau value.

4.2 Analysis of Hydrogen Bond Characteristics in Protein-Solvent Interface.

The H-bond characteristics between added solvent and hh-CytC at 298.15 K temperature and 1 atm pressure were computed and analyzed. Table-9 and Figure-33 have revealed that the total number of H-bonds between solvent molecules (both MeOH and H_2O) and hh-CytC has decreased with increasing MeOH concentration which may be because of the methyl group in MeOH that excludes certain space around hh-CytC so that there is a lower number of solvent molecules in the solvation layer around the protein available for hydrogen bonding. Moreover, a protein has a higher probability of hydrogen bonding with H_2O compared to MeOH because of the same methyl group. And the data has also shown that hh-CytC has more H-bonds with pure water solvent compared to pure methanol solvent. In presence of MeOH, some of the H-bonding sites at the surface of hh-CytC remain free which may form intra-protein H-bonds if they are geometrically accessible.

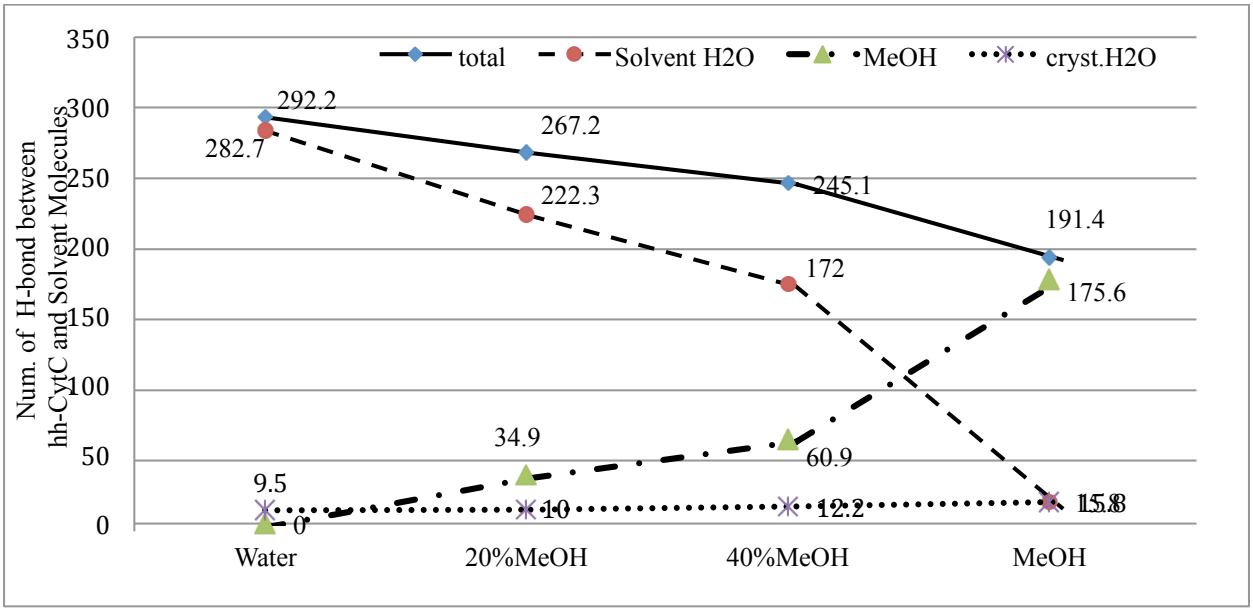


Figure-33: Graph of number of H-bonds in different simulation system, depicting the variation of number of H-bond between solvent and hh-CytC at interface at 298.15 K temperature and 1 atm pressure.

Table-9: Data of Hydrogen Bond Characteristics between hh-CytC and added solvents, H_2O and MeOH in different solvent composition at 298.15 K temperature and 1 atm pressure.

Solvent	**$N_{H\text{-}Bond}$**	**H-bond angle, $\theta_{A\text{-}D\text{-}H}$ (deg.)**	**H-bond distance (Å), $d_{D\text{-}A}$**	**H-bond lifetime, $\tau_{H\text{-}bond}$, (ps)**
Between hh-CytC and added H_2O solvent				
Water	282.7±3.2	16.82±1.94	3.01±0.21	11.5
20%MeOH in Water	222.3±4.1	16.51±2.33	3.00±0.86	32.3
40%MeOH in Water	172.0±3.2	16.46±2.69	2.99±0.45	14.5
Between hh-CytC and added MeOH solvent				
20%MeOH in Water	34.9±3.9	16.48±2.91	3.01±0.21	11.7
40%MeOH in Water	60.9±3.4	16.40±2.77	2.99±0.13	7.6
MeOH	175.6±2.1	16.16±1.95	2.97±0.09	16.6

Even though there is a significant variation in the number of H-bond between solvents molecules and hh-CytC in different solvent compositions, the analysis of hydrogen bond angles and donor-acceptor distance as shown in Figures - 34 & 35 respectively and data from Table-9 indicate that hydrogen bond geometries are virtually identical which has varied very slightly among all solvent compositions.

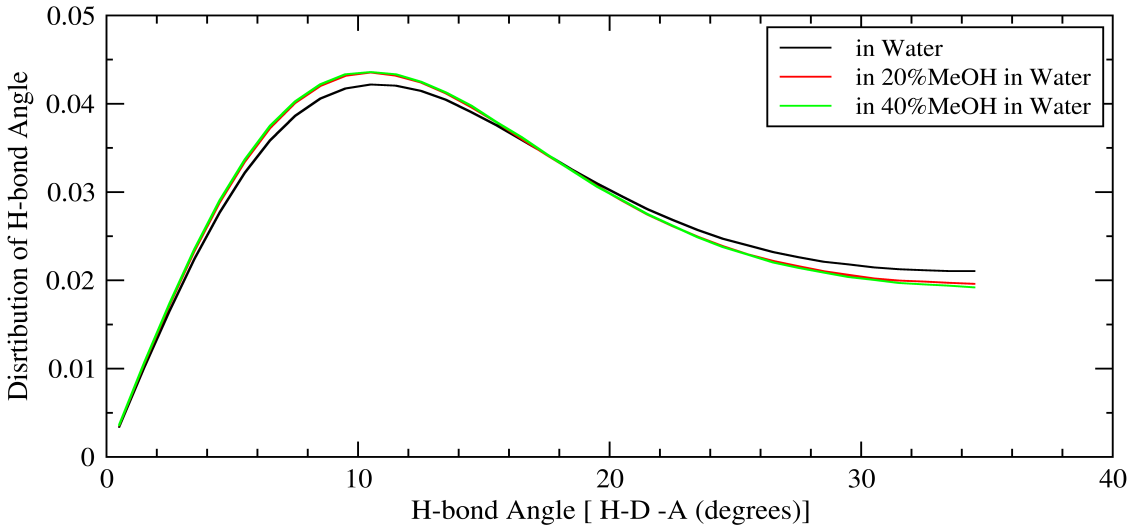


Figure-34: Distribution of Hydrogen Bond Angle, $\theta_{H\text{-}D\text{-}A}$, between hh-CytC and solvent H_2O molecules at solvent-protein interface (solvation layer) in different solvent compositions at 298.15 K and 1atm pressure.

The solvent molecules in solvation layer hop around at the surface of protein by diffusion between sites on the protein surface and/or exchange with bulk solvent. The intermittent H-bonding lifetime and auto-correlation function between solvent molecules and hh-CytC was computed from GROMACS (Table-8). Hydrogen bonds between solvent and hh-CytC survive longer than H-bonds between solvent molecules in the bulk and, moreover, the H-bond lifetime between added solvent H_2O and hh-CytC, $\tau_{hh\text{-}CytC\text{-}H2O}$, is shorter than the H-

bond lifetime between MeOH and hh-CytC, $\tau_{hh\text{-}CytC\text{-}MeOH}$, in pure solvents, but this is reversed in mix-solvents.

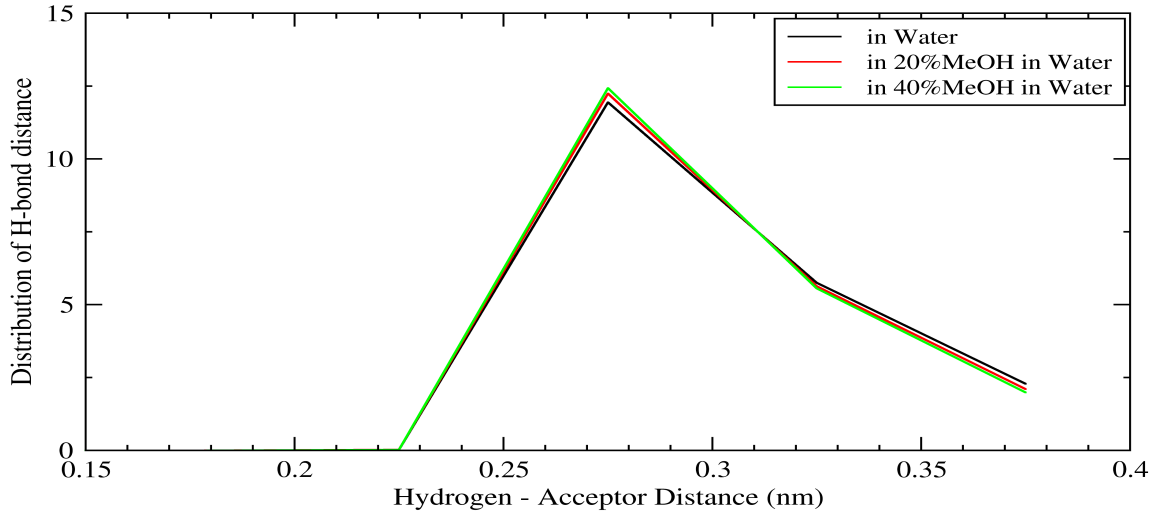


Figure-35: Distribution of Hydrogen Bond Donor-Acceptor Distance, $d_{D\text{-}A}$, between hh-CytC and solvent H_2O molecules at solvent-protein interface (solvation layer) in different solvent compositions at 298.15 K and 1atm pressure.

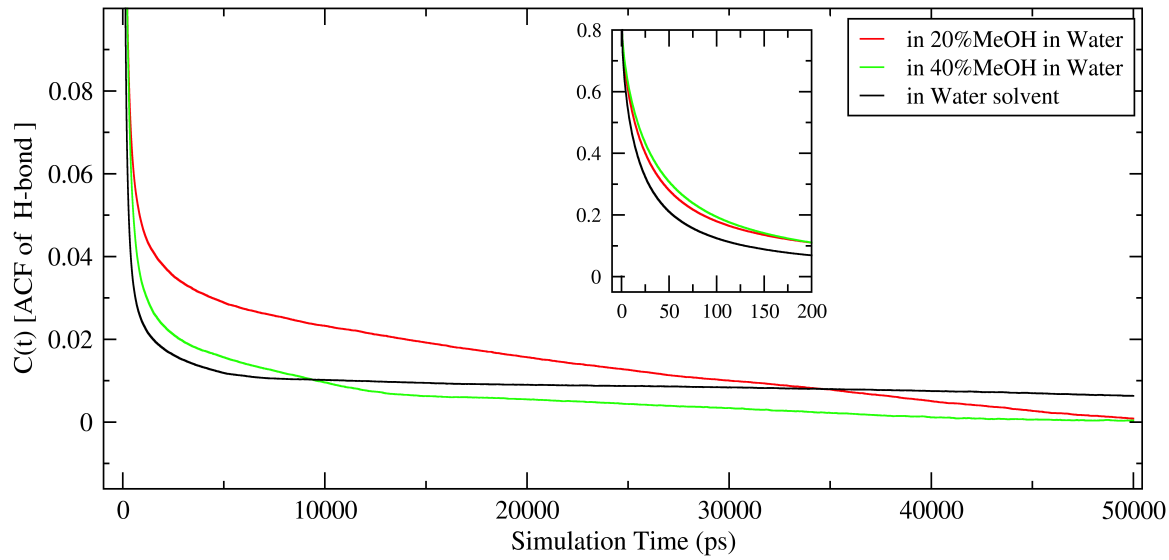


Figure-36: Time Auto-Correlation Functions of Hydrogen Bond between hh-CytC and H_2O (added solvent) in different solvent compositions at 298.15 K and 1atm pressure.

In short time scale, the H-bond ACF between H_2O (solvent, added) and hh-CytC show similar slower decay in mix-solvents compared to that in pure water solvent as shown in the inset of Figure-36; but after longer time, ACF in the mix-solvents decays down to a lower value than in pure water solvent; indicating that H-bond existence between solvent H_2O and hh-CytC may be relatively easy but less stable initially (easily formed and easily broken) in pure solvent even though after long time H-bond become less stable in mix-solvents than in pure water solvent. The long H-bond lifetime and slow decay of ACF in 20% aqueous MeOH solvent shows unusual persistent interaction of solvent H_2O and hh-CytC. After 10 ns, ACF attains plateau in pure water solvent indicating equilibrium H-bond condition; but in mix-solvents, ACF decays relatively slowly and ultimately achieves lower plateau after 50 ns.

On the other hand, ACF of H-bond between MeOH and hh-CytC as shown in Figure-37 decays faster in mix-solvents compared pure MeOH solvent and attains plateau after 20 ns, but in pure methanol very slow decay was observed attaining plateau after 45 ns. Obviously, with increasing MeOH concentration, the chances of MeOH to form H-bond with protein could be higher, but the protein's behavior and conformation should also play crucial role in MeOH-protein interaction.

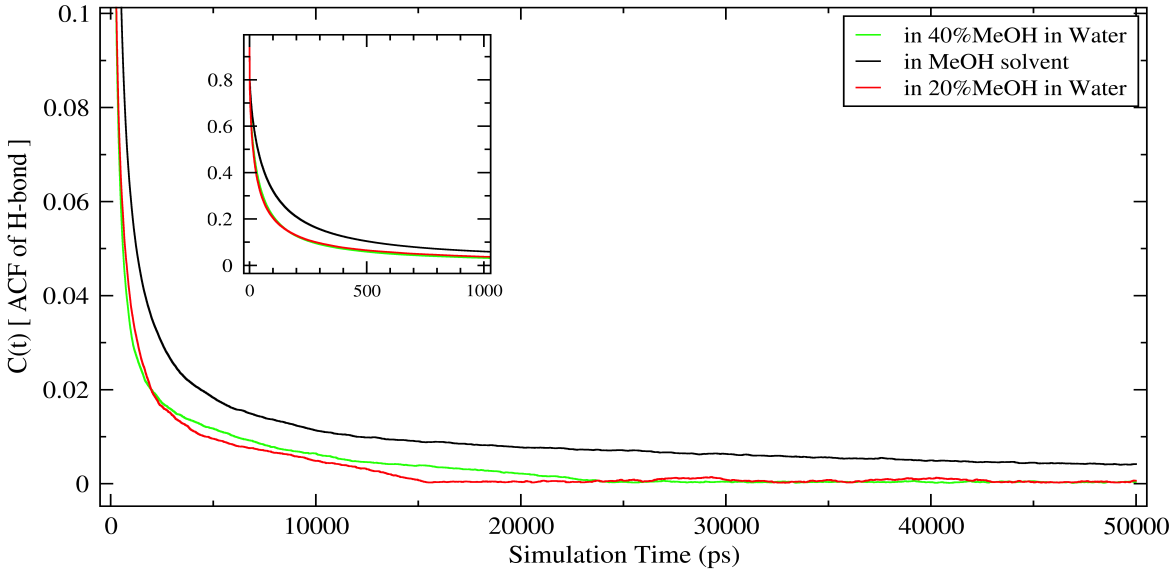


Figure-37: Time Auto-Correlation Functions of Hydrogen Bond between hh-CytC and MeOH in different solvents at 298.15 K and 1atm in constant NPT condition.

When we compare ACF of H-bond of H_2O and MeOH with hh-CytC in Figures - 36 & 37 respectively, hydrogen bonding between solvent water molecules and hh-CytC is faster and shorter (higher H-bond dynamics) in pure solvents compared to H-bonding between MeOH and hh-CytC, but opposite in mix solvents. The initial slower decay of H-bond ACF for MeOH compared to that of H_2O indicate less compatible interaction of hh-CytC with methanol. The ACF after initial sharp decay transient period is governed by continuous rearrangement of the protein-solvent H-bond network leading to a plateau of the ACF that infers the equilibrium hydrogen bonding states independent of time and surrounding conditions [59]. Figures-36 and 37 depict that H_2O has a higher plateau value of ACF compared to that of MeOH.

Even though all the $C_{HB}(t)$ or H-bond ACF display an initial sub-picosecond or picosecond transient decay, comparing all the ACF figures, the protein-solvent H-bond ACF decays somewhat faster than intra-protein H-bond ACF, which is much slower than bulk

solvent H-bond ACF indicating the role of hydrogen bonding and solvent mobility in modulating protein conformations in exploring biophysical phenomena.

5. Study of Crystallographic Water in Horse Heart Cytochrome C.

Figure-38 displays the radial distribution of crystallographic water molecules (cryst.H_2O) around the hh-CytC surface. Th small feature at 1.25 Å should be because of the buried water molecules. There are two prominent RDF peaks at 2 Å and 4.2 Å indicating the first and second solvation layer positions of cryst.H_2O from protein surface. The RDF peak of cryst.H_2O increases with increasing MeOH percentage, indicating that cryst.H_2O molecules are closer to protein with increasing number of MeOH molecules (or a decreasing number of solvent H_2O molecules which can replace them).

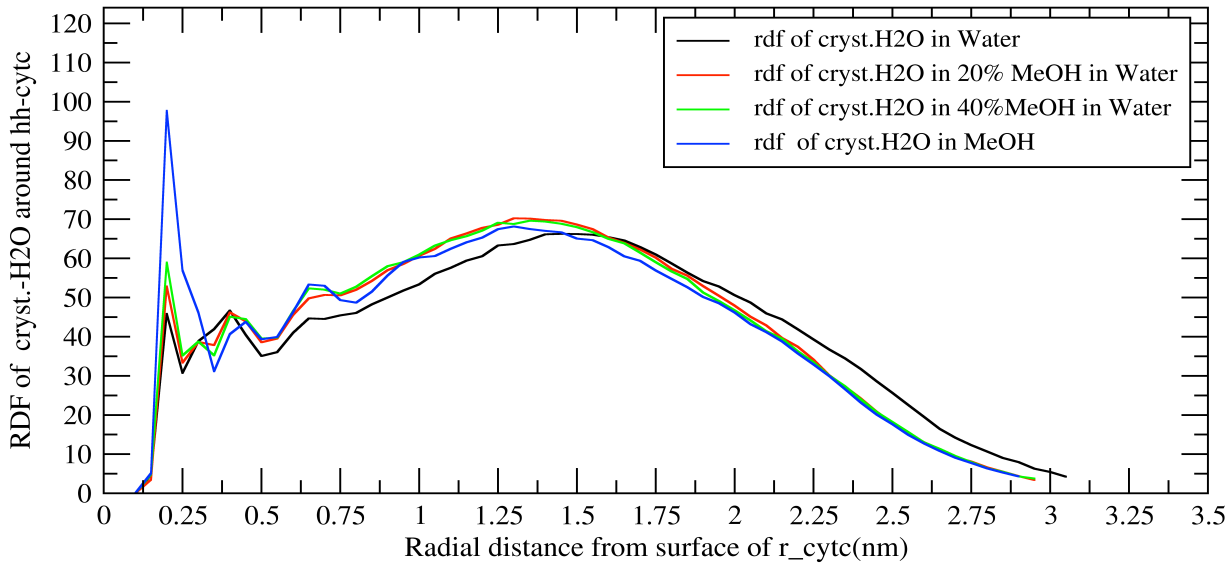


Figure-38: Radial Distribution Function, RDF, of crystallographic water around the surface of hh-CytC in different solvents at 298.15 K temperature and 1 atm pressure at constant NPT.

The cryst.H_2O molecules lose their identity in the presence of solvent water molecules. A small RDF peak was also observed at 6.5 Å and it is more prominent in MeOH solvent. Figures – 39(A, B, C, D) and Figures – 40 (E & F) are the van Hove Distribution Function (VHDF) of cryst.H_2O molecules which give average distance moved by cryst.H_2O during simulation. Some of the cryst.H_2O molecules diffuse away from protein vicinity in first equilibration step (warming up of simulation system) as shown by Figures – 39(A) & 39(B), indicating that not all 124 cryst.H_2O mentioned in X-ray crystal structure are structurally and functionally important and they may be equivalently replaced by solvent molecules. The cryst.H_2O molecules which are retained in distance of 5 Å in second equilibration step, VHDF shown Figure - 39(C) & 39(D), should have special importance in structure and function of hh-CytC. Observing both RDF and VHDF of cryst.H_2O, diffusion of cryst.H_2O molecules far from protein surface to bulk was found to be hindered in presence of MeOH, or water preferentially may solvate in some regimes when hh-CytC is in aqueous methanol. The VHDF of cryst.H_2O for first 1000 picoseconds in data production step, shown in Figures - 40(E) & 40(F), shows that there should be some molecules which did not move at all in 1000 ps after equilibration, indicating that some cryst.H_2O molecules are constrained which might be the structural water molecules as mentioned in literature [3, 9, 15]. But, in contrast to the equilibration steps, the diffusion of cryst.H_2O was found to be faster in pure water solvent compared to pure methanol after equilibration in data production step. The VHDF peaks at 2-5 Å in pure water and 2-7 Å in pure methanol may infer the movement of cryst.H_2O in solvation layer or buried internally. So, the VHDF and RDF have neatly indicated that many of the cryst.H_2O molecules which were at 1-4.5 Å away from hh-CytC in X-ray crystal structure slowly diffuse into bulk solvent, whereas some of the cryst.H_2O are intimately associated with protein.

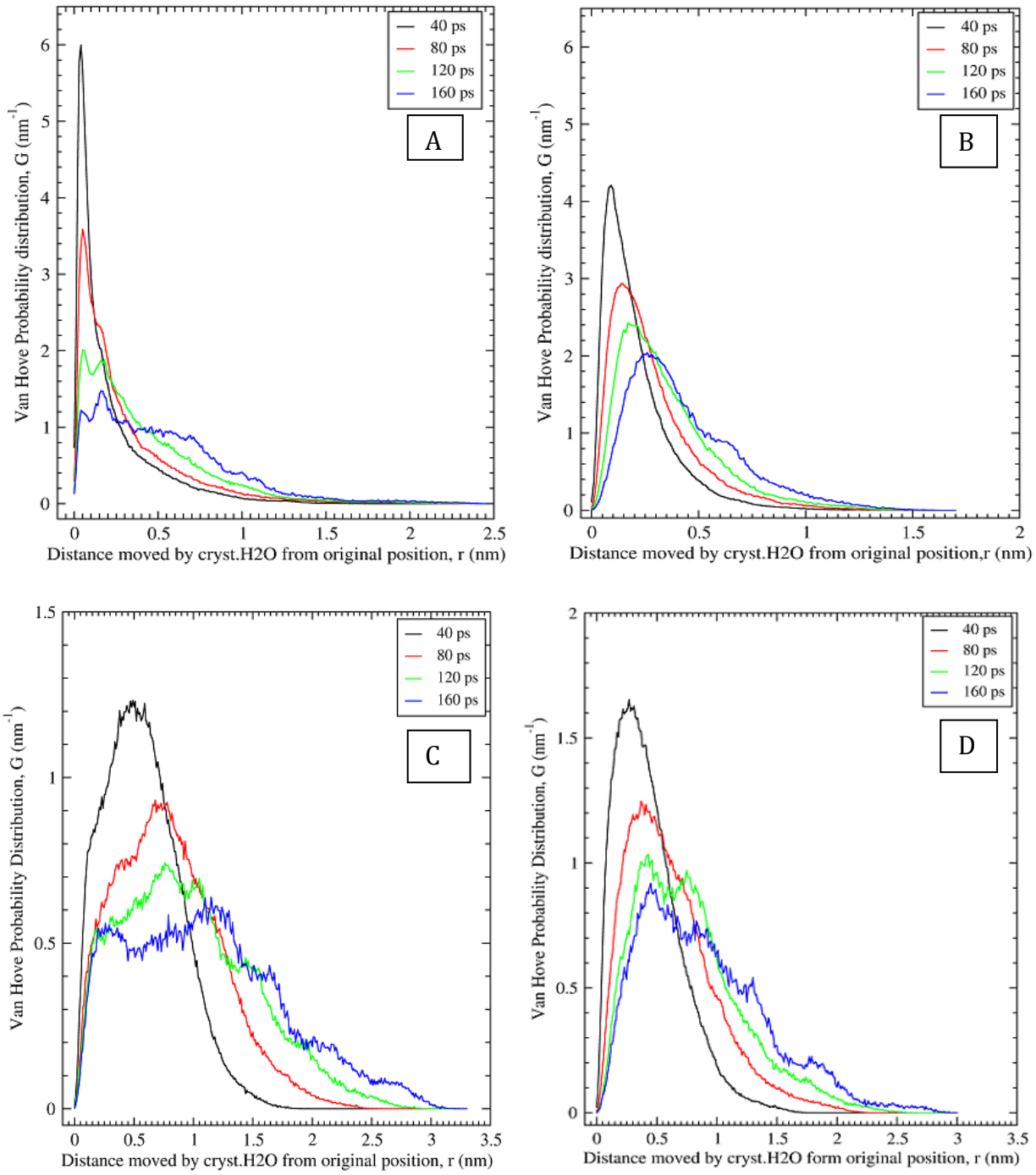


Figure-39: van Hove Distribution Function (VHDF) of cryst. water molecules at 298.15 K and 1 atm pressure in constant NPT; (A): in pure water in first equilibration step (top-left), (B): in pure methanol in first equilibration step (top-right), (C): in pure water in second equilibration step (down-left), (D): in pure methanol in second equilibration step (down-right).

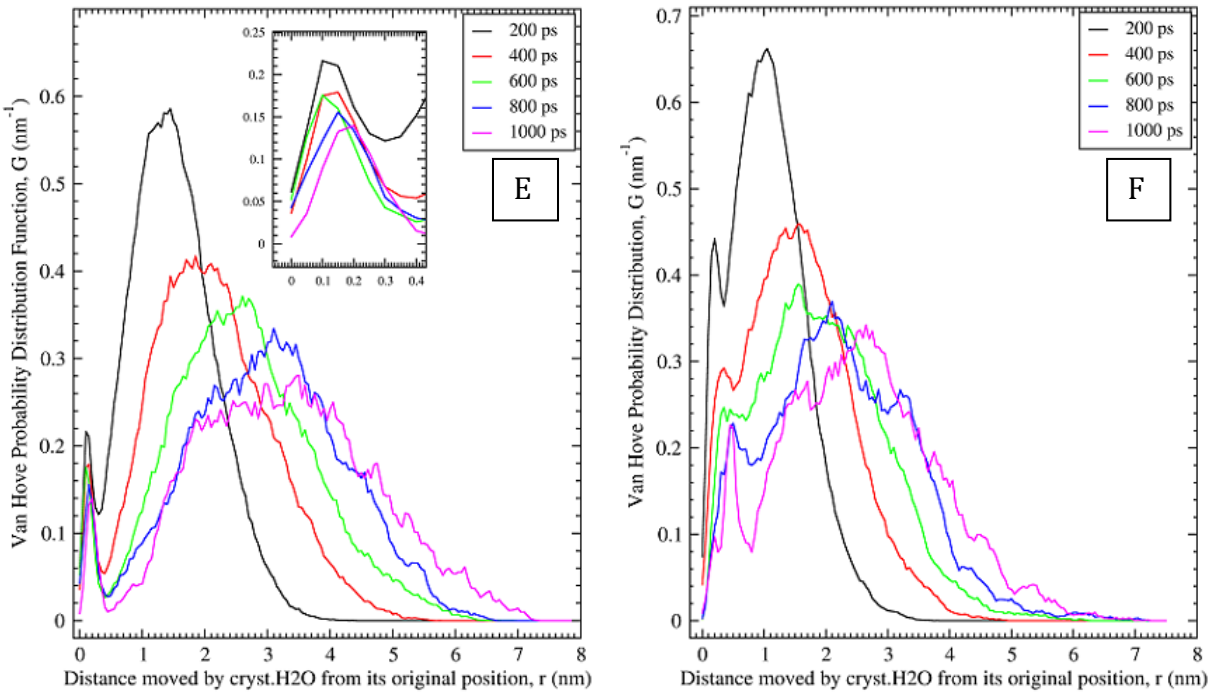


Figure-40: van Hove Distribution Function (VHDF) of crystallographic water molecules in data production step in one ensemble of simulation at 298.15 K and 1 atm pressure in constant NPT; (E): in pure water (left), (F): in pure methanol (right).

The characteristics of hydrogen bonding between cryst.H_2O and hh-CytC were analyzed and data are shown in Table-10. The number of H-bond between cryst.H_2O and hh-CytC decreases within the first 2 ns in all of the simulation systems as illustrated in Figure-41 for 20% aqueous MeOH solvent, and it reaches an equilibrium value of number of H-bond around 10-20.

The average number of H-bonds between cryst.H_2O and hh-CytC has increased with increasing MeOH concentration, and this is because the solvent water molecules equivalently and more easily replace cryst.H_2O from H-bonding sites of protein surface compared to MeOH. Minimum four H-bonds were observed in all system in 100 ns simulation indicating that these

residual H-bonds between cryst. H_2O and hh-CytC should be due to more persistent buried or trapped surface water molecules at any instant.

Table-10: Data of Hydrogen Bonding Characteristics between crystallographic water molecules and hh-CytC in different solvents at 298.15 K and1atm.

Solvent	$N_{H\text{-}Bond}$	**H-bond angle,** $\theta_{A\text{-}D\text{-}H}$ (^{0})	**H-Bond distance,** d_{D-A} **(Å)**	**H-bond lifetime,** $\tau_{H\text{-}bond}$ **(ps)**
Water	9.5±1.1	16.02±2.89	2.95±0.26	264.5
20%MeOH in Water	10.0±1.1	16.51±2.22	2.97±0.31	1146.6
40%MeOH in Water	12.2±1.3	16.84±2.12	2.95±0.21	454.6
MeOH	15.8±1.4	15.51±1.79	2.94±0.11	184.7

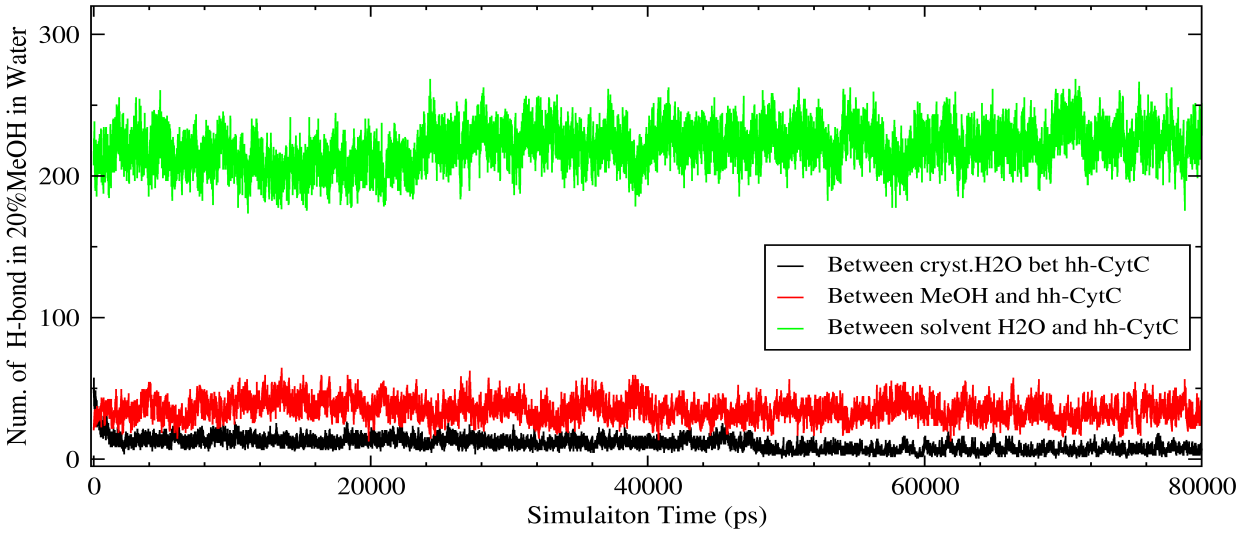


Figure-41: Illustration of time evolution of number of H-bonds between hh-CytC and solvent molecules and cryst.H_2O molecules in 20% aqueous MeOH solvent in one analyzed ensemble at 298.15 K and 1atm pressure.

The fluctuations may be simply the outcome of probability of closeness of donor-accepter partners because of random dynamics along simulation time. The analysis of

hydrogen bond donor-acceptor distance and angle between cryst.H_2O and hh-CytC as shown in figures- 42 & 43 respectively and Table-10 indicates that hydrogen bond geometries remain relatively constant among all solvent compositions.

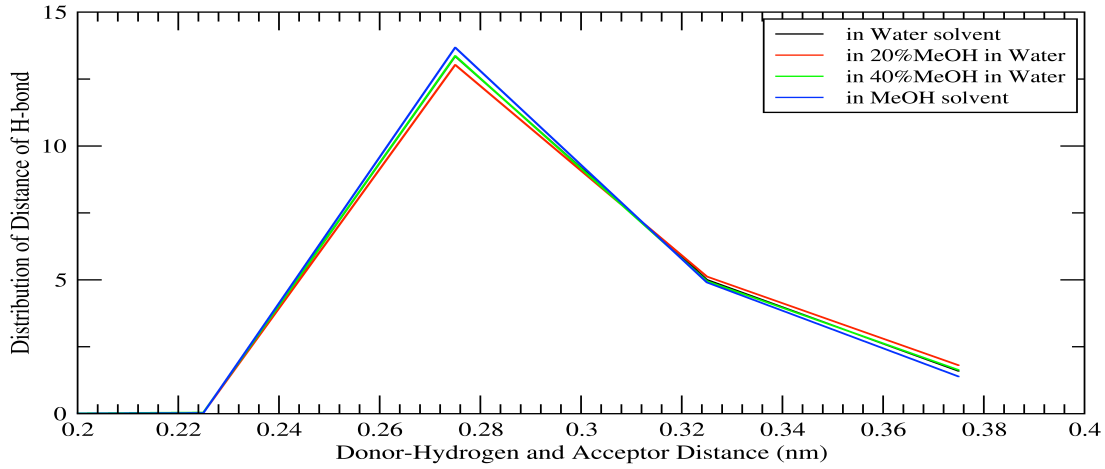


Figure-42: Distribution of Hydrogen Bond Donor-Acceptor Distance, $d_{D\text{-}A}$, between crystallographic water molecules and hh-CytC in different solvents in analyzed ensemble at 298.15 K and 1atm.

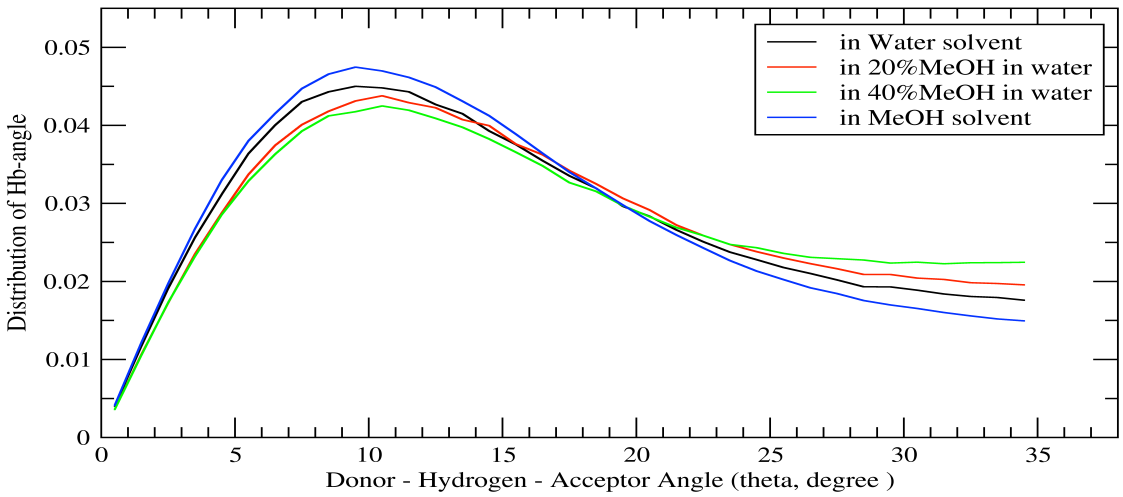


Figure-43: Distribution of Hydrogen Bond Angle, $\theta_{H\text{-}D\text{-}A}$, between crystallographic water molecules and hh-CytC in different solvents in analyzed ensemble at 298.15 K and 1atm.

The calculated intermittent H-bond lifetime between cryst.H_2O and hh-CytC is in hundreds of picoseconds range and varies unexpectedly in different solvents. It is the longest in 20% aqueous MeOH solvent (1146.6 ps) with a value of 454.5 ps in 40% aqueous MeOH solvent, and 264.5 ps and 184.7 ps for pure water and pure methanol solvents respectively. These data of long picosecond scale H-bond lifetime clearly indicate that some of the cryst.H_2O molecules are constantly in close contact with protein at the surface or buried internally within hh-CytC as stated by previous research [3, 6, 9] forming persistent H-bonds with protein having longer H-bond lifetime. Even though the number of H-bonds between cryst.H_2O molecules in pure water solvent are fewer than that in MeOH solvents, their H-bond lifetime in pure water solvent is higher than that in MeOH solvent which, maybe, due to the effect of MeOH in surface-structure flexibility in hh-CytC as indicated by Lindemann's disorder index and H-bonding of MeOH with cryst.H_2O or due to slower diffusion of cryst.H_2O in the bulk in presence of MeOH. The unexpectedly high H-bond lifetime of cryst.H_2O with protein in mix-solvent is very difficult to interpret. It may be due to the non-ideal behavior of water-methanol solvent mixtures. In mixture, the dielectric constant of MeOH drops by 80-90% so that it cannot perturb existing H-bond in the protein surface with cryst.H_2O which may result long H-bond lifetime. The ACF of H-bond between cryst.H_2O and hh-CytC (Figure-44) has depicted that hydrogen bond is more correlated in 40% aqueous MeOH than in 20% aqueous MeOH composition solvent even though H-bond lifetime is longer in 20% aqueous MeOH solvent.

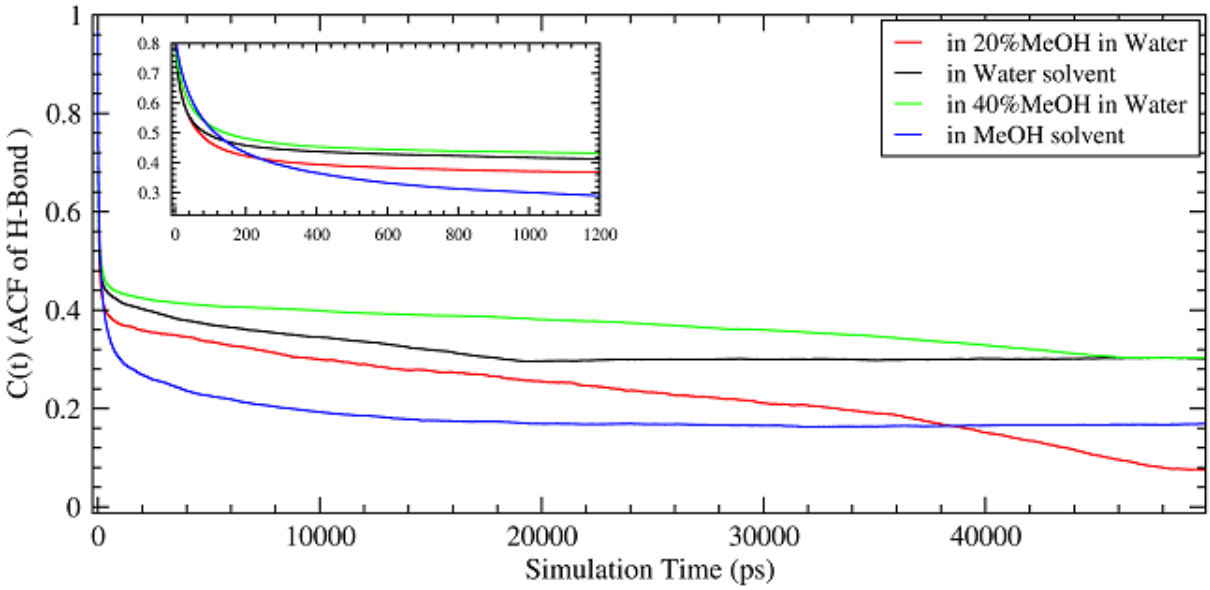


Figure-44: Time Auto-Correlation Function of Hydrogen Bond between crystallographic water molecules and hh-CytC in different solvents at 298.15 K and 1atm pressure.

The crystallographic water molecules which are identified by Bushnell, G.W. *et al* [3] and observed in NMR experiment [9] in the conserved positions of protein are mentioned in Table-11. These cryst.H_2O molecules are hydrogen bonded with protein atoms or may be trapped in hydrophobic surfaces in the sites where they located. Since these water molecules have been proposed to play a role in the mechanism of action of hh-CytC, they were tracked individually in our studies. Three cryst.H_2O on the surface of the protein: $107H_2O$, $128H_2O$, and $142H_2O$ has lost their contact with the protein in equilibration steps, diffusing rapidly in bulk solvent losing their identity, and their positions are replaced by different solvent H_2O or MeOH molecules (about 30-45 solvent molecules in our 100 ns simulation time) that are in the solvent layer around the surface of hh-CytC. These surface H-bonding sites of proteins are important in stabilizing local segments of polypeptide chain [3, 16], but are not occupied by specific cryst.H_2O or solvent molecules for a long time in our simulation, indicating that the

identity of any single H_2O or any other solvent molecules in these surface H-bonding sites are not so important in structural stability of hh-CytC, rather that solvation layer H-bonding should be important to maintain the stability of protein.

Table-11: Conserved crystallographic water molecules in defined positions or conserved sites in hh-CytC [3].

Water Number in Protein Crystal structure	**B-factor (A^2)**	**Protein Atoms involved in H-bonding (Conserved Positions for bound water)**	**Location of Position**
107H_2O	29.6	70ASN-OD1, 72LYS-N(backbone), 82PHE-O(backbone)	Surface
112H_2O	36.4	52ASN-OD1, 67TYR-OH, 78THR-OG1	Buried
125H_2O	22.6	39LYS-O(backbone), 42GLN-N(backbone), 105HEM-Propionate-O1A, 38ARG-guanidino	Buried
128H_2O	23.6	79LYS-0(backbone), 81ILE-N(backbone)	Surface
142H_2O	31.0	36PHE-N(backbone), 102THR-O(backbone)	Surface

The two crystallographic water molecules, 112H_2O and 125H_2O, were buried inside the protein near to heme. These two internal cryst.H_2O molecules seems to play more crucial role in electron transfer mechanism in hh-CytC [3]. We tracked the dynamics of individual molecules and measured the distance between these cryst.H_2O molecules and their H-bonding protein atoms in the conserved sites where they reside as mentioned in Table-11. Moreover, previous research has tried to estimate experimentally the bound water-bulk water exchange time with value between 10^{-8} to 10^{-2} s supporting the proton exchange mechanism between bound water and amide groups [15, 16]. We also observed that the buried cryst.H_2O molecules, in our 100 ns simulation, were typically found in nanoseconds timescales.

The first buried crystallographic water, cryst.112H_2O shown in Figure-45 is H-bonded with sidechain hydroxyl groups of 52ASN, 67TYR and 78THR amino acids which construct the conserved site next to heme group inside protein matrix for water molecules. Figures - 45 to 55 display the distances measured between cryst.112H_2O with its H-bonding partners inside protein in different solvents.

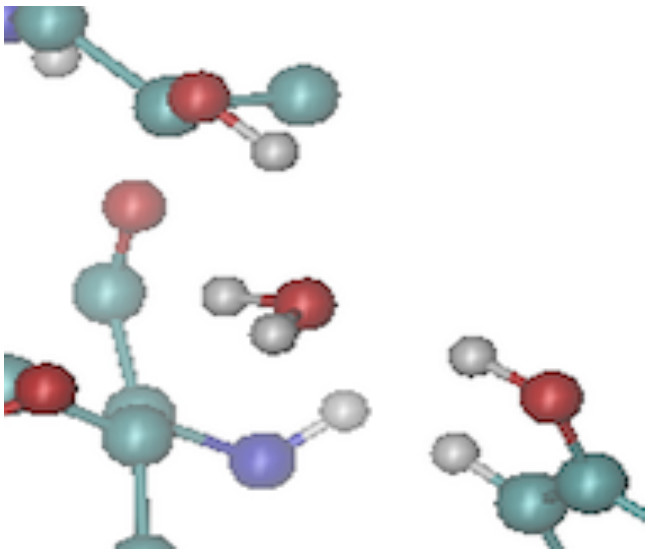


Figure-45: The buried crystallographic water molecule, cryst.112H_2O in the cavity of 52ASN-OD1, 67TYR-OH, and 78THR-OG1 in hh-CytC as defined in X-ray crystal structure [3].

In water solvent, cryst.112H_2O in our one analyzed simulation ensemble has resided in this centrally located conserved cavity for the whole 100 nanoseconds simulation time within H-bonding distance, closer to the side-chain aromatic hydroxyl group of 67TYR (Figure-46). But its distance was constantly oscillating in 2-6 Å in range with sidechain amide group of 52ASN and hydroxyl group of 78THR. This cryst.112H_2O looks in high electrostatic pressure where its mobility should be controlled by local conformation of protein or vice versa. Since this site is made up of all the highly conserved residues of cytochromes, viz. 52ASN, 67TYR,

78THR, the cryst.112H_2O should be more important functionally [8]. In all three simulation ensembles in pure water solvent (Figure-47), we did observe the cryst.112H_2O resided in this conserved site within hh-CytC during whole 100 ns simulation.

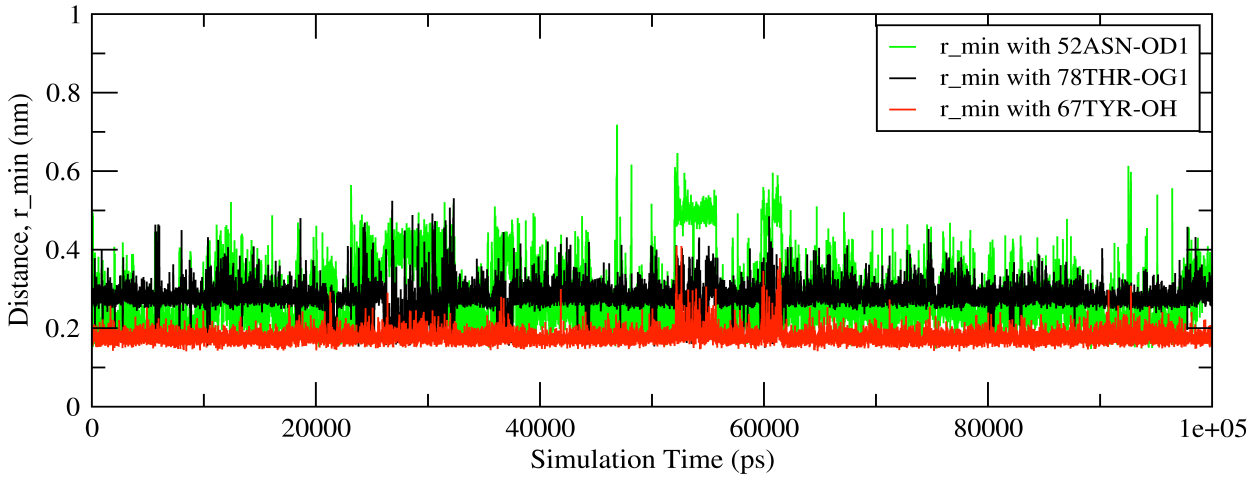


Figure-46: Distance between crystallographic water, cryst.112H_2O and its H-bonding partners inside hh-CytC in pure water solvent in one simulation ensemble at 298.15 K and 1atm.

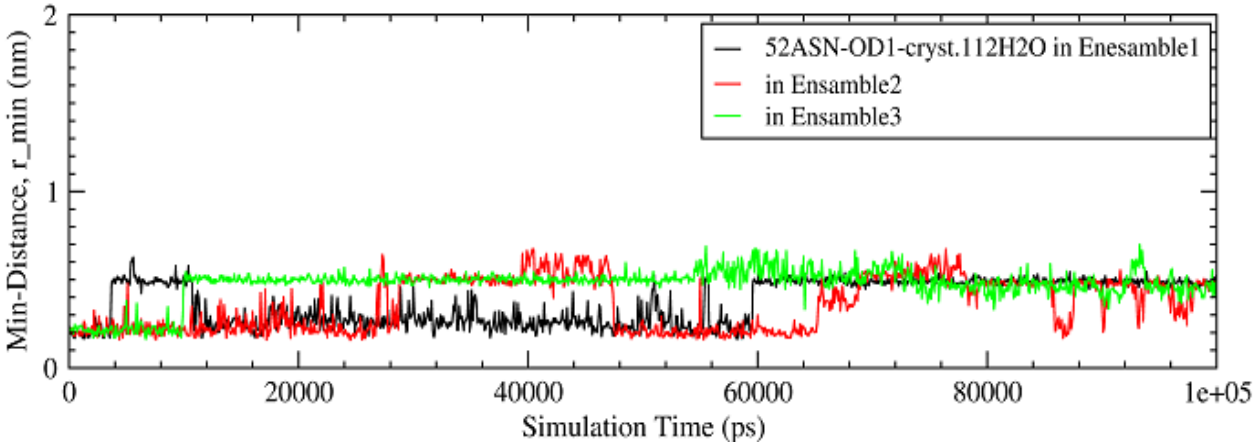


Figure-47: Distance between crystallographic water, cryst.112H_2O and one of its H-bonding partners (52ASN-OD1) inside hh-CytC in three different ensembles of pure water solvent at 298.15 K and 1atm.

In the other solvents, similar results were observed. In 20% aqueous MeOH solvent, the distances between cryst.H_2O and its H-bonding partners were highly fluctuating in the first 50 ns (Figure-48), from which we may predict that in this composition of binary solvent, hh-CytC may have followed the more unstable conformational transitions in this simulation and protein has been taking long time to achieve equilibrated conformation internally. After 50 ns, the cryst.112H_2O was has shifted toward 4 Å from 2 Å indicating that either the volume of the cavity made of these protein residue has increased where the buried water's position fluctuates between H-bonding partners, or cryst.112H_2O is slightly displaced from the conserved site due to conformational changes in protein. As in pure water, the cryst.112H_2O remained in the conserved site during whole simulation time in all three simulation ensembles as shown in Figure-49.

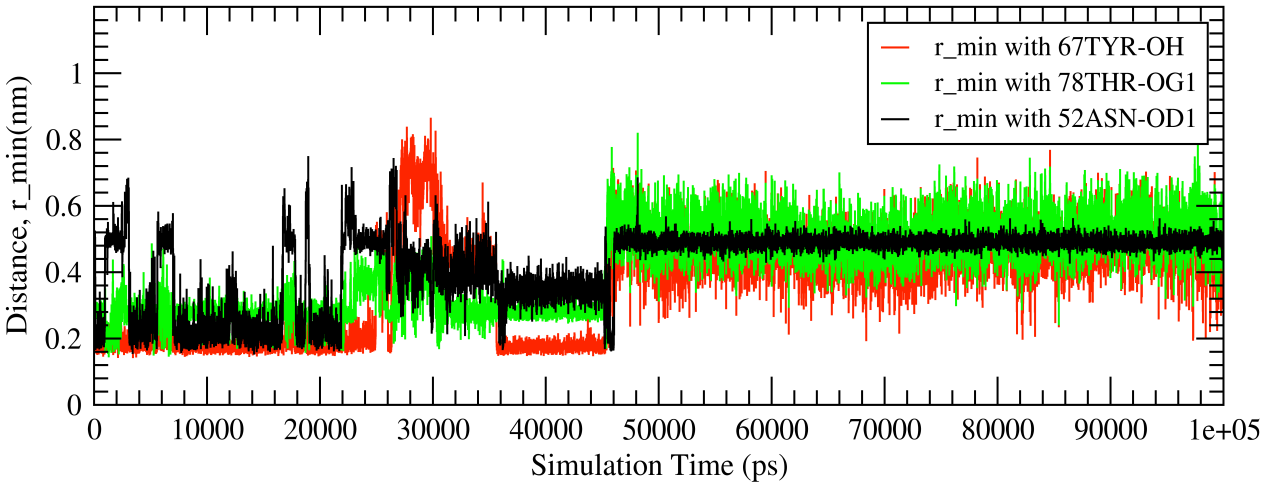


Figure-48: Distance between crystallographic water, cryst.112H_2O and its H-bonding partners in hh-CytC in 20% aqueous MeOH solvents in one analyzed ensemble at 298.15 K and 1atm pressure.

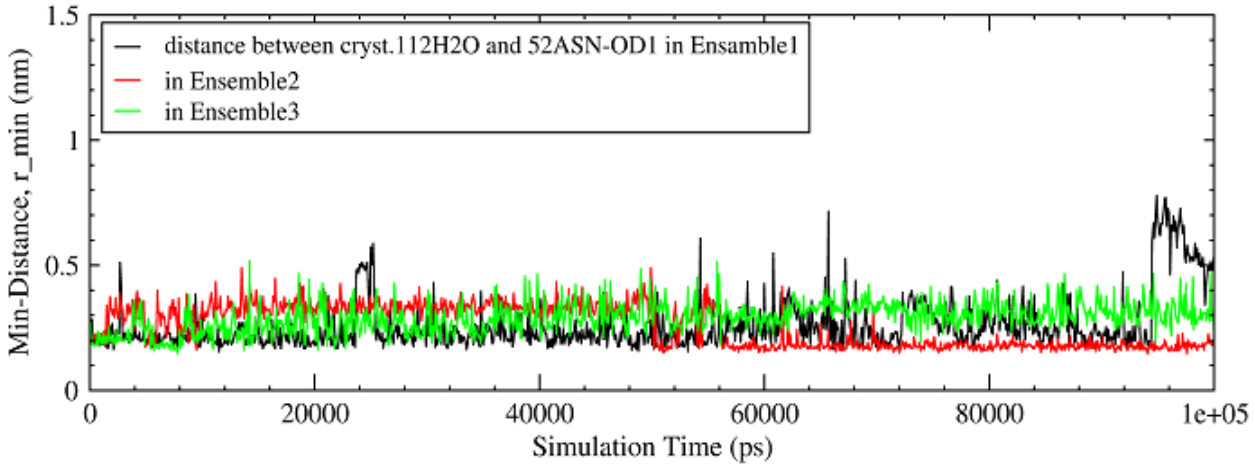


Figure-49: Distance between crystallographic water, cryst.112H_2O and one of its H-bonding partners (52ASN-OD1) in hh-CytC in three different simulation ensembles of 20% aqueous MeOH solvent at 298.15 K and 1atm pressure.

In 40% aqueous MeOH solvent (Figure-50), the cryst.112H_2O is displaced from its conserved site after 42 ns and has escaped from the cavity after 50 ns (d > 10 Å) in our analyzed ensemble. It is replaced by another solvent water molecule within 800 - 1200 picoseconds in our analyzed ensemble, as shown in Figure-51, even though this conserved site is located centrally inside protein and this water molecule resides in this cavity for the remaining 45 ns of this simulation (Figure-52). Figure-53 reveals that the residence time of cryst.112H_2O in this conserved site is different in different simulation ensembles, which may be an outcome of different mechanisms of conformational transition followed by hh-CytC in different simulation ensembles. We did not observe more than one water molecule in this site and no MeOH molecules even replaced water molecules in our three ensembles of 100ns simulations. This behavior of hh-CytC suggests that the protein has a more open structure in water-methanol binary mixture, which lets water or solvent molecules enter the interior of the protein. Insertion of solvent molecules inside the protein is the primary step of protein

denaturation [4, 11]; how long it takes in binary mixture of different water-methanol compositions may help us to judge the efficiency of protein activity that might be another part of our research.

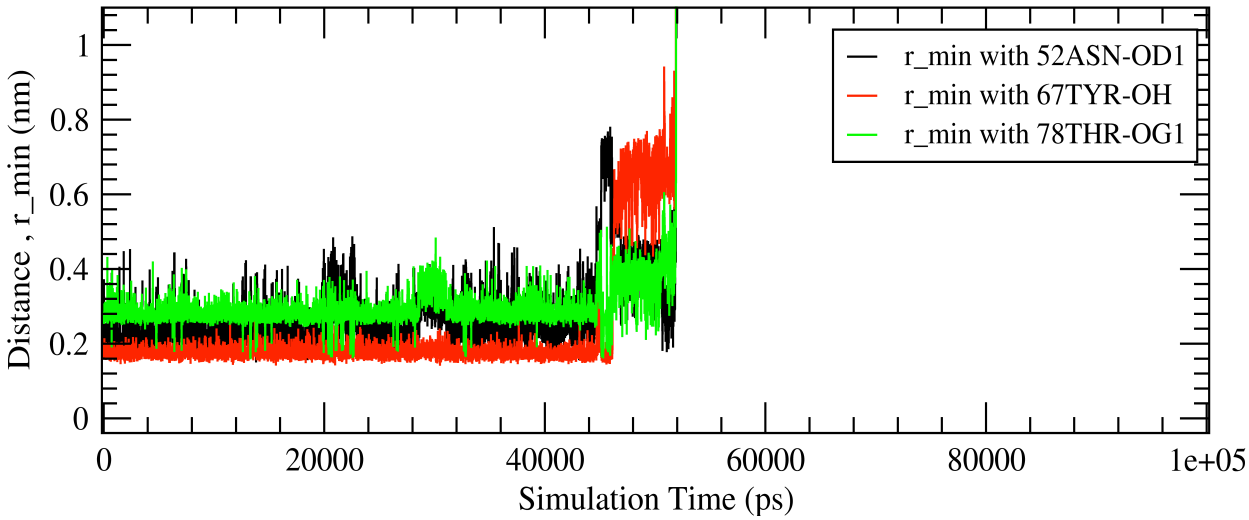


Figure-50: Distance between crystallographic water, cryst.112H_2O and its H-bonding partners in hh-CytC in 40% aqueous MeOH solvent in one analyzed ensemble at 298.15 K and 1 atm.

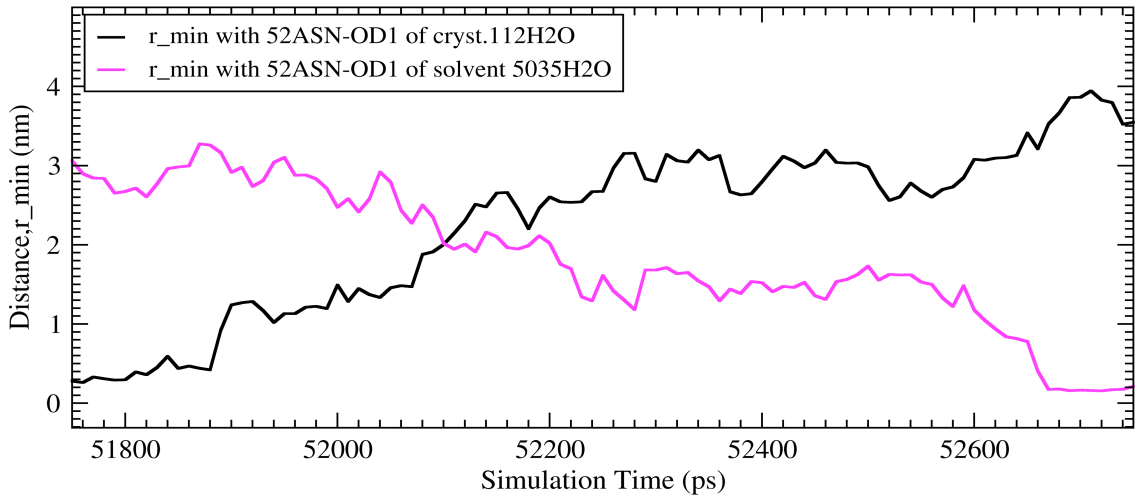


Figure-51: Illustration of an exchange event between cryst.112H_2O and solvent 5035H_2O in conserved site of 52ASN-OD1, 67TYR-OH and 78THR-OG1 inside hh-CytC in 40% aqueous MeOH solvent in one ensemble at 298.15 K and 1 atm.

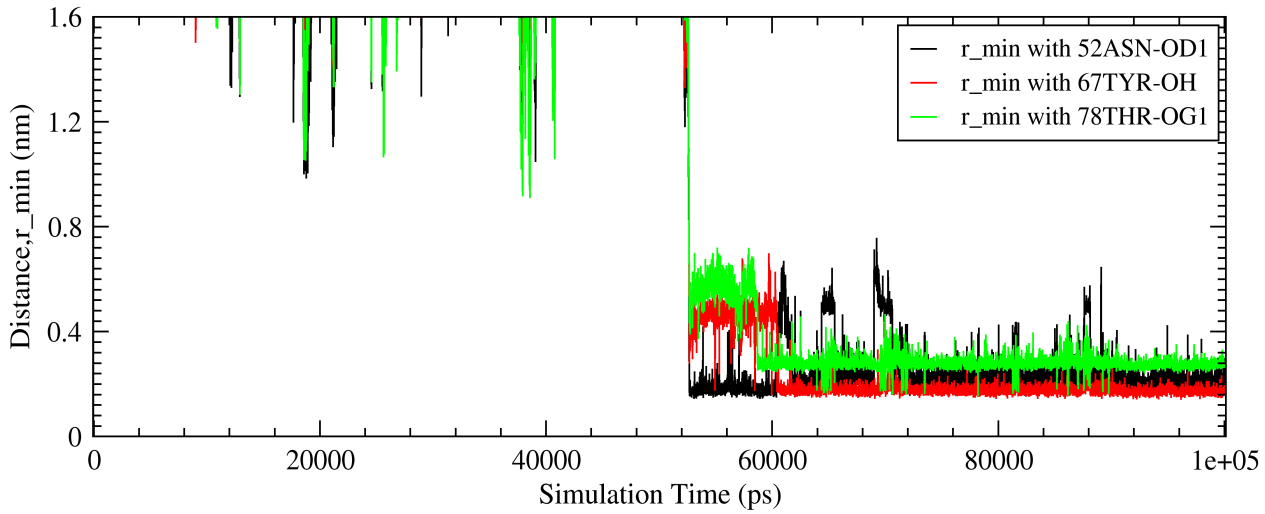


Figure-52: Distance between a solvent water molecule 5035H_2O that replaces cryst.112H_2O from the centrally located conserved site and its H-bonding partners inside hh-CytC in 40% aqueous MeOH solvent in one analyzed ensemble at 298.15 K and 1atm.

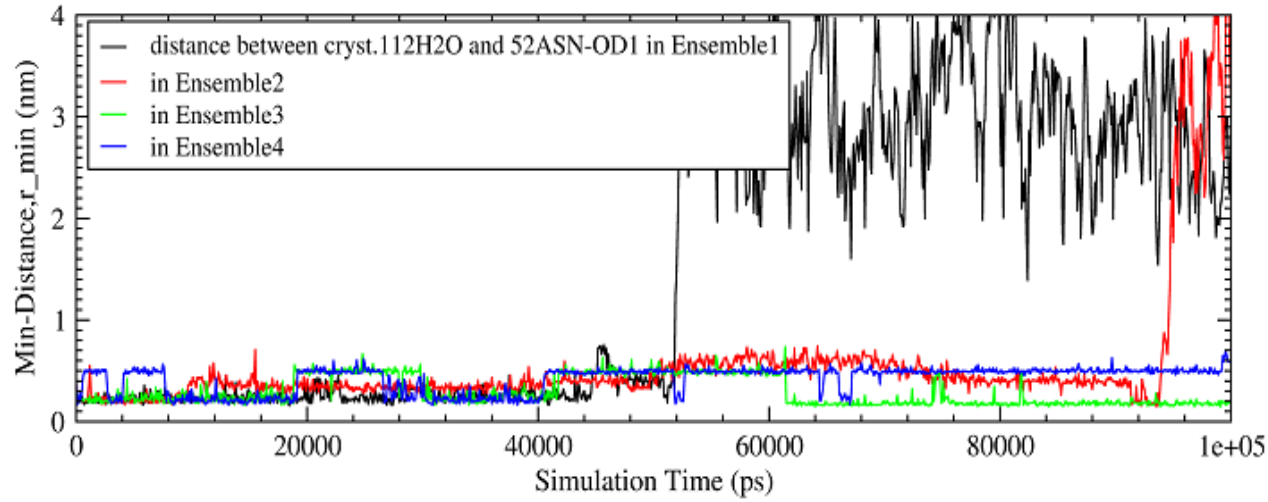


Figure-53: Distance between crystallographic water, cryst.112H_2O and one of its H-bonding partners (52ASN-OD1) of hh-CytC in four different simulation ensembles in 40% aqueous MeOH solvent at 298.15 K and 1atm.

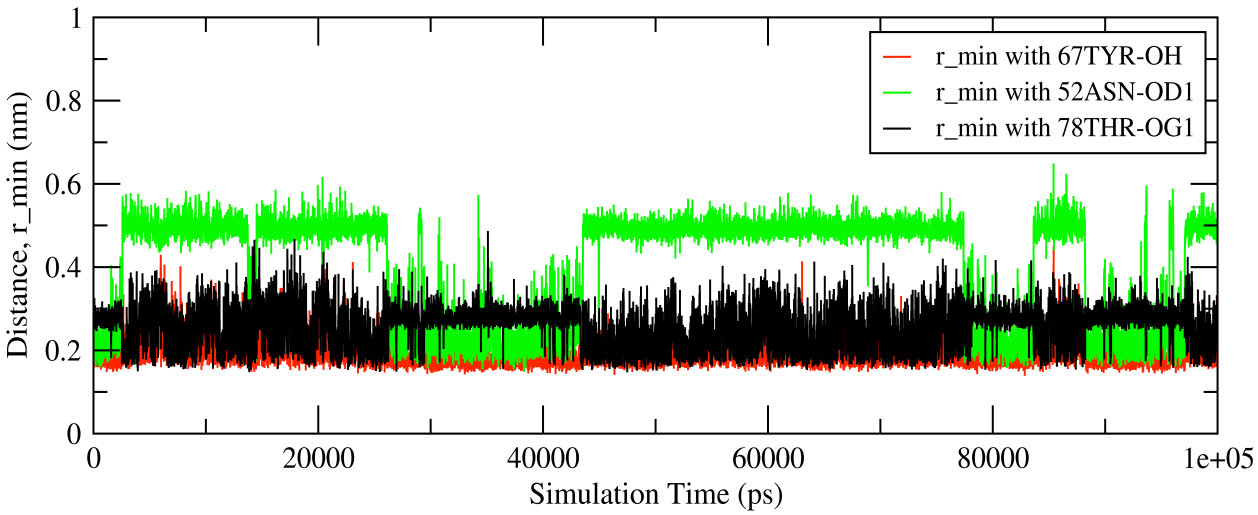


Figure-54: Distance between crystallographic water, cryst.112H_2O and its H-bonding partners in hh-CytC in pure methanol in one analyzed ensemble at 298.15 K and 1 atm.

In pure methanol solvent, cryst.112H_2O resides in the interior of protein during all three simulations ensembles for the whole 100 ns simulation as shown in Figure-54 & 55. The may be which may be because the comparatively large MeOH molecule cannot replace this buried water molecule even though conformation of protein swells slightly in pure methanol. The distances from cryst.112H_2O to its H-bonding partners inside protein were highly fluctuating which may be due to increased internal flexibility of protein as indicated by Lindemann's disorder index for backbone of hh-CytC. These results clearly indicate that the outer solvent environment has influenced the internal structure of protein.

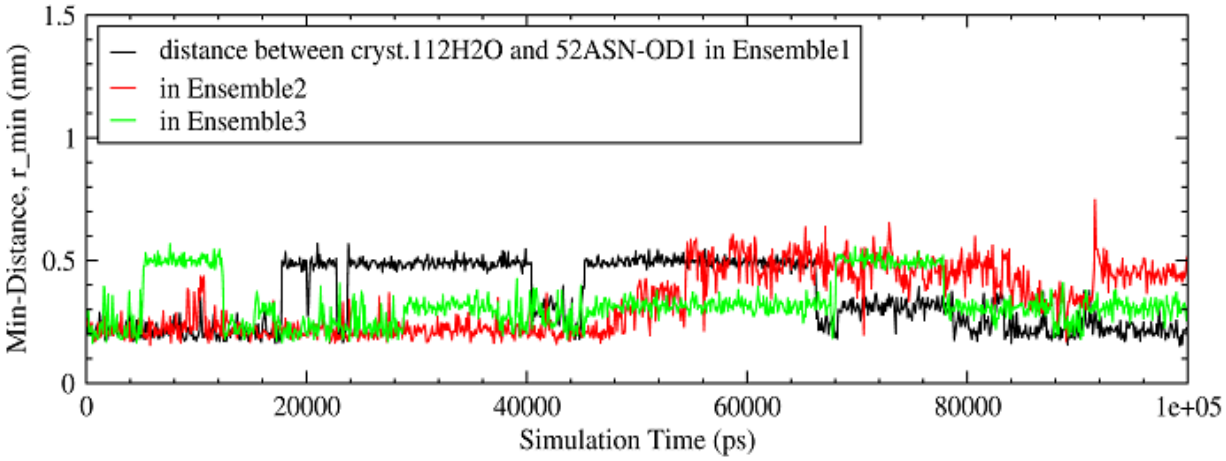


Figure-55: Distance between crystallographic water, cryst.112H_2O and one of its H-bonding partners (52ASN-OD1) of hh-CytC in three different simulation ensembles in pure methanol solvent at 298.15 K and 1atm.

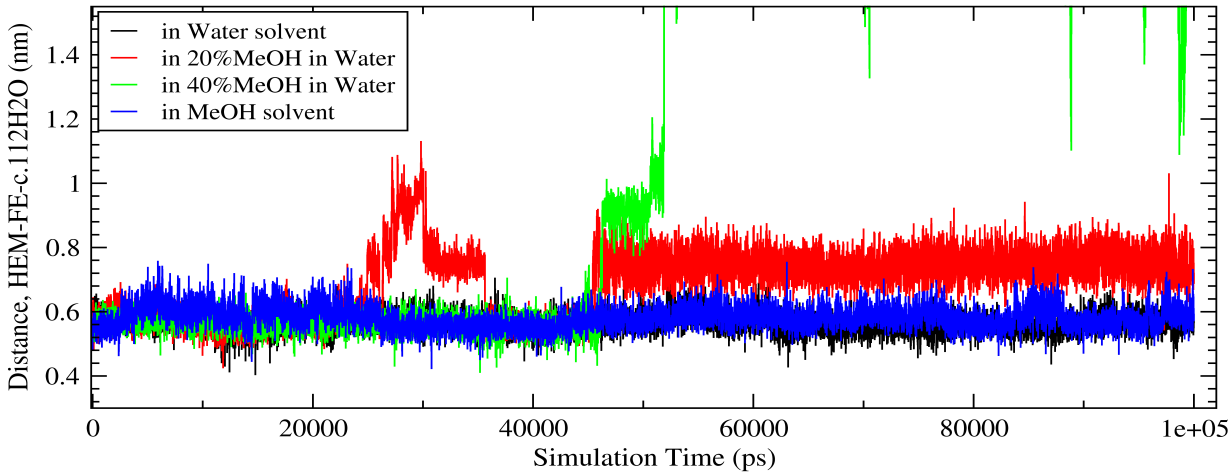


Figure-56: Distance between Fe (Iron in Heme group) and cryst.112H_2O in different solvents in one analyzed simulation ensembles at 298.15 K and 1atm.

Since GROMACS uses crystal structure of hh-CytC (*1HRC.pdb*) as the reduced form of hh-CytC with -2 charge in heme group (+2 charge of Fe) and total +7 charge including terminal charges, and the literature has mentioned that position of cryst.112H_2O from heme-Fe depends

on the oxidation state of hh-CytC [3, 6, 18, 33], we also observed the variation of distance between cryst.H_2O and Heme-Fe (Figure-56) even though we did not include any redox partners of hh-CytC in our simulation. In water and in methanol, its distance is 5.91 ± 0.65 Å and 6.20 ± 0.71 Å respectively in our analyzed ensemble. But in mix-solvents, it is about 7 Å when it resides in this conserved cavity. This fluctuation in the distance cryst.112H_2O and HEM-FE in our simulation should be simply the outcome of structural fluctuation of protein rather than depending on oxidation state since oxidation state remains constant in our work. So, these observations of long residence of cryst.112H_2O inside hh-CytC even in water-methanol binary mixture and pure methanol in our 100 ns simulation suggest that cryst.112H_2O should definitely have an important role in the mechanism of conformational transitions of hh-CytC and its function.

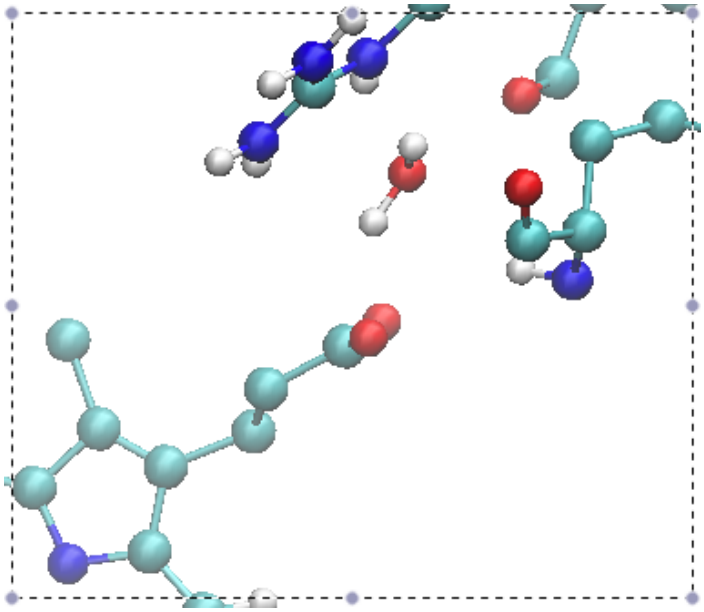


Figure-57: The buried crystallographic water molecule, cryst.125H_2O in conserved site of Heme propionate, 42GLN-N(backbone), 39LYS-O(backbone) and 38ARG-guanidino in hh-CytC as defined in X-ray crystal structure.

Likewise, we have also studied the dynamics of another crystallographic water molecule, cryst.125H_2O, which is buried internally in a conserved site near heme-propionate-A of hh-CytC as shown in Figure-57. This conserved site of water was found to be more vulnerable to the external solvent environment compared to previous centrally located conserved site. Figures- 58 to 67 provide the distance measured between cryst.125H_2O and its H-bonding partners in the conserved site inside the protein, viz. Heme propionate, 42GLN-N(backbone), 39LYS-O(backbone) and 38ARG-guanidino. The residence time of cryst.125H_2O in this conserved site is found to be different in different solvent compositions.

In pure water solvent in (Figure-58), it resided about 20 ns before it escaped to solvent in one analyzed ensemble. This site remained empty for about 6 ns, but about 5-7 solvent molecules were found 6-7.5 Å away from this conserved site where these solvent molecules might have been competing to occupy in this site. After 25 ns of simulation, a solvent water molecule succeeded to occupy this site in 6 ns as depicted in Figure-59 and resided till the end of 100 ns simulation (Figure-60). This phenomenon demonstrates that for these conserved sites, water-protein interactions are more favorable than the formation of intra-protein hydrogen bonds, which might be a structural or conformational need to adopt water molecules in these sites in the mechanism of protein folding. Figure-61 indicates that cryst.125H_2O was easily replaced by a solvent H_2O molecule in early stage of simulation in all three ensembles, indicating that this conserved site is easily accessible to solvent. But the exchange time was found to be longer than the centrally buried conserved site (occupied by cryst.112H_2O), indicating that solvent molecules spent more time competing to reach this conserved site.

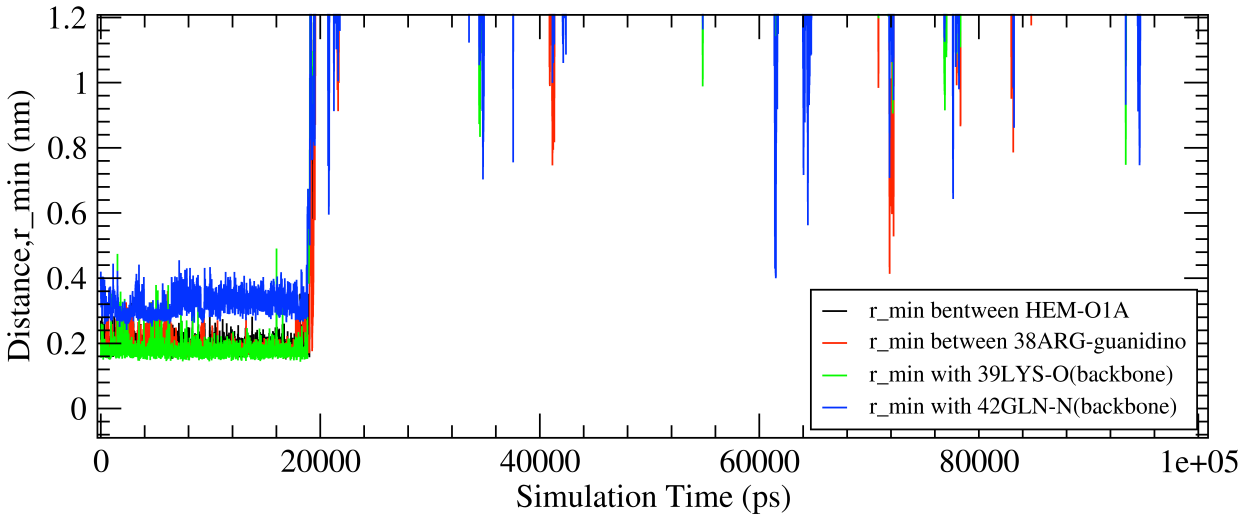


Figure-58: Distance between crystallographic water, 125H_2O and its H-bonding partners in hh-CytC in the pure water solvent in one analyzed ensemble at 298.15 K and 1atm.

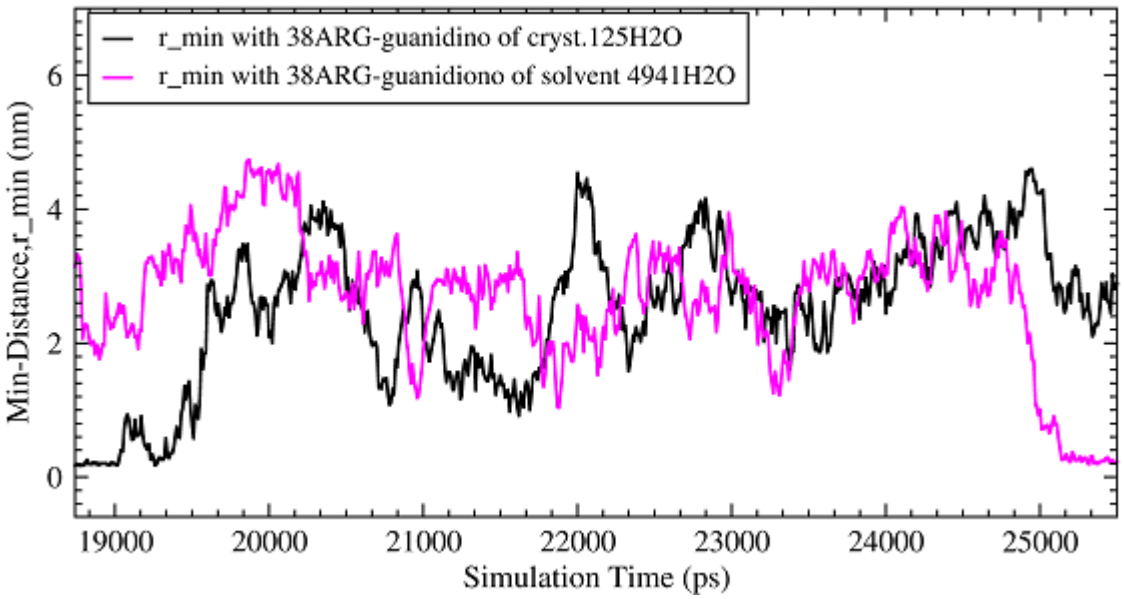


Figure-59: Distance fluctuation in the mechanism of exchange between one solvent molecule 4941H_2O and cryst.125H_2O in the conserved site (Heme propionate, 42GLN-N (backbone), 39LYS-O (backbone) and 38ARG-guanidino) within hh-CytC in pure water solvent in one simulation ensembles at 298.15 K and 1 atm pressure.

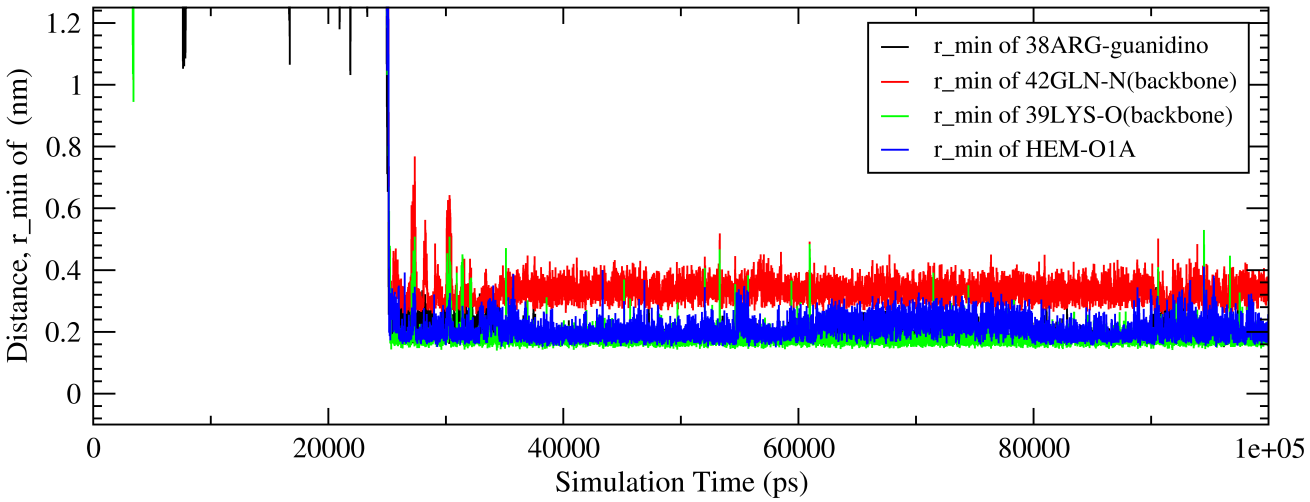


Figure-60: Distance between solvent water molecule (one solvent water molecule, 4941H_2O that exchanged with cryst.125H_2O) in one simulation ensemble and its H-bonding partners in hh-CytC in water solvent at 298.15 K and 1atm.

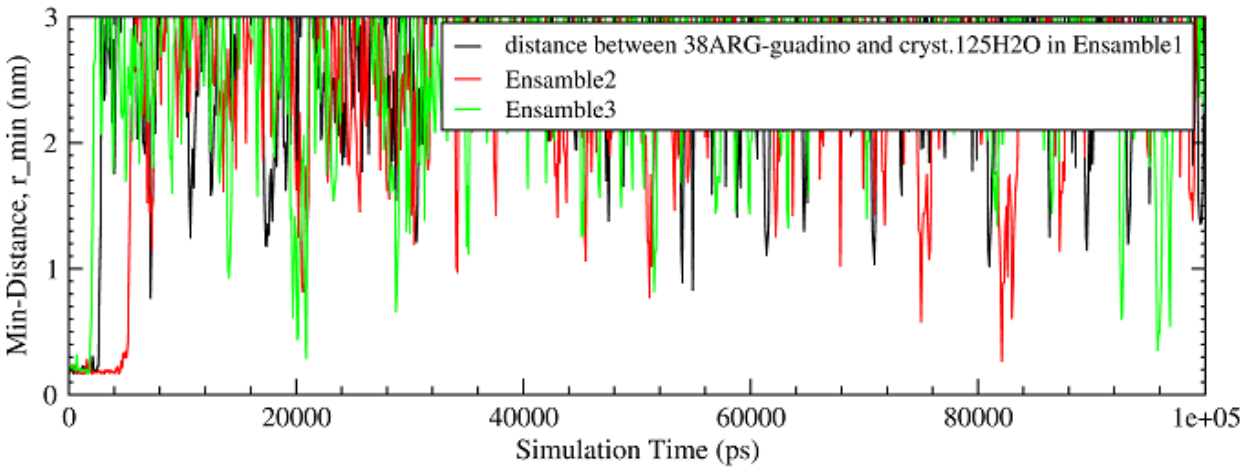


Figure-61: Distance between crystallographic water, 125H_2O and one of its H-bonding partners (38ARG-gunidino) in hh-CytC in pure water solvent in three simulation ensembles at 298.15 K and 1atm.

Similarly, in binary solvents, only solvent water molecules were found to replace cryst.125H_2O in this site in all ensembles of 100 ns simulation. Even though MeOH molecules

have a good probability of H-bonding, we did not find any MeOH molecules in 7 Å distance near to this conserved site, and here we may bluntly say that this preference of H_2O molecules for this conserved site should be because of the size of solvent molecules where MeOH is bigger than H_2O to fit to this conserved site, and hence H_2O can compete with MeOH with higher hydrogen bonding probability. In 20% aqueous MeOH solvent, cryst.125H_2O resides about 45 ns in the analyzed ensemble (Figure-62), but in all three ensembles, the residence time of cryst.125H_2O in this conserved site was found different in the same condition. When we tracked the solvent molecules, we observed that four solvent H_2O molecules were closer at distance of 5 Å all the simulation time (Figure-63), and four H_2O molecules and two MeOH molecules were at distance 7.5 Å away from this conserved site.

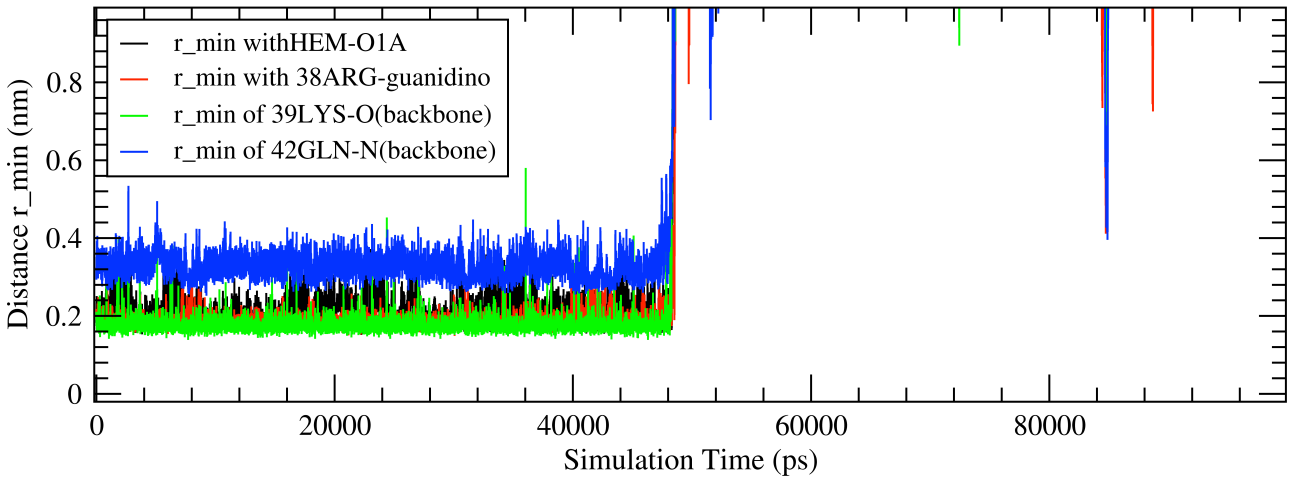


Figure-62: Distance between crystallographic water, 125H_2O and its H-bonding partners in hh-CytC in one analyzed ensemble of 20% aqueous MeOH solvent at 298.15 K and 1atm.

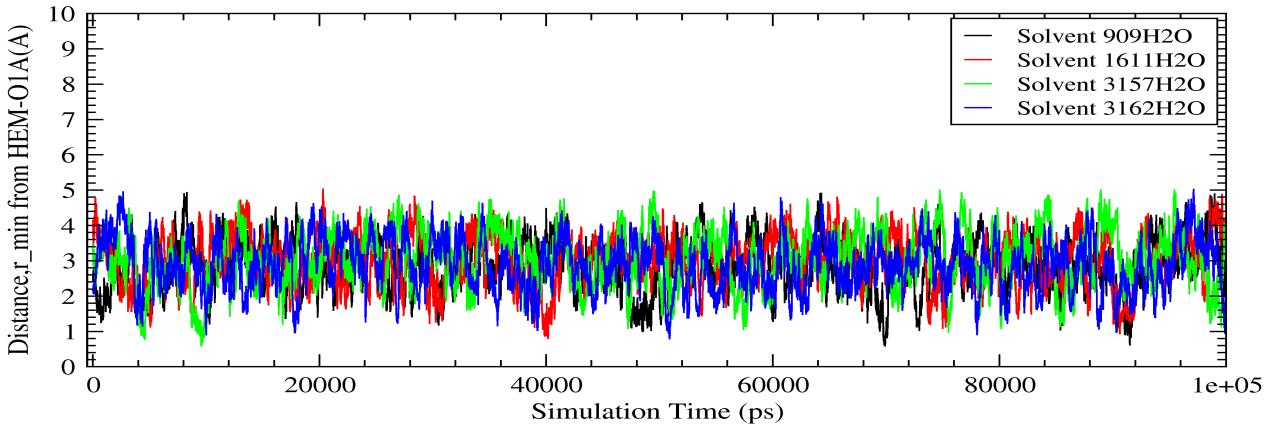


Figure-63: Distance between four solvent water molecules (that replace 125H_2O from conserved site inside protein) and HEM-propionate-A of hh-CytC in 20% aqueous MeOH solvent in one analyzed ensemble at 298.15 K and 1atm.

In 40% aqueous MeOH solvent (Figure-64), we did not observe cryst.125H_2O replaced by any solvent molecules in one analyzed ensemble but we did succeed to track three solvent H_2O and one MeOH molecules at 5 Å closer to this conserved site during simulation. In other ensembles of 40% aqueous MeOH solvent (Figure-65), we did observe that only the solvent H_2O has replaced cryst.125H_2O from its bound site in different time.

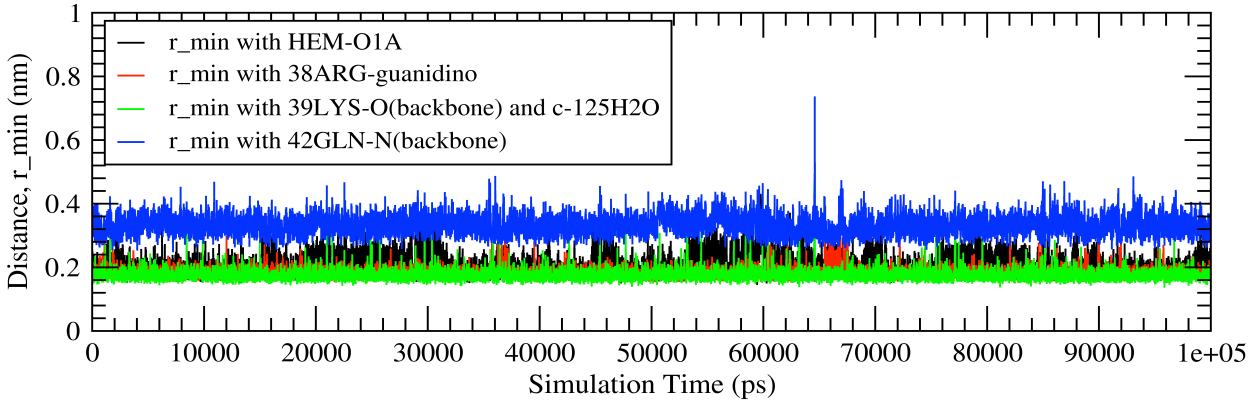


Figure-64: Distance between crystallographic water, 125H_2O and its H-bonding partners in hh-CytC in 40% aqueous MeOH solvent in one ensemble at 298.15 K and 1atm.

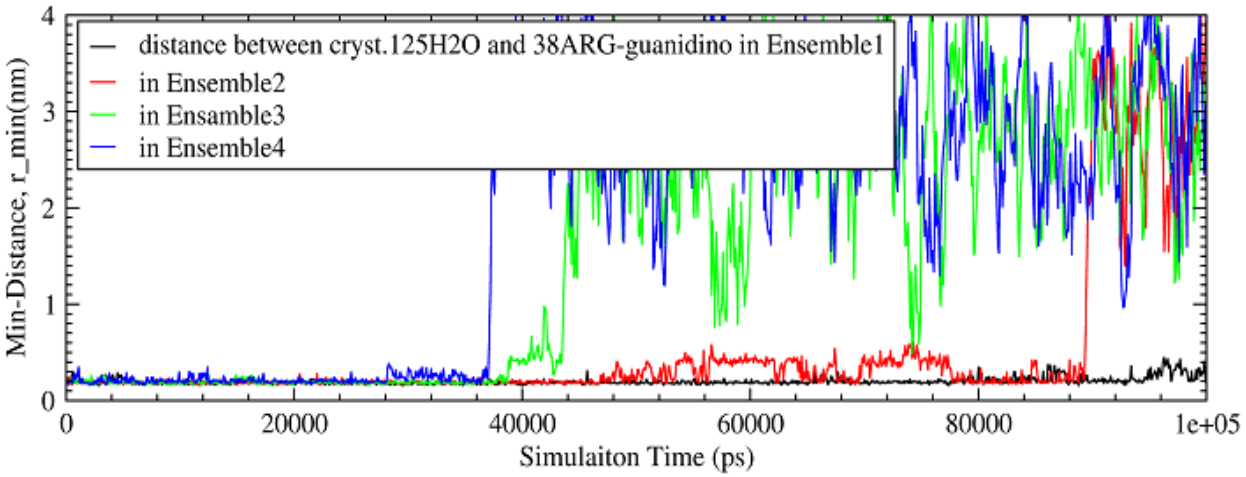


Figure-65: Distance between crystallographic water, 125H_2O and one of its H-bonding partners in hh-CytC (38ARG-guanidino) in 40% aqueous MeOH solvent in four simulation ensembles at 298.15 K and 1 atm pressure.

In pure methanol solvent, 1 to 5 methanol molecules occupied the space within 7.5 - 10 Å from this site from starting of the simulation and the cryst.125H_2O slowly comes out of this site within 5 ns and remains within 10 Å away from this site most of the time (Figure-66). Along with MeOH molecules, cryst.125H_2O try to get this site in next 30 ns, but after 35 ns simulation one MeOH molecule got closer (2-8 Å) to this site (Figure-67). It is still unclear why cryst.125H_2O was stripped off from this site by itself in pure methanol solvent without replacement by a MeOH molecule completely. Our understanding is that excess internal structural flexibility of protein at the expense of electrostatic or H-bonding instability may contribute the cryst.125H_2O moving away from the conserved site. Since, in literature, cryst.125H_2O is supposed to mediate the charged interaction between heme propionate and 38ARG [3], it appears the residence of water molecule in that conserved site for a nano-second time scale is functionally justified even in water methanol mix solvents. But, if this site is

replaced by MeOH, it is important to understand how much or how long the hh-CytC is functionally viable in redox reaction.

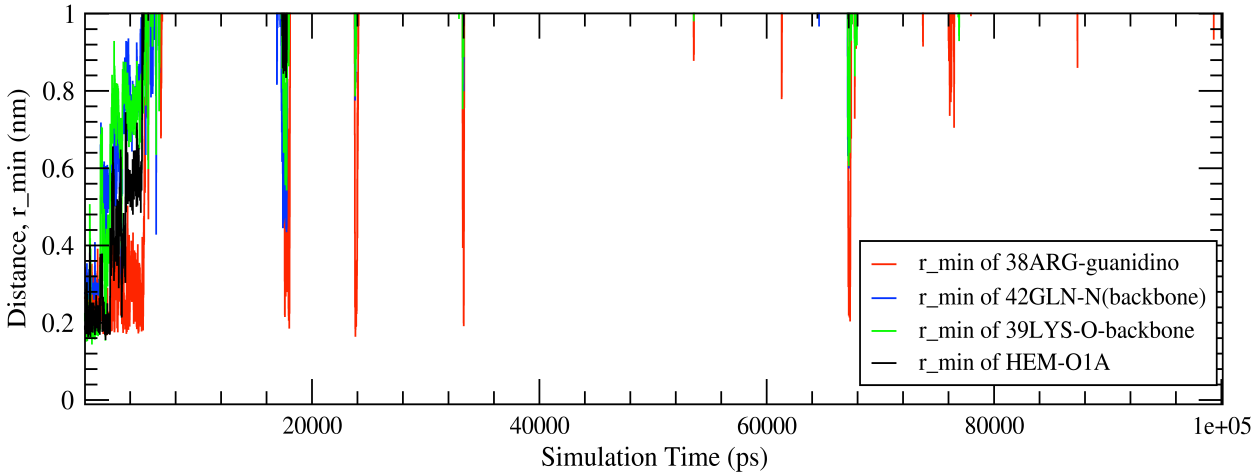


Figure-66: Distance between crystallographic water, 125H_2O and its H-bonding partners in hh-CytC in one analyzed ensemble in pure methanol solvent at 298.15 K and 1 atm pressure.

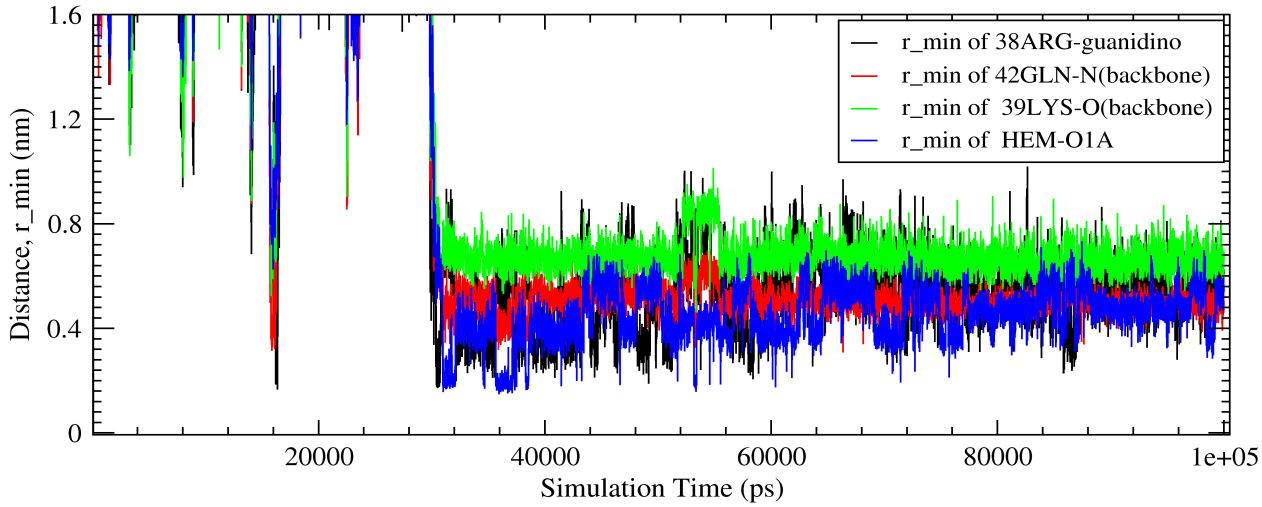


Figure-67: Distance between a solvent MeOH molecule, 394MeOH, which replace the cryst.125H_2O from its conserved sites and its H-bonding partners in hh-CytC in Methanol at 298.15 K and 1 atm pressure.

So, concisely, our understanding is that these crystallographic water molecules are not permanently conserved in the position they occupied, but rather these positions are highly conserved for a water molecule for hydrogen bond which may be due to conformational need while they may have included as residual water internally during early protein folding stages. The water molecules in these conserved positions have long nanoscales residence time with hydrogen bond lifetimes of 100 to 1000's of picoseconds. The residence time of these water molecules or the time to replace these water molecules by solvent molecules depends on the paths or mechanisms of conformational transitions followed by hh-CytC in simulation. The internal flexibility of the protein and mobility of these buried water molecules are controlled mutually, where these water molecules may act as bearings in wheels or buffer of internal motion to maintain flexibility of protein as well as pathway of electron transfer process.

CHAPTER IV

CONCLUSIONS

1. Summary of Present Work

We studied horse heart Cytochrome C (hh-CytC) by molecular dynamic simulations in our laboratory using GROMACS software and its most recent force field, *gromos53a6*, in four solvents namely, water, methanol and two binary mixtures, 20 percent methanol in water and 40 percent methanol in water at 298.15 Kelvin temperature and 1 atmospheric pressure at constant NPT condition. We focused our research on four aspects of structural properties of horse heart Cytochrome C and all the solvents compositions themselves in long 100 ns simulations. First, we studied the macroscopic properties of all solvents and compared with experimental results and literature values for the reproducibility of our experiments. Second, we computed and analyzed structural parameters of horse-heart cytochrome C in different solvents from molecular dynamics simulation vis a vis available X-ray crystal structure to understand the effect of solvent environment on protein structure. We tried to evaluate and characterize the effect. Third, we studied solvent dynamics and the dynamics of horse heart cytochrome C in different solvents, and the solvent–protein hydrogen bonding properties in interface. Fourth, we studied the significance of crystallographic water in relation to structure and functional properties of horse heart cytochrome C. In particular, we tracked the two crystallographic water molecules buried in conserved sites inside the protein. We computed the positional and hydrogen bonding properties of these crystallographic water molecules to advance in further understanding of their role in protein function.

We found that our methodology was quite effective for computing reproducibly the macroscopic properties of the solvents; using SPC/E water model and OPLS-UA model for methanol. The solvent properties which we calculated, such as dielectric constant, dipole moment, solvent density and viscosity, diffusion coefficient and hydrogen bond lifetime were quite reproducible and comparable to experimental values in our simulation conditions. The C-α backbone RMSD of horse heart Cytochrome C is higher in water-methanol binary solvents and in pure methanol relative to the RMSD in pure water solvent, even though we observed the same pattern of C-α backbone RMSF of protein residue in all solvent compositions with a slight incremental increase in presence of methanol. The heme prosthetic group buried inside the protein was found to be quite stable irrespective of solvent properties. From the estimation of Lindemann's disorder index, we found that the protein is more solid-like in the core or backbone and more liquid-like at the surface and in its overall structure. Compared to the protein in pure water, protein has more flexibility at both the surface and internal core in mix-solvent. In contrast, in pure methanol, it showed more solid-like structure in its internal structure and more liquid-like on the surface. The number of internal hydrogen bonds in horse heart Cytochrome C increases with increasing amount of methanol solvent. We calculated the internal hydrogen bond lifetime, which is hundreds of picoseconds in range and is different in different solvents with only very slight changes in H-D-A (Hydeogen-Donor-Acceptor) hydrogen bond angles and D-A (Donor-Acceptor) distances. The slight increase in the radius of gyration of horse heart cytochrome C with high fluctuations of hydrophobic solvent accessible surface area (Hydrophobic SASA) indicates the role of hydrophobic interactions of the methyl group of CH_3OH in the conformational transition of the protein.

The number of hydrogen bonds between horse heart Cytochrome C and solvent molecules decreases with increasing methanol percentage in solvents, which should be because of excluded volume of methyl group near the protein surface resulting in fewer number of molecules in solvation layer around the protein surface. The hydrogen bond lifetime between solvent molecules and protein was found to be 2-3 times higher than solvent-solvent hydrogen bond lifetimes, indicating the affinity of solvent molecules to exist in the solvation layer around protein. We hardly conclude the diffusion properties of hh-CytC from our single molecule simulation experiment.

Most of the crystallographic water molecules diffuse into bulk solvent within the first few nanoseconds of simulation time. Few surface crystallographic water molecules reside in their position on a picoseconds time scale. Notably, two crystallographic water molecules, which were buried in conserved sites inside the protein, resided with a time scale of nanoseconds, having hydrogen bond lifetimes of 100's of picoseconds. The water molecules in these sites were replaced by solvent water molecules in our simulation, which implies that these conserved sites are more important to keep the water molecules that are structurally and functionally important. The residence time of these buried crystallographic water molecules in the conserved sites might depend on the path or mechanism of conformational transition of hh-CytC during simulation and, consequently, the accessibility of solvent in these conserved sites. In binary mixture, only the solvent H_2O molecules, but not MeOH molecules, were found successful to replace these buried crystallographic water molecules, which may be the result of size effect along with probability of hydrogen bonding. The MeOH molecules were observed near to these conserved sites but rarely close enough for hydrogen bonding.

The simulation of hh-CytC without crystallographic water in the same condition increases the C-α backbone RMSD and RMSF of protein indicating its higher conformational flexibility even though this change was not big enough to show protein unfolding behavior in our 100 ns simulation. The effect was more prominent in the water and water-methanol binary solvents having higher water in compositions. Moreover, based on the data of Lindemann's disorder index of hh-CytC, the protein showed more liquid-like internal core along with overall higher molten structure of protein in absence of crystallographic water molecules. These buried crystallographic water molecules may also involve in buffering action in protein's internal dynamics in conformational transitions.

2. Future Outlook

The role of water molecules in maintaining certain protein structure or conformation is vital for specificity in protein function. Hydration of protein to certain extent is essential for its function to harness non-aqueous enzymology. Study of these water molecules, which are intimately associated with a protein, is essential in understanding protein's folding mechanism and their malleability.

The crystal structure of a protein obtained from X-ray and NMR methods has included some water molecules that are located in defined positions buried internally or at the protein surface of protein structure. These crystallographic water molecules have resided in certain conserved sites for nanoseconds of time and should have structural and functional value. As Cytochrome C is a cornerstone of biophysical research, we used the horse heart Cytochrome C as a model protein in water-methanol binary mixtures of different compositions to establish a new paradigm in understanding the role of solvent and crystallographic water molecules to normalize internal flexibility. The study of these crystallographic water molecules individually

using molecular dynamics simulation is very important to specify their role and has scope in further understanding of protein folding mechanisms.

Since solvent compositions influence the structure of protein, modification of flexibility of internal core of protein is one of the parameters to estimate the entropic contribution in the thermodynamics of protein catalysis. We observed a significant effect of buried crystallographic water molecules in variation of internal motion of hh-CytC, which might be a milestone in the study of solvent assisted protein function and to investigate substrate-protein compatible solvents.

A protein may become functionally inactive without any significant change in its overall secondary or tertiary structure. The characteristics of H-bond lifetime and salt-bridge interactions in a protein in different environments should be essential parameters to explain the nuances of protein activity in different solvents. In terms of hydrogen bonding energy, it would be interesting to perform these *in silico* experiments at different temperatures which give a way to calculate the thermodynamic parameters for optimum enzyme activity.

REFERENCES

REFERENCES

1. Salemme, F.R. Structure and function of cytochrome c. *Ann. Rev. Biochem.* **1977**, *46*, 299 – 329.

2. Battistuzzi, G.; Borsari, M.; Sola, M. Redox properties of cytochrome c. *Antioxid. Redox Sign*. **2004**, *3,* 278 – 291.

3. Bushnell, G.W.; Louie, G.V.; Brayer, D. G. High-resolution three-dimensional structure of horse heart cytochrome c. *J. Mol. Biol.* **1990**, *214*, 585 – 595.

4. Pulawski, W.; Filipec S.; Zwolinska, A.; Sebinski, A.; Krzysko, K.; Garduno-Juarez, R.; Viswanathan, S.; Ranugopalakrishnan V.; Low-temperature molecular dynamics simulation of horse heart cytochrome c and comparison with inelastic neutron scattering data. *Eur. Biophys. J.* **2013**, *42*, 291 – 300.

5. Rasnik, I.; Sharp K. A.; Fee; J. A.; Vanderkooi, J. M. Spectral analysis of cytochrome c: effect of heme conformation, axial ligand, peripheral substituents, and local electric fields. *J. Phys. Chem. B,* **2001**, 105, 282 – 286.

6. Banci, L.; Bertini, I.; Huber, J. G.; Spyroulias, G.A.; Turano, P. Solution structure of reduced horse heart cytochrome c. *J. Biol. Inorg. Chem.* **1999**, *4*, 21 – 31.

7. De Biase, P. M.; Paggi, D.A.; Doctorovich F.; Hildebrandt, P.; Estrin, D. A.; Murgida, D. H.; Marti, M. A. Molecular basis for the electric field modulation of cytochrome c structure and function. *J. Am. Chem. Soc.* **2009**, *131*, 16248 – 16256.

8. Zaidi, S.; Hassan, Md. I.; Islam, A.; Ahmad, F. The role of key residues in structure, function, and stability of cytochrome - c. *Cell. Mol. Life Sci.* **2014**, *71*, 229 – 255.

9. Qi, P.X.; Urbauer, J.L.; Fuentes, E. J.; Leopold, M. F.; Wand, A. J. Structural water in oxidized and reduced horse heart cytochrome c. *Nat. Struct. Biol.* **1994**, *1*, 378 – 382.

10. Mark,P.; Nilsson, L.; Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A*, **2001**, *105*, 9954 – 9960.

11. Cruz, A.; Ramirez,E.; Santana, A.; Barletta, G.; Lopez, G.E. Molecular dynamics study of subtilisin Carlsberg in aqueous and nonaqueous solvents. *Molecular Simulation.* **2009**, *35*, 205 – 212.

12. Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. An extended dynamical hydration shell around proteins. *PANS.* **2007**, *104*, 20749 – 20752.

13. Soares, C.M.; Teixeira, V. H.; Baptsta, A. M. Protein structure and dynamics in nonaqueous Solvents: Insights from molecular dynamics simulation studies. *Biophys. J.* **2003**, *84*, 1628 – 1641.

14. Wedberg, R.; Abildskov, J.; Peters, G. H. Protein dynamics in organic media at varying water activity studied by molecular dynamics simulation. *J. Phys. Chem. B,* **2012**, *116*, 2575 – 2585.

15. Wuthrich, K.; Otting, G. Studies of protein hydration in aqueous solution by direct NMR observation of individual protein - bound water molecules, *J. Am. Chem. Soc.* **1989**, *111*, 1871 – 1875.

16. Kriwacki, R. W.; Hill, R. B.; Flanagan, J. M.; Caradonna, J. P.; Prestegard, J. H.; New NMR methods for the characterization of bound waters in macromolecules. *J. Am. Chem. Soc.* **1993**, *115*, 8907 – 8911.

17. Humphrey, W.; Dalke, A; Schulten, K. "VMD – Visual Molecular Dynamics". *J. Molec. Graphics.* **1996**, *14*, 33 – 38. http://www.ks.uiuc.edu/Research/vmd.

18. van der Sopel, D.; Lindhhl, E.; Hess, B.; van Buuren, Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, T.M.; Feenstra, K. A.; vanDrunen R.; Berendsen, H.J.C. Gromacs User Manual 4.5.5-6. **2010**. www.gromacs.org.

19. Hess B, Bekker H, Berendsen HJC, Fraaije J. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18*, 1463 – 1472.

20. Berendsen, H. J. C.; Grigera, J. R.; Straatsm. T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269 – 6271.

21. Jorgensen, W. L. Optimized intermolecular potential functions for liquid alcohols. *J. Phys. Chem.* **1986**, *90*, 1276 – 1284.

22. Adamovic, I.; Gordon, M.S.; Methanol – Water Mixtures: A micro-solvation study using the effective fragment potential method. *J. Phys. Chem. A*, **2006**, *110*, 10267 – 10273.

23. Pascal, T. A.; Goddard III, W.A. Hydrophobic segregation, phase transitions and the anomalous thermodynamics of water/methanol mixtures. *J. Phys. Chem. B,* **2012**, *116*, 13905 – 13912.

24. Bortolotti, A. C.; Amadei, A.; Aschi, M.; Borsari, M.; Corni, S.; Sola, M.; Daidone, I.; The reversible opening of water channels in cytochrome - c modulates the heme iron reduction potential. *J. Am. Chem. Soc.* **2012**, *134*, 13670 – 13678.

25. Cramer, C. J. *Essentials of computational chemistry, theories and models*, 2nd ed.; John Wiley & Sons: San Francisco, 2004.

26. van Gunsteren, W. F.; Mark, A. E. Validation of molecular dynamics simulation. *J. Chem. Phys.* **1998**, *108*, 6109 – 6116.

27. Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press: New York, 1987.

28. Schlick, T. *Molecular modeling and simulation: An interdisciplinary guide*, 2nd ed.; Springer: New York, 2010.

29. Miller, J.; Molecular Dynamics. *ENCYCLOPEDIA OF LIFE SCIENCE*. Nature Publishing Group, **2001.** www.els.net.

30. van Gunsteren, W.F.; Berendsen, H. J. C. Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992 – 1023.

31. van Gunsteren, W.F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D.P.; Glattli, A.; Hunenberger, H. P.; Kastenholz, M.A.; Oostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F. A.; Yu, H.B. Biomolecular modeling: Goals, problems, perspectives. *Angew. Chem. Int. Ed.* **2006**, *45*, 4064 – 4092.

32. Prabhu, N. V.; Dalosto, D. S.; Sharp, K. A.; Wright, W. W.; Vanderkooi, J. M. Optical spectra of Fe (II) cytochrome c interpreted using molecular dynamics simulations and quantum mechanical calculations. *J. Phys. Chem. B,* **2002**, *106*, 5561 – 5571.

33. Banci, L.; Bertini, I.; Gray, H. B.; Luchinat, C.; Reddig, T.; Rosato, A.; Turano, P. Solution structure of oxidized horse heart cytochrome c. *Biochemistry*. **1997**, *36*, 9867 – 9877.

34. Raccatano, D. Computer simulations study of biomolecules in non-aqueous or cosolvent / water mixture solutions. *Current Protein and Peptide Science.* **2008**, *9*, 407 – 426.

35. Haynes, W. M.; *et al*, Eds. *CRC Handbook of Chemistry and Physics, 93rd Edition*; CRC press, Taylor & Francis Group: New York, 2012 – 2013.

36. Frenkel, D.; Smit, B. *Understanding Molecular Simulation, From Algorithm to Applications*; Academic Press: San Diego, 1996.

37. Van Vliet, K. J.; Walton, E. B. Equilibration of experimentally determined protein structures for molecular dynamics simulation. *Phys. Rev. E.* **2006**, *74*, 061901(8).

38. Palmer, B. J.; Transverse-current autocorrelation-function calculations of the shear viscosity for molecular liquids. *Phys. Rev. E,* **1994**, *49*, 359 – 366.

39. Wakeham, W.A.; Sokolov, M.; Kestin, J. Viscosity of liquid water in the range - 8 ^{0}C to 150 ^{0}C. *J. Phys. Chem. Ref. Data.* **1978**, *7*, 941 – 947.

40. Crippen M. G.; Maiorov V.N.; Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *J. Mol. Biol.* **1994**, *235*, 625 – 635.

41. Zhou, J.; Zheng, J.; Jiang, S.; Molecular simulation studies of the orientation and conformation of cytochrome c adsorbed on self-assembled monolayers. *J. Phys. Chem. B,* **2004**, *108*, 17418 – 17424.

42. Xu, D.; Tsai C.; Nussinov, R. Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Eng.* **1997**, *10*, 999 – 1012.

43. Bosshard H. R.; Marti, D. N.; Jelesarov, I.; Protein stabilization by salt bridges: Concepts, experimental approaches and clarification of some misunderstandings. *J. Mol. Recognit.* **2004**, *17*, 1 – 16.

44. Luzar, A. Resolving the hydrogen bond dynamics conundrum. *J. Chem. Phys.* **2000**, *113*, 10663 – 10675.

45. Spoel van der, D.; van Maaren, J. P.; Larsson, P.; Timneanu, N. Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. *J. Phys. Chem. B,* **2006**, *110*, 4393 – 4398.

46. Antipova, M. L.; Petrenko, V. E. Hydrogen bond lifetime for water in classic and quantum molecular dynamics. *Russ. J. Phys. Chem. A*, **2013**, *87*, 1170 – 1174.

47. Frembgen-Kesner, T.; Elcock, A. H.; Striking effects of hydrodynamic interactions on the simulated diffusion and folding of proteins. *J. Chem. Theory Comput.* **2009**, *5*, 242 – 256.

48. Zhou, Y.; Vitkup, D.; Carplus, M. F. Native proteins are surface-molten solids: Application of the lindemann criterion for the solid versus liquid State. *J. Mol. Biol*. **1999**, *285*, 1371 – 1375.

49. Rueda, M.; Ferrer-Costas, C.; Meyer, T.; Perez, A.; Campus, J.; Hospital, A.; Gelpi, L. J.; Orozco, M.; A consensus view of protein dynamics. *Natl. Acad. Sci. USA*. **2007**, *104*, 796 – 801.

50. González, M.A.; Abascal, J.L.F; The shear viscosity of rigid water models. *J. Chem. Phys.* **2010**, *132*, 096101 – 2.

51. Derlacki, Z.J.; Easteal, A.J.; Edge, A. V. J.; Woolf, L. A. Diffusion coefficients of methanol and water and the mutual diffusion coefficient in methanol-water solution at 278 K and 298 K. *J. Phys. Chem.* **1985**, *89*, 5318 – 5322.

52. Wang, J.; Hou, T. Application of molecular dynamics simulations in molecular property prediction II: Diffusion coefficient. *J. Comput. Chem.* **2011**, *32*, 3505 – 3519.

53. Hess, B. Determining the shear viscosity of model liquids from molecular dynamics simulations. *J. Chem. Phys.* **2002**, *116*, 209 – 217.

54. Pletneva, E.V.; Gray, B. H.; Winkler, J.R.; Snapshots of cytochrome c folding. *PNAS.* **2005**, *113*, 18397 – 18402.

55. Jana, B.; Pal, Subrata; Bagchi, B.; Hydrogen bond breaking mechanism and water representational dynamics in the hydration layer of lysozyme. *J. Phys. Chem. B,* **2008**, *112*, 9112 – 9117.

56. Jain, R.; Sharma, D.; Kumar, R. Effect of alcohols on the stability and low-frequency local motions that control the slow changes in structural dynamics of ferrocytochrome c. *J. Biochem.* **2013**, *154*, 341 – 354.

57. Parak, F.G.; McMahon, H.; Frauenfelder, H.; Fenimore, P. W. Slaving: Solvent fluctuations dominate protein dynamics and functions. *PANS.* **2002**, *99*, 16047 – 16051.

58. Hayward, S.; Kitao, A.; Hirata, F.; Go, N.; Effect of solvent on collective motions in globular protein. *J. Mol. Biol.* **1993**, *234*, 1207 – 1217.

59. Tarek, M.; Tobias, D. J.; Role of protein-water hydrogen dynamics in the protein dynamical transition. *Phy. Rev. Lett.* **2002**, *88*, 138101(9).

60. Sterpone, F.; Stirnemann, G.; Hynes, J. T.; Laage, D. Water hydrogen-bond dynamics around amino acids: The key role of hydrophilic hydrogen-bond acceptor groups. *J. Phys. Chem. B,* **2010**, *114*, 2083 – 2089.

61. Colombo, G.; Carrea, G. Modeling enzyme reactivity in organic solvents and water through computer simulations. *J. Biotechnol.* **2002**, *96*, 23 – 33.

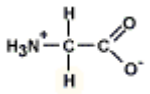
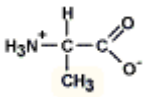
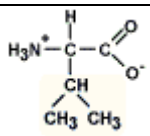
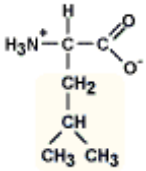
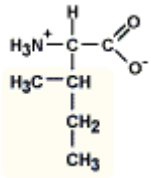
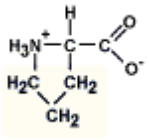
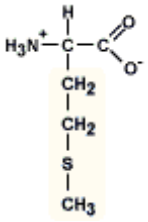
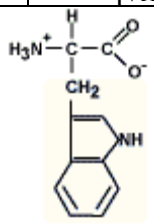
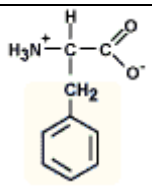
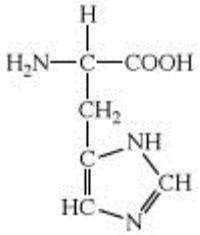
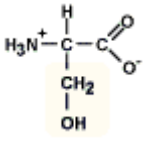
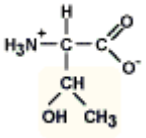
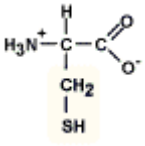
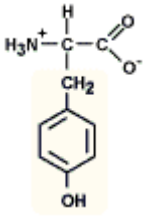
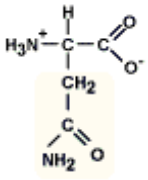
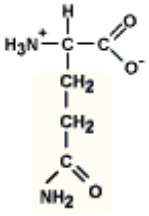
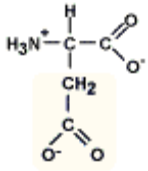
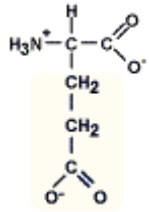
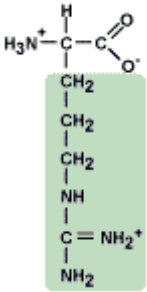
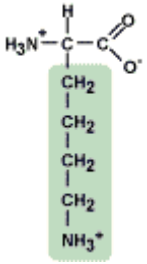
62. Ghosh, R.; Roy, S.; Bagchi, B. Solvent sensitivity of protein unfolding: Dynamical study of chicken villin headpiece subdomain in water-ethanol binary mixture. *J. Phys. Chem. B*, **2013**, *117*, 15625 – 15638.

63. Hopkins, P.; Fortini, A.; Archer, A. J.; Schmidt, M. The van Hove distribution function for Brownian hard spheres: Dynamical test particle theory and computer simulations for bulk dynamics. *J. Chem. Phys.* **2010**, *133*, 224505 (18).

APPENDICES

APPENDIX-I

List of Amino Acids [28, google edited pictures]

NON POLAR AMINO ACIDS					
 GLYCINE **[GLY]**	 ALANINE **[ALA]**	 VALINE **[VAL]**	 LEUCINE **[LEU]**	 ISOLEUCINE **[ILE]**	 PROLINE **[PRO]**
 METHIONINE **[MET]**	 TRYPTOPHAN **[TRP]**	 PHANYLALANINE **[PHE]**	 HISTIDINE **[HIS]**		
POLAR AMINOACIDS					
 SERINE **[SER]**	 THREONINE **[THR]**	 CYSTINE **[CYS]**	 TYROSINE **[TYR]**	 ASPERGINE **[ASN]**	 GLUTAMINE **[GLN]**
NEGATIVELY CHARGED AMINO ACIDS		**POSITIVELY CHARGED AMINO ACIDS**			
 ASPERTIC ACID **[ASP]**	 GLUTAMIC ACID **[GLU]**	 ARGININE **[ARG]**	 LYSINE **[LYSH]**		

APPENDIX-II

Calculation for Solvent Composition.

CALCULATION OF NUMBER OF MOLECULES OF WATER AND METHANOL IN A BOX OF MIXTURE OF REQUIRED PROPORTION.

Density of Pure Methanol at 20^0C = 0.7917 gram cm^{-3}

Molar Mass of Methanol, M_{MeOH} = 32.04186 gram

$$Molar\ Volume\ of\ MeOH = \frac{32.04186}{0.7917} = 40.472224\ cm^3$$

$$1\ Molecule\ Volume\ of\ MeOH = \frac{40.472224 \times 10^{24}}{6.02214 \times 10^{23}} = 67.20572\ Angstrom^3$$

Density of Pure Water at 20^0C = 0.9982067 gram cm^{-3}

Molar Mass of Water, M_{H2O} = 18.01528 gram

$$Molar\ Volume\ of\ Water = \frac{18.01528}{0.9982067} = 18.0477645\ cm^3$$

$$1\ Molecule\ Volume\ of\ Water = \frac{18.0477645 \times 10^{24}}{6.02214 \times 10^{23}} = 29.968823\ Angstrom^3$$

For 40% by mass of MeOH in mixture with water in 22.00 $Å^3$

$$\frac{\left[\frac{N_{MeOH} \times M_{MeOH}}{N_A}\right]}{\left[\frac{N_{H2O} \times M_{H2O}}{N_A}\right]} = \frac{4}{6}$$

$$\frac{N_{MeOH}}{N_{H2O}} = \frac{2}{3}\left(\frac{M_{H2O}}{M_{MeOH}}\right)$$

Density of Mixture (ρ) = 0.9347 gram cm^{-3}

$$N_{MeOH} \times \frac{M_{MeOH}}{N_A} + \frac{M_{H2O}}{N_A} \times N_{H2O} = \rho \times V_{Box}$$

$$N_{MeOH} = \frac{\rho \times V_{box}}{10^{24} \times M_{MeOH}} 6.02214 \times 10^{23} - N_{H2O} \times \frac{M_{H2O}}{M_{MeOH}}$$

$$\frac{5}{2} N_{MeOH} = \rho \times V_{Box}\ (Å^3) \times 0.0187946$$

$$N_{MeOH} = 74.82 \approx 75 \qquad N_{H2O} = 199.62 \approx 200$$

APPENDIX-III

Simulation Parameters

RUN CONTROL PARAMETERS

```
integrator          = md ; simple leap-frog molecular dynamics
                         ; algorithm.
tinit               = 0  ; initial time is zero with start of
                         ; data production.
dt                  = 0.002 ; 2.0 femtosecond integration step.

nsteps              = 50000000; 100 ns simulation time.

comm_mode           = linear; Mode for Centre of Mass motion
                      removal.
nstcomm             = 10 ; number of steps for center of mass
                           motion removal.
comm_grps           = CL   MeOH    Water   Protein   HEM
```

ENERGY MINIMIZATION OPTIONS

```
integrator        = steep   ; A steepest descent algorithm
                            ; for energy minimization.
emtol             = 50      ; stepwise minimizing force
                            ; tolerance from 1000 to 50 KJ/MOL.
emstep            = 0.002
nsteps            = 100000
;
niter             = 50      ; sytem relaxation frequency.
fcstep            = 0       ; no constraint is used for
minimization.
```

OUTPUT CONTROL OPTIONS::for data production,collect in every 1 picosecond.

```
nstxout           = 500    ; coordinate.
xtc_precision     = 50     ; max.precision to write coordinate.
nstvout           = 500    ; velocity.
nstfout           = 500    ; force.
nstlog            = 500    ; write in logfile.
nstcalcenergy     = -1     ;
```

```
nstenergy         = 500    ; energy.

Selection of energy groups-Energy Groups are devided to update
energy in different category.

energygrps        = CL  MeOH   A-SOL   C-SOL   Protein   HEM

OUTPUT CONTROL FOR EQUILIBRATION STEP :: 0.02 picoseconds.


NEIGHBORSEARCHING PARAMETERS
nstlist           = 10   ; nblist update frequency
ns_type           = grid ; using grid algorithm for neighbour
                           group searching
                         ; will be fast and efficient.
pbc               = xyz  ; Apply Periodic Boundary Condition in
                           all directions.
periodic_molecules    = no
rlist              = 1.4 ; nm, do not find neighbour group
                           beyond.

OPTIONS FOR ELECTROSTATICS AND VDW

coulombtype       = PME  ; Particle Mesh Ewald method for
                           calculating electrostatistics.
rcoulomb-switch   = 0    ; do direct cut off after 1.4
                           nanometer.
rcoulomb          = 1.4  ;  do calculate columbic energy up to
                           1.4 nanometer.
vdw-type          = Cut-off ; direct space cut-off for Van der
                              Waals energy calculations.
rvdw              = 1.4     ; same cut-off.

DispCorr          = EnerPres ; apply long-range dispersion
                               correction for energy and
                               pressure.

fourierspacing     = 0.10     ; grid spacing for PME is  1.0
Angstroms.

TEMPERATURE COUPLING
tcoupl             = berendsen
nsttcouple         = -1
nh-chain-length    = 10
; Groups to couple separately
tc_grps            = CL   MeOH   Water   Protein   HEM
; Time constant (ps) and reference temperature (K)
```

```
tau_t                = 0.2     0.2       0.2       0.2       0.2
ref_t                = 298.15  298.15    298.15    298.15
298.15
```

PRESSURE COUPLING

```
pcoupl               = berendsen
Pcoupltype           = Isotropic
nstpcouple           = -1
; Time constant (ps), and reference P (bar)
tau-p                = 0.2
ref-p                = 1.01325
; Scaling of reference coordinates:
refcoord_scaling     = All
```

SIMULATED ANNEALING:: TO WARMUP SYSTEM IN FIRST EQUILIBRATION STEP

```
; Type of annealing for each temperature group
annealing            = single  single  single  single  single
; Number of time points to use for specifying annealing in each
group
annealing_npoints    = 31  31  31  31  31
; List of times at the annealing points for each group
annealing_time       =  0    4    8   12   16   18   22   26
                        28   32   36  42   48   52   56   60
                        66   72   78  84   90   96   102  108
                        114  120  126 132  138  144  150
                        0    4    8   12   16   18   22   26
                        28   32   36  42   48   52   56   60
                        66   72   78  84   90   96   102  108
                        114  120  126 132  138  144  150
                        0    4    8   12   16   18   22   26
                        28   32   36  42   48   52   56   60
                        66   72   78  84   90   96   102  108
                        114  120  126 132  138  144  150
                        0    4    8   12   16   18   22   26
                        28   32   36  42   48   52   56   60
                        66   72   78  84   90   96   102  108
                        114  120  126 132  138  144  150
                        0    4    8   12   16   18   22   26
                        28   32   36  42   48   52   56   60
                        66   72   78  84   90   96   102  108
                        114  120  126 132  138  144  150

; Temp. at each annealing point, for each group.
annealing_temp       = 20   30   40   50   60   70   80    90
                       100  110  120  130  140  150  160   170
```

```
                    180   190  200  210  220  230  240   250
                    260   270  280  290  295  298  298.15
                     20   30   40   50   60   70   80    90
                    100   110  120  130  140  150  160   170
                    180   190  200  210  220  230  240   250
                    260   270  280  290  295  298  298.15
                     20   30   40   50   60   70   80    90
                    100   110  120  130  140  150  160   170
                    180   190  200  210  220  230  240   250
                    260   270  280  290  295  298  298.15
                     20   30   40   50   60   70   80    90
                    100   110  120  130  140  150  160   170
                    180   190  200  210  220  230  240   250
                    260   270  280  290  295  298  298.15
```

```
; GENERATE VELOCITIES FOR STARTUP RUN
gen_vel                 = yes
gen_temp                = 20
gen_seed                = 2013

SIMULATED ANNEALING:: IN SECOND EQUILIBRATION STEP AND DATA
PRODUCTION
annealing               = no
```

OPTIONS FOR BONDS:: CONSTRAINTS

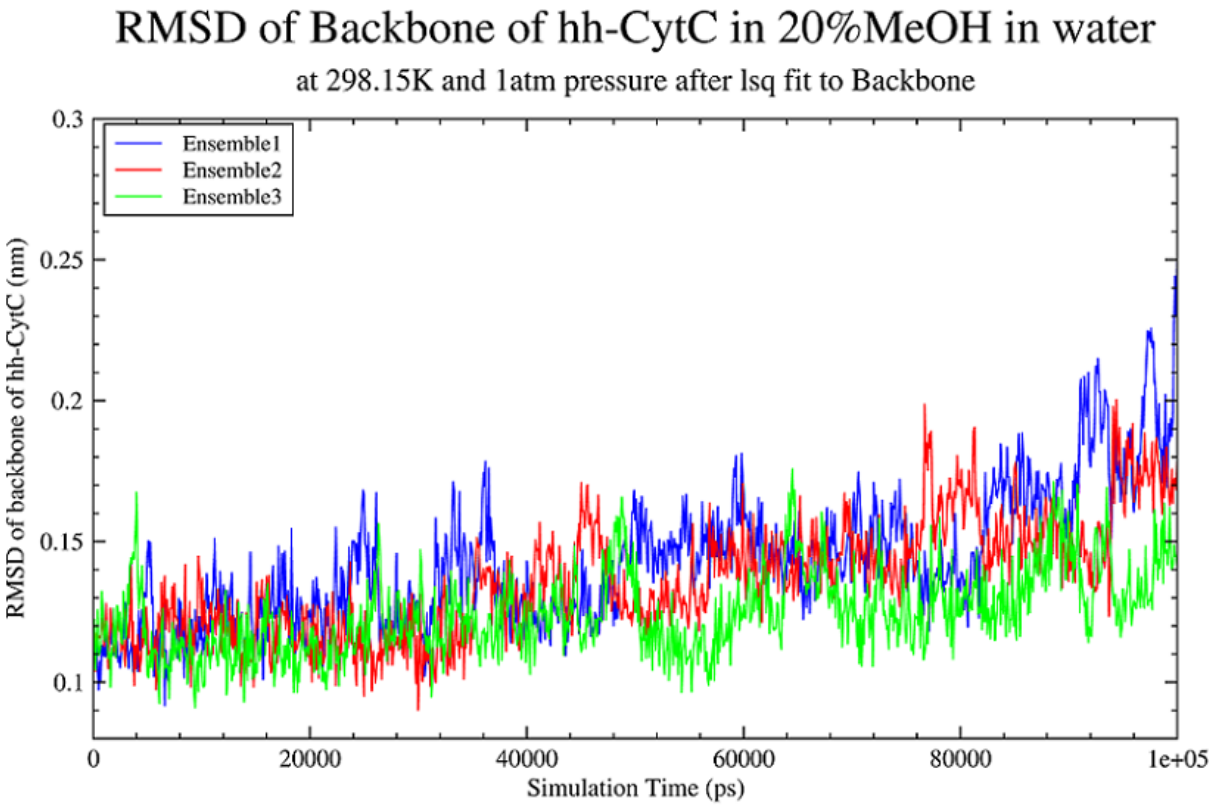
```
constraints             = h-bonds
constraint-algorithm    = Lincs
; Do not constrain the start configuration
continuation            = yes ; in data production.
lincs-order             = 4
lincs-iter              = 1 ; for equilibration and data
production.
lincs-iter              = 4 ; for energy minimization.
;Allow LINCS to write warning if a bond rotates over more
degrees than
lincs-warnangle         = 30

-----------------------------------------------------------------
-----------------------------------------------------------------
```

APPENDIX-IV

RMSD of hh-CytC in three different ensembles

Figure depicts the variation of RMSD of C-α backbone in different ensembles in data production step of MD simulation in 20%MeOH water-methanol binary mixture in same NPT condition, indicating that hh-CytC follows different mechanism of conformational transition in same condition. Increased RMSD at the end of simulation may infer the possibility of unfolding of hh-CytC in future.

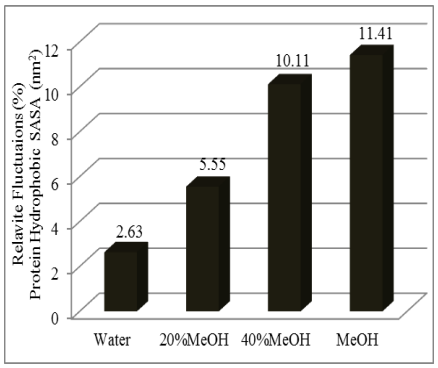


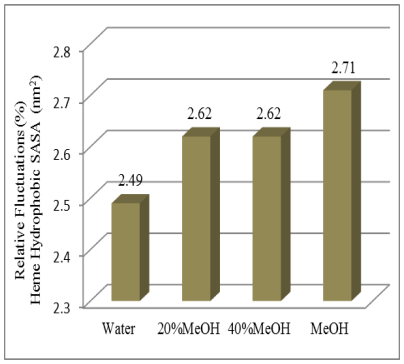
APPENDIX-V

Solvent Accessible Surface Area (SASA) in Horse Heart Cytochrome C

Solvent Accessible Surface Area (SASA) was calculated from gromacs program with probe 0.14nm in different solvents at 298.15K, 1atm.

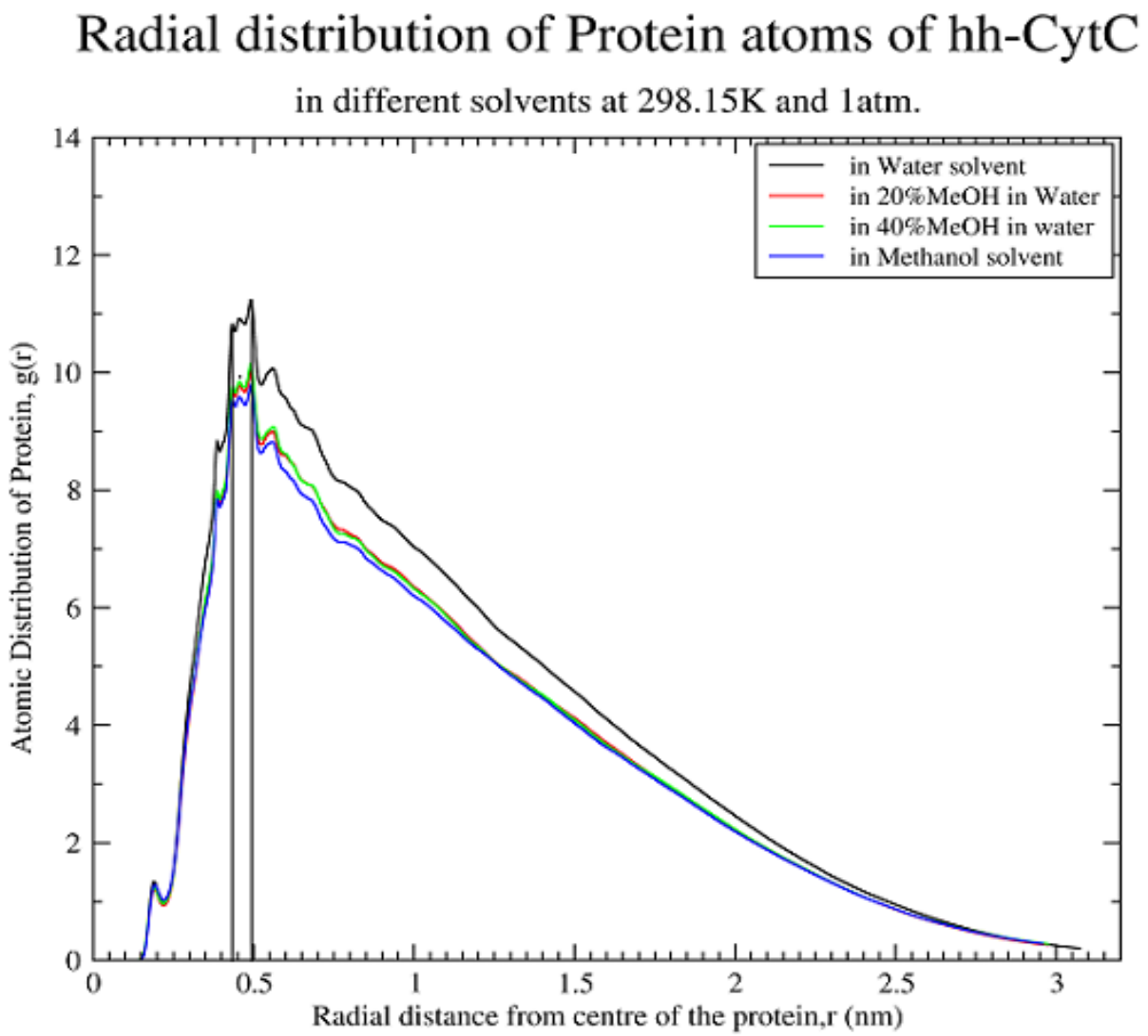
Solvent	Surface Area	Protein (nm^2)	hh-CytC(nm^2)	HEM(nm^2)
Water	**Hydrophobic**	39.541±1.038	42.433±0.971	5.973±0.149
	Hydrophilic	27.562±0.798	29.169±0.868	1.672±0.072
	Total	67.103±1.430	71.757±1.444	7.654±0.175
20%MeOH in Water	**Hydrophobic**	40.344±2.241	43.755±1.027	5.916±0.155
	Hydrophilic	27.906±0.701	29.818±0.840	1.711±0.071
	Total	68.181±1.746	73.573±1.746	7.652±0.174
40%MeOH in Water	**Hydrophobic**	39.020±3.965	42.510±0.978	5.986±0.157
	Hydrophilic	28.104±0.582	29.033±0.807	1.680±0.07
	Total	68.132±1.393	72.878±1.406	7.63±0.177
MeOH	**Hydrophobic**	39.202±4.471	42.930±1.017	5.937±0.161
	Hydrophilic	27.563±0.593	29.702±0.908	1.658±0.079
	Total	69.140±1.750	70.207±1.759	7.63±0.177





APPENDIX-VI

Atomic Radial Distribution inside Protein in Different Solvents

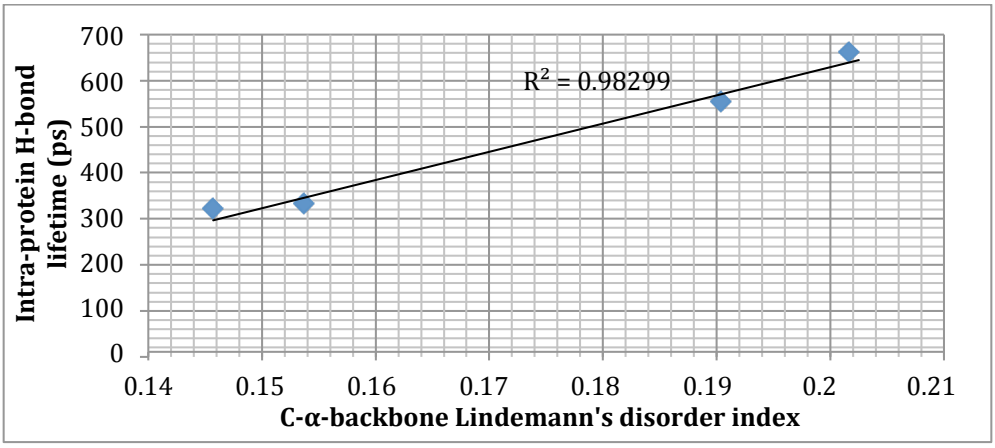


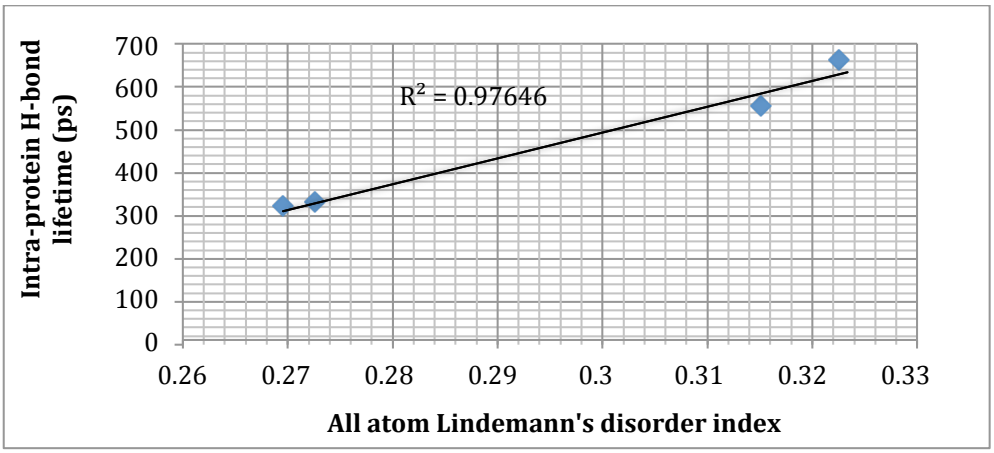
Estimation of t value of "the most-probable non-bonded near-neighbour distance", (*a´*) from radial distribution graph is based on the method adopted by M.Karplus and et *al* [48]. In our calculation for "Lindemann's disorder index" for hh-CytC in different solvent systems is 0.48 nm (4.80 Å).

APPENDIX-VII

Correlation Analysis between Lindemann's disorder index and intra-protein H-bond lifetime.

Following graphs indicate that there is strong correlation between Lindemann's disorder index and intra-protein H-bond lifetime. With increase of protein internal motion or C-α-backbone Lindemann's disorder index, short-chain or intrahelix (or within secondary structure or intra-foldon) H-bond become more probable with long H-bond lifetime (or persistent) compared to long-chain or inter-foldon H-bond. In our opinion this is possible only in folded stable or meta-stable protein structures of protein.





APPENDIX-VIII

FORTRAN Program to Track Solvent Molecules in Conserved Sites in hh-CytC.

```
c      FORTRAN-77 PRORTRAM TO FIND SOLVENT MOLECULES THAT ARE
H-  BONDED            AND RESIDED IN THE CONSERVED SITES INSIDE
PROTEIN
c      FOR 20%MeOH in water at 298.15K and 1 atm. Pressure.
       IMPLICIT NONE

       PROGRAM      FIND@SOLVENT

c      ASSSIGN THE PAREMETER TO READ FROM COORDINATE FILES

        INTEGER,PARAMETER :: frame = 100000
        INTEGER,PARAMETER :: n_me =  718
        INTEGER,PARAMETER :: n_w = 5222
        INTEGER ::  i,j,k,a,b,e,f
        INTEGER :: at_nom(3*n_me,frame),res_nm(n_me,frame)
        INTEGER :: at_now(3*n_w,frame),res_nw(n_w,frame)
        INTEGER :: c1,c2
        REAL ::  c3,c4

        REAL :: x_O1A(frame),y_O1A(frame),z_O1A(frame),
     $          x_O2A(frame),y_O2A(frame),z_O2A(frame)

        REAL :: r_O1A(n_me,frame),r_O2A(n_me,frame),
     $  x_om(n_me,frame),y_om(n_me,frame),z_om(n_me,frame),
     $  x_hm(n_me,frame),y_hm(n_me,frame),z_hm(n_me,frame),
     $  x_me(n_me,frame),y_me(n_me,frame),z_me(n_me,frame)

        REAL :: r_1a(n_w,frame),r_2a(n_w,frame),
     $  x_ow(n_w,frame),y_ow(n_w,frame),z_ow(n_w,frame),
     $  x_hw1(n_w,frame),y_hw1(n_w,frame),z_hw1(n_w,frame),
     $  x_hw2(n_w,frame),y_hw2(n_w,frame),z_hw2(n_w,frame)

        CHARACTER*4 :: ATOM
        CHARACTER*3 ::
O1A,O2A,MeO,HEM,Omet,HMet,CMet,SOL,OW,HW1,HW2

c       READ PARAMETER FROM THE PDB-FILE OF DATA PRODUCITON
STEP,100ns
```

```
c       SIMULATION WITH TREJECTORY WRITTEN IN 1PS

C       SOURCE: READ THE FILE OF COORDINATE TRAJECTORY TO FIND
MOLECULE

       OPEN (UNIT=100,file='5data-trjconv-wm20.pdb'
     $  , action='read',status='old')

c       COORDINATE FILE TO TEST FIRST IF PROGRAMM CODE WORK OR
NOT
c        OPEN(UNIT=100,file='test.pdb'
c     $  ,action='read',status='old')

c       ESCAPE SOME STEP IF NEEDED
c        do f = 1,758120
c         READ(100,*)
c        enddo

c       READ  DATA FROM EACH LINE

        do i = 1,frame
          do a = 1,1093
           READ(100,*)
         enddo

c       TAKE  HEM-PROPIONATE AS A REFERENCE H-BONDING SITE AND
READ IT

      READ(100,1000)
ATOM,c1,O1A,HEM,c2,x_O1A(i),y_O1A(i),z_O1A(i),c3,c4
 1000
FORMAT(BN,A4,2X,I5,2X,A3,1X,A3,2X,I4,6X,3(F6.3,2X),2(F4.2,2X),1
0X)
      READ(100,1010)
ATOM,c1,O2A,HEM,c2,x_O2A(i),y_O2A(i),z_O2A(i),c3,c4
 1010
FORMAT(BN,A4,2X,I5,2X,A3,1X,A3,2X,I4,6X,3(F6.3,2X),2(F4.2,2X),1
0X)
        do b = 1,30
        READ(100,*)
        enddo

c       READ THE MEOH COORDINATE FIRST AS IN SEQUENCE IN
COORDINATE FILE
        do j = 1,n_me
        READ(100,1100) ATOM,at_nom(j,i), Omet, MeO,
     &  res_nm(j,i), x_om(j,i), y_om(j,i), z_om(j,i),c3,c4
```

```
 1100
FORMAT(BN,A4,2X,I4,1X,A4,1X,A3,2X,I4,6X,3(F6.3,2X),2(F4.2,2X),1
0X)
        READ(100,1200) ATOM,at_nom(j,i), HMet, MeO,
     &  res_nm(j,i), x_hm(j,i), y_hm(j,i), z_hm(j,i),c3,c4
 1200
FORMAT(BN,A4,2X,I4,1X,A4,1X,A3,2X,I4,6X,3(F6.3,2X),2(F4.2,2X),1
0X)
        READ(100,1300) ATOM,at_nom(j,i), CMet, MeO,
     &  res_nm(j,i), x_me(j,i), y_me(j,i), z_me(j,i),c3,c4
 1300
FORMAT(BN,A4,2X,I4,1X,A4,1X,A3,2X,I3,6X,3(F6.3,2X),2(F4.2,2X),1
0X)
        enddo

c       READ THE H2O COORDINATE THEN
        do k = 1, n_w
       READ(100,2100) ATOM,at_now(k,i), OW, SOL,
     &  res_nw(k,i), x_ow(k,i), y_ow(k,i), z_ow(k,i),c3,c4
 2100
FORMAT(BN,A4,2X,I5,2X,A3,1X,A3,2X,I4,6X,3(F6.3,2X),2(F4.2,2X),1
0X)
      READ(100,2200) ATOM,at_now(k,i), HW1,SOL,
     &  res_nw(k,i), x_hw1(k,i), y_hw1(k,i), z_hw1(k,i),c3,c4
 2200
FORMAT(BN,A4,2X,I5,2X,A3,1X,A3,2X,I4,6X,3(F6.3,2X),2(F4.2,2X),1
0X)
      READ(100,2300) ATOM,at_now(k,i),HW2,SOL,
     &  res_nw(k,i), x_hw2(k,i), y_hw2(k,i), z_hw2(k,i),c3,c4
 2300
FORMAT(BN,A4,2X,I5,2X,A3,1X,A3,2X,I4,6X,3(F6.3,2X),2(F4.2,2X),1
0X)
       enddo
       do e = 1,9
        READ(100,*)
       enddo
       enddo

c    CALCULATE THE DISATNCE BETWEEN MEOH OR WATER  AND H-
BONDING SITE
      do i = 1,frame
       do j = 1,n_me
        r_O1A(j,i)=sqrt((x_O1A(i)-x_om(j,i))**2+(y_O1A(i)-
y_om(j,i))**2
     &  +(z_O1A(i)-z_om(j,i))**2)
        r_O2A(j,i)=sqrt((x_O2A(i)-x_om(j,i))**2+(y_O2A(i)-
y_om(j,i))**2
```

```
     &  +(z_O2A(i)-z_om(j,i))**2)
       enddo
       do k = 1,n_w
       r_1a(k,i)=sqrt((x_O1A(i)-x_ow(k,i))**2+(y_O1A(i)-
y_ow(k,i))**2
     &  +(z_O1A(i)-z_ow(k,i))**2)
       r_2a(k,i)=sqrt((x_O2A(i)-x_ow(k,i))**2+(y_O2A(i)-
y_ow(k,i))**2
     &  +(z_O2A(i)-z_ow(k,i))**2)
       enddo
      enddo
c   CALCULATE THE  ANGLE BETWEEN MEOH OR WATER AND H-BONDING
SITE
c   WE USE THE GROMACS PROGRAM TO CALCULATE THE ANGLE H-D-A
BETWEEN
c   DONOR AND ACCEPTOR/

c   WRITE THE MEOH MOLECULES IN THE CONSERVED SITE THAT REPLACE
c   CRYST.H2O

        OPEN(UNIT=200,file='meoh-hem-OA1-allsteps.pdb'
     &  ,action='write',status='new')
       do i = 1,frame
         do j = 1,n_me
          IF ( r_O1A(j,i).LT.6.0 .AND. r_O2A(j,i).LT.6.0) THEN
        WRITE(200,*)  i*1.0
        WRITE(200,3100) at_nom(j,i),MeO,res_nm(j,i),r_O1A(j,i),
     &  r_O2A(j,i)
 3100   FORMAT(BN,2X,I5,2X,A4,2X,I4,4X,2(F11.8))
        ENDIF
        enddo
       enddo

c     WRITE THE SOLVENT H2O MOLECULES IN THE CONSERVED SITE
THAT
c     REPLACE THE CRYST.H2O

       OPEN(UNIT=400,file='h2o-hem-OA1-allsteps.pdb'
     & ,action='write',status='new')
        do i = 1,frame
         do k = 1,n_w
          IF ( r_1a(k,i).LT.7.0 .AND. r_2a(k,i).LT.7.0) THEN
        WRITE(400,*)  i*1.0
        WRITE(400,4100) at_now(k,i),SOL,res_nw(k,i),r_1a(k,i),
     &  r_2a(k,i)
 4100   FORMAT(BN,2X,I5,2X,A3,I4,4x,2(F11.8))
         ENDIF
```

```
        enddo
        enddo
        END PROGRAM FIND-WATER
```

==========

APPENDIX-IX

Topology File Horse Heart Cytochrome C in 20% aqueous MeOH solvent.

```
;    File 'r_cytc.top' was generated
;    By user: onbekend (0)
;    On host: onbekend
;    At date:  Dec 10 10:30:21 2013
;
;    This is a standalone topology file
;
;    It was generated using program:
;    pdb2gmx - VERSION 4.5.5
;     USED FOR SIMULATION OF HORSE-HEART CYTOCHROME C IN
20%MEOH IN WATER
;
;    Command line was:
;    pdb2gmx -f r_cytc.pdb -o r_cytc.gro -p r_cytc.top -i
     posre_r_cytc.itp -merge all -water spce
;
;    Force field was read from the standard Gromacs share
directory.
;
; Include forcefield parameters
#include "gromos53a6.ff/forcefield.itp"

[ moleculetype ]
; Name            nrexcl
 MeOH               2

[ atoms ]
;   nr       type  resnr residue  atom   cgnr     charge
mass  typeB    chargeB      massB
; residue   0 MeOH rtp CH3OH q  0.0
     1       OMet       1  MeOH   Omet      1     -0.674
15.9994   ; qtot -0.674
     2          H       1  MeOH   HMet      1      0.408
1.008   ; qtot -0.266
     3       CMet       1  MeOH   CMet      1      0.266
15.035   ; qtot 0

[ bonds ]
;  ai    aj funct            c0            c1            c2
c3
```

```
    1     2     2    gb_1
    1     3     2    gb_27
    2     3     2    gb_51

[ angles ]
;  ai    aj    ak funct            c0            c1
c2            c3
    2     1     3     2    ga_47
;    1     2     3     2
;    1     3     2     2
;
[ moleculetype ]
; Name            nrexcl
r_cytc             3

[ atoms ]
;   nr       type  resnr residue  atom   cgnr     charge
mass  typeB    chargeB      massB
; residue   1 GLY rtp GLY  q +1.0
     1         NL      1    GLY      N      1      0.129
14.0067   ; qtot 0.129
     2          H      1    GLY     H1      1      0.248
1.008   ; qtot 0.377
     3          H      1    GLY     H2      1      0.248
1.008   ; qtot 0.625
     4          H      1    GLY     H3      1      0.248
1.008   ; qtot 0.873
     5        CH2      1    GLY     CA      2      0.127
14.027   ; qtot 1
     6          C      1    GLY      C      3       0.45
12.011   ; qtot 1.45
     7          O      1    GLY      O      3      -0.45
15.9994   ; qtot 1
;
; residue   2 ASP rtp ASP  q -1.0
     8          N      2    ASP      N      4      -0.31
14.0067   ; qtot 0.69
     9          H      2    ASP      H      4       0.31
1.008   ; qtot 1
    10        CH1      2    ASP     CA      5          0
13.019   ; qtot 1
    11        CH2      2    ASP     CB      5          0
14.027   ; qtot 1
    12          C      2    ASP     CG      6       0.27
12.011   ; qtot 1.27
    13         OM      2    ASP    OD1      6     -0.635
15.9994   ; qtot 0.635
```

```
    14          OM      2    ASP    OD2      6     -0.635
15.9994   ; qtot 0
    15           C      2    ASP      C      7       0.45
12.011   ; qtot 0.45
    16           O      2    ASP      O      7      -0.45
15.9994
;
; residue   3 VAL rtp VAL  q  0.0
    17           N      3    VAL      N      8      -0.31
14.0067   ; qtot -0.31
    18           H      3    VAL      H      8       0.31
1.008   ; qtot 0
    19         CH1      3    VAL     CA      9          0
13.019   ; qtot 0
    20         CH1      3    VAL     CB      9          0
13.019   ; qtot 0
    21         CH3      3    VAL    CG1      9          0
15.035   ; qtot 0
    22         CH3      3    VAL    CG2      9          0
15.035   ; qtot 0
    23           C      3    VAL      C     10       0.45
12.011   ; qtot 0.45
    24           O      3    VAL      O     10      -0.45
15.9994   ; qtot 0
;
; residue   4 GLU rtp GLU  q -1.0
    25           N      4    GLU      N     11      -0.31
14.0067   ; qtot -0.31
    26           H      4    GLU      H     11       0.31
1.008   ; qtot 0
    27         CH1      4    GLU     CA     12          0
13.019   ; qtot 0
    28         CH2      4    GLU     CB     12          0
14.027   ; qtot 0
    29         CH2      4    GLU     CG     12          0
14.027   ; qtot 0
    30           C      4    GLU     CD     13       0.27
12.011   ; qtot 0.27
    31          OM      4    GLU    OE1     13     -0.635
15.9994   ; qtot -0.365
    32          OM      4    GLU    OE2     13     -0.635
15.9994   ; qtot -1
    33           C      4    GLU      C     14       0.45
12.011   ; qtot -0.55
    34           O      4    GLU      O     14      -0.45
15.9994   ; qtot -1
;
```

```
; residue   5 LYS rtp LYSH q +1.0
    35          N      5    LYS      N     15      -0.31
14.0067   ; qtot -1.31
    36          H      5    LYS      H     15       0.31
1.008   ; qtot -1
    37        CH1      5    LYS     CA     16          0
13.019   ; qtot -1
    38        CH2      5    LYS     CB     16          0
14.027   ; qtot -1
    39        CH2      5    LYS     CG     17          0
14.027   ; qtot -1
    40        CH2      5    LYS     CD     17          0
14.027   ; qtot -1
    41        CH2      5    LYS     CE     18      0.127
14.027   ; qtot -0.873
    42         NL      5    LYS     NZ     18      0.129
14.0067   ; qtot -0.744
    43          H      5    LYS    HZ1     18      0.248
1.008   ; qtot -0.496
    44          H      5    LYS    HZ2     18      0.248
1.008   ; qtot -0.248
    45          H      5    LYS    HZ3     18      0.248
1.008   ; qtot 0
    46          C      5    LYS      C     19       0.45
12.011   ; qtot 0.45
    47          O      5    LYS      O     19      -0.45
15.9994   ; qtot 0
;
; residue   9 ILE rtp ILE  q  0.0
    79          N      9    ILE      N     33      -0.31
14.0067   ; qtot 1.69
    80          H      9    ILE      H     33       0.31
1.008   ; qtot 2
    81        CH1      9    ILE     CA     34          0
13.019   ; qtot 2
    82        CH1      9    ILE     CB     35          0
13.019   ; qtot 2
    83        CH2      9    ILE    CG1     35          0
14.027   ; qtot 2
    84        CH3      9    ILE    CG2     35          0
15.035   ; qtot 2
    85        CH3      9    ILE     CD     35          0
15.035   ; qtot 2
    86          C      9    ILE      C     36       0.45
12.011   ; qtot 2.45
    87          O      9    ILE      O     36      -0.45
15.9994   ; qtot 2
```

```
;
; residue  10 PHE rtp PHE  q  0.0
    88          N     10    PHE      N     37      -0.31
14.0067   ; qtot 1.69
    89          H     10    PHE      H     37       0.31
1.008   ; qtot 2
    90        CH1     10    PHE     CA     38          0
13.019   ; qtot 2
    91        CH2     10    PHE     CB     38          0
14.027   ; qtot 2
    92          C     10    PHE     CG     38          0
12.011   ; qtot 2
    93          C     10    PHE    CD1     39      -0.14
12.011   ; qtot 1.86
    94         HC     10    PHE    HD1     39       0.14
1.008   ; qtot 2
    95          C     10    PHE    CD2     40      -0.14
12.011   ; qtot 1.86
    96         HC     10    PHE    HD2     40       0.14
1.008   ; qtot 2
    97          C     10    PHE    CE1     41      -0.14
12.011   ; qtot 1.86
    98         HC     10    PHE    HE1     41       0.14
1.008   ; qtot 2
    99          C     10    PHE    CE2     42      -0.14
12.011   ; qtot 1.86
   100         HC     10    PHE    HE2     42       0.14
1.008   ; qtot 2
   101          C     10    PHE     CZ     43      -0.14
12.011   ; qtot 1.86
   102         HC     10    PHE     HZ     43       0.14
1.008   ; qtot 2
   103          C     10    PHE      C     44       0.45
12.011   ; qtot 2.45
   104          O     10    PHE      O     44      -0.45
15.9994   ; qtot 2
;
; residue  12 GLN rtp GLN  q  0.0
   113          N     12    GLN      N     48      -0.31
14.0067   ; qtot 1.69
   114          H     12    GLN      H     48       0.31
1.008   ; qtot 2
   115        CH1     12    GLN     CA     49          0
13.019   ; qtot 2
   116        CH2     12    GLN     CB     49          0
14.027   ; qtot 2
```

```
   117        CH2     12    GLN     CG     49          0
14.027   ; qtot 2
   118          C     12    GLN     CD     50       0.29
12.011   ; qtot 2.29
   119          O     12    GLN    OE1     50      -0.45
15.9994   ; qtot 1.84
   120         NT     12    GLN    NE2     50      -0.72
14.0067   ; qtot 1.12
   121          H     12    GLN   HE21     50       0.44
1.008   ; qtot 1.56
   122          H     12    GLN   HE22     50       0.44
1.008   ; qtot 2
   123          C     12    GLN      C     51       0.45
12.011   ; qtot 2.45
   124          O     12    GLN      O     51      -0.45
15.9994   ; qtot 2
;
; residue  14 CYS rtp CYS2 q  0.0
   138          N     14    CYS      N     57      -0.31
14.0067   ; qtot 2.69
   139          H     14    CYS      H     57       0.31
1.008   ; qtot 3
   140        CH1     14    CYS     CA     58          0
13.019   ; qtot 3
   141        CH2     14    CYS     CB     58          0
14.027   ; qtot 3
   142          S     14    CYS     SG     58          0
32.06   ; qtot 3
   143          C     14    CYS      C     59       0.45
12.011   ; qtot 3.45
   144          O     14    CYS      O     59      -0.45
15.9994   ; qtot 3
;
; residue  15 ALA rtp ALA  q  0.0
   145          N     15    ALA      N     60      -0.31
14.0067   ; qtot 2.69
   146          H     15    ALA      H     60       0.31
1.008   ; qtot 3
   147        CH1     15    ALA     CA     61          0
13.019   ; qtot 3
   148        CH3     15    ALA     CB     61          0
15.035   ; qtot 3
   149          C     15    ALA      C     62       0.45
12.011   ; qtot 3.45
   150          O     15    ALA      O     62      -0.45
15.9994
;
```

```
; residue  17 CYS rtp CYS2 q  0.0
   163          N     17    CYS      N     67      -0.31
14.0067   ; qtot 2.69
   164          H     17    CYS      H     67       0.31
1.008   ; qtot 3
   165        CH1     17    CYS     CA     68          0
13.019   ; qtot 3
   166        CH2     17    CYS     CB     68          0
14.027   ; qtot 3
   167          S     17    CYS     SG     68          0
32.06   ; qtot 3
   168          C     17    CYS      C     69       0.45
12.011   ; qtot 3.45
   169          O     17    CYS      O     69      -0.45
15.9994   ; qtot 3
;
; residue  18 HIS rtp HIS1 q  0.0
   170          N     18    HIS      N     70      -0.31
14.0067   ; qtot 2.69
   171          H     18    HIS      H     70       0.31
1.008   ; qtot 3
   172        CH1     18    HIS     CA     71          0
13.019   ; qtot 3
   173        CH2     18    HIS     CB     71          0
14.027   ; qtot 3
   174          C     18    HIS     CG     72          0
12.011   ; qtot 3
   175         NR     18    HIS    ND1     72      -0.05
14.0067   ; qtot 2.95
   176          H     18    HIS    HD1     72       0.31
1.008   ; qtot 3.26
   177          C     18    HIS    CD2     72          0
12.011   ; qtot 3.26
   178         HC     18    HIS    HD2     72       0.14
1.008   ; qtot 3.4
   179          C     18    HIS    CE1     72          0
12.011   ; qtot 3.4
   180         HC     18    HIS    HE1     72       0.14
1.008   ; qtot 3.54
   181         NR     18    HIS    NE2     72      -0.54
14.0067   ; qtot 3
   182          C     18    HIS      C     73       0.45
12.011   ; qtot 3.45
   183          O     18    HIS      O     73      -0.45
15.9994   ; qtot 3
;
; residue  19 THR rtp THR  q  0.0
```

```
   184          N     19    THR      N     74      -0.31
14.0067   ; qtot 2.69
   185          H     19    THR      H     74       0.31
1.008   ; qtot 3
   186        CH1     19    THR     CA     75          0
13.019   ; qtot 3
   187        CH1     19    THR     CB     76      0.266
13.019   ; qtot 3.266
   188         OA     19    THR    OG1     76     -0.674
15.9994   ; qtot 2.592
   189          H     19    THR    HG1     76      0.408
1.008   ; qtot 3
   190        CH3     19    THR    CG2     77          0
15.035   ; qtot 3
   191          C     19    THR      C     78       0.45
12.011   ; qtot 3.45
   192          O     19    THR      O     78      -0.45
15.9994   ; qtot 3
;
; residue  23 GLY rtp GLY  q  0.0
   224          N     23    GLY      N     91      -0.31
14.0067   ; qtot 2.69
   225          H     23    GLY      H     91       0.31
1.008   ; qtot 3
   226        CH2     23    GLY     CA     92          0
14.027   ; qtot 3
   227          C     23    GLY      C     93       0.45
12.011   ; qtot 3.45
   228          O     23    GLY      O     93      -0.45
15.9994   ; qtot 3
;
;residue  30 PRO rtp PRO  q  0.0
   288          N     30    PRO      N    119          0
14.0067   ; qtot 5
   289        CH1     30    PRO     CA    120          0
13.019   ; qtot 5
   290       CH2r     30    PRO     CB    120          0
14.027   ; qtot 5
   291       CH2r     30    PRO     CG    121          0
14.027   ; qtot 5
   292       CH2r     30    PRO     CD    121          0
14.027   ; qtot 5
   293          C     30    PRO      C    122       0.45
12.011   ; qtot 5.45
   294          O     30    PRO      O    122      -0.45
15.9994   ; qtot 5
;
```

```
; residue  31 ASN rtp ASN  q  0.0
   295          N     31    ASN      N    123      -0.31
14.0067   ; qtot 4.69
   296          H     31    ASN      H    123       0.31
1.008   ; qtot 5
   297        CH1     31    ASN     CA    124          0
13.019   ; qtot 5
   298        CH2     31    ASN     CB    124          0
14.027   ; qtot 5
   299          C     31    ASN     CG    125       0.29
12.011   ; qtot 5.29
   300          O     31    ASN    OD1    125      -0.45
15.9994   ; qtot 4.84
   301         NT     31    ASN    ND2    125      -0.72
14.0067   ; qtot 4.12
   302          H     31    ASN   HD21    125       0.44
1.008   ; qtot 4.56
   303          H     31    ASN   HD22    125       0.44
1.008   ; qtot 5
   304          C     31    ASN      C    126       0.45
12.011   ; qtot 5.45
   305          O     31    ASN      O    126      -0.45
15.9994   ; qtot 5
;
;residue  32 LEU rtp LEU  q  0.0
   306          N     32    LEU      N    127      -0.31
14.0067   ; qtot 4.69
   307          H     32    LEU      H    127       0.31
1.008   ; qtot 5
   308        CH1     32    LEU     CA    128          0
13.019   ; qtot 5
   309        CH2     32    LEU     CB    128          0
14.027   ; qtot 5
   310        CH1     32    LEU     CG    129          0
13.019   ; qtot 5
   311        CH3     32    LEU    CD1    129          0
15.035   ; qtot 5
   312        CH3     32    LEU    CD2    129          0
15.035   ; qtot 5
   313          C     32    LEU      C    130       0.45
12.011   ; qtot 5.45
   314          O     32    LEU      O    130      -0.45
15.9994   ; qtot 5
;
; residue  38 ARG rtp ARG  q +1.0
   365          N     38    ARG      N    153      -0.31
14.0067   ; qtot 4.69
```

```
   366          H     38    ARG      H    153        0.31
1.008   ; qtot 5
   367        CH1     38    ARG     CA    154           0
13.019   ; qtot 5
   368        CH2     38    ARG     CB    154           0
14.027   ; qtot 5
   369        CH2     38    ARG     CG    154           0
14.027   ; qtot 5
   370        CH2     38    ARG     CD    155        0.09
14.027   ; qtot 5.09
   371         NE     38    ARG     NE    155       -0.11
14.0067   ; qtot 4.98
   372          H     38    ARG     HE    155        0.24
1.008   ; qtot 5.22
   373          C     38    ARG     CZ    155        0.34
12.011   ; qtot 5.56
   374         NZ     38    ARG    NH1    155       -0.26
14.0067   ; qtot 5.3
   375          H     38    ARG   HH11    155        0.24
1.008   ; qtot 5.54
   376          H     38    ARG   HH12    155        0.24
1.008   ; qtot 5.78
   377         NZ     38    ARG    NH2    155       -0.26
14.0067   ; qtot 5.52
   378          H     38    ARG   HH21    155        0.24
1.008   ; qtot 5.76
   379          H     38    ARG   HH22    155        0.24
1.008   ; qtot 6
   380          C     38    ARG      C    156        0.45
12.011   ; qtot 6.45
   381          O     38    ARG      O    156       -0.45
15.9994   ; qtot 6
;
; residue  48 TYR rtp TYR  q  0.0
   465          N     48    TYR      N    197       -0.31
14.0067   ; qtot 6.69
   466          H     48    TYR      H    197        0.31
1.008   ; qtot 7
   467        CH1     48    TYR     CA    198           0
13.019   ; qtot 7
   468        CH2     48    TYR     CB    198           0
14.027   ; qtot 7
   469          C     48    TYR     CG    198           0
12.011   ; qtot 7
   470          C     48    TYR    CD1    199       -0.14
12.011   ; qtot 6.86
```

```
   471          HC     48    TYR    HD1    199        0.14
1.008   ; qtot 7
   472           C     48    TYR    CD2    200       -0.14
12.011   ; qtot 6.86
   473          HC     48    TYR    HD2    200        0.14
1.008   ; qtot 7
   474           C     48    TYR    CE1    201       -0.14
12.011   ; qtot 6.86
   475          HC     48    TYR    HE1    201        0.14
1.008   ; qtot 7
   476           C     48    TYR    CE2    202       -0.14
12.011   ; qtot 6.86
   477          HC     48    TYR    HE2    202        0.14
1.008   ; qtot 7
   478           C     48    TYR     CZ    203       0.203
12.011   ; qtot 7.203
   479          OA     48    TYR     OH    203      -0.611
15.9994   ; qtot 6.592
   480           H     48    TYR     HH    203       0.408
1.008   ; qtot 7
   481           C     48    TYR      C    204        0.45
12.011   ; qtot 7.45
   482           O     48    TYR      O    204       -0.45
15.9994   ;
; residue  57 ILE rtp ILE  q  0.0
   560           N     57    ILE      N    238       -0.31
14.0067   ; qtot 7.69
   561           H     57    ILE      H    238        0.31
1.008   ; qtot 8
   562         CH1     57    ILE     CA    239           0
13.019   ; qtot 8
   563         CH1     57    ILE     CB    240           0
13.019   ; qtot 8
   564         CH2     57    ILE    CG1    240           0
14.027   ; qtot 8
   565         CH3     57    ILE    CG2    240           0
15.035   ; qtot 8
   566         CH3     57    ILE     CD    240           0
15.035   ; qtot 8
   567           C     57    ILE      C    241        0.45
12.011   ; qtot 8.45
   568           O     57    ILE      O    241       -0.45
15.9994   ; qtot 8
;
; residue  80 MET rtp MET  q  0.0
   811           N     80    MET      N    348       -0.31
14.0067   ; qtot 7.69
```

```
   812          H     80    MET      H    348       0.31
1.008   ; qtot 8
   813        CH1     80    MET     CA    349          0
13.019   ; qtot 8
   814        CH2     80    MET     CB    349          0
14.027   ; qtot 8
   815        CH2     80    MET     CG    350      0.241
14.027   ; qtot 8.241
   816          S     80    MET     SD    350     -0.482
32.06   ; qtot 7.759
   817        CH3     80    MET     CE    350      0.241
15.035   ; qtot 8
   818          C     80    MET      C    351       0.45
12.011   ; qtot 8.45
   819          O     80    MET      O    351      -0.45
15.9994   ;
; residue 104 GLU rtp GLU  q -2.0
  1063          N    104    GLU      N    455      -0.31
14.0067   ; qtot 10.69
  1064          H    104    GLU      H    455       0.31
1.008   ; qtot 11
  1065        CH1    104    GLU     CA    456          0
13.019   ; qtot 11
  1066        CH2    104    GLU     CB    456          0
14.027   ; qtot 11
  1067        CH2    104    GLU     CG    456          0
14.027   ; qtot 11
  1068          C    104    GLU     CD    457       0.27
12.011   ; qtot 11.27
  1069         OM    104    GLU    OE1    457     -0.635
15.9994   ; qtot 10.63
  1070         OM    104    GLU    OE2    457     -0.635
15.9994   ; qtot 10
  1071          C    104    GLU      C    458       0.27
12.011   ; qtot 10.27
  1072         OM    104    GLU     O1    458     -0.635
15.9994   ; qtot 9.635
  1073         OM    104    GLU     O2    458     -0.635
15.9994   ; qtot 9
; residue 105 HEM rtp HEME q -2.0
  1074         FE    105    HEM     FE    459        0.4
55.847   ; qtot 9.4
  1075         NR    105    HEM     NA    459       -0.1
14.0067   ; qtot 9.3
  1076         NR    105    HEM     NB    459       -0.1
14.0067   ; qtot 9.2
```

```
  1077         NR    105    HEM     NC    459       -0.1
14.0067   ; qtot 9.1
  1078         NR    105    HEM     ND    459       -0.1
14.0067   ; qtot 9
  1079          C    105    HEM    CHA    460       -0.1
12.011   ; qtot 8.9
  1080         HC    105    HEM    HHA    460        0.1
1.008   ; qtot 9
  1081          C    105    HEM    C1A    461          0
12.011   ; qtot 9
  1082          C    105    HEM    C2A    461          0
12.011   ; qtot 9
  1083          C    105    HEM    C3A    461          0
12.011   ; qtot 9
  1084          C    105    HEM    C4A    461          0
12.011   ; qtot 9
  1085        CH3    105    HEM    CMA    462          0
15.035   ; qtot 9
  1086        CH2    105    HEM    CAA    463          0
14.027   ; qtot 9
  1087        CH2    105    HEM    CBA    463          0
14.027   ; qtot 9
  1088          C    105    HEM    CGA    464       0.27
12.011   ; qtot 9.27
  1089         OM    105    HEM    O1A    464     -0.635
15.9994   ; qtot 8.635
  1090         OM    105    HEM    O2A    464     -0.635
15.9994   ; qtot 8
  1091          C    105    HEM    CHB    465       -0.1
12.011   ; qtot 7.9
  1092         HC    105    HEM    HHB    465        0.1
1.008   ; qtot 8
  1093          C    105    HEM    C1B    466          0
12.011   ; qtot 8
  1094          C    105    HEM    C2B    466          0
12.011   ; qtot 8
  1095          C    105    HEM    C3B    466          0
12.011   ; qtot 8
  1096          C    105    HEM    C4B    466          0
12.011   ; qtot 8
  1097        CH3    105    HEM    CMB    467          0
15.035   ; qtot 8
  1098        CR1    105    HEM    CAB    468          0
13.019   ; qtot 8
  1099        CH3    105    HEM    CBB    468          0
15.035   ; qtot 8
```

```
  1100          C    105    HEM    CHC    469       -0.1
12.011   ; qtot 7.9
  1101         HC    105    HEM    HHC    469        0.1
1.008   ; qtot 8
  1102          C    105    HEM    C1C    470          0
12.011   ; qtot 8
  1103          C    105    HEM    C2C    470          0
12.011   ; qtot 8
  1104          C    105    HEM    C3C    470          0
12.011   ; qtot 8
  1105          C    105    HEM    C4C    470          0
12.011   ; qtot 8
  1106        CH3    105    HEM    CMC    471          0
15.035   ; qtot 8
  1107        CR1    105    HEM    CAC    472          0
13.019   ; qtot 8
  1108        CH3    105    HEM    CBC    472          0
15.035   ; qtot 8
  1109          C    105    HEM    CHD    473       -0.1
12.011   ; qtot 7.9
  1110         HC    105    HEM    HHD    473        0.1
1.008   ; qtot 8
  1111          C    105    HEM    C1D    474          0
12.011   ; qtot 8
  1112          C    105    HEM    C2D    474          0
12.011   ; qtot 8
  1113          C    105    HEM    C3D    474          0
12.011   ; qtot 8
  1114          C    105    HEM    C4D    474          0
12.011   ; qtot 8
  1115        CH3    105    HEM    CMD    475          0
15.035   ; qtot 8
  1116        CH2    105    HEM    CAD    476          0
14.027   ; qtot 8
  1117        CH2    105    HEM    CBD    476          0
14.027   ; qtot 8
  1118          C    105    HEM    CGD    477       0.27
12.011   ; qtot 8.27
  1119         OM    105    HEM    O1D    477     -0.635
15.9994   ; qtot 7.635
  1120         OM    105    HEM    O2D    477     -0.635
15.9994   ; qtot 7
;
;ADDED BY DEV
;[bonds]
;
#define gb_53     0.232   0.361e+6
```

```
; S(MET-80)  -   FE(HEME)  (REF: J.Phys.Chem. B. Vol 106, No
21, 2002  literature mentioned that S(Met-80)-Fe , Kb = 250
Kcal/mol for 0.232 nm bond length used for charmm27.
; converted to gromacs value using following formula
; Kb(gro) = (250 x 4.18/((0.232)^4))
;bond is less stronger than N(HIS-18)-FE ) ; sd of this bond
fluactuation does not
;show significant change bond length.
;
;
#define gb_54     0.204          0.600e+06
; NE1(HIS-18)  -  Fe(HEME) (REF:gb_34)
;
;[angles]
;
;Added by DEV from VMD angle estimation on 1HRC.pdb.
;
#define ga_55   175.05      56600.00
; NR(HIS-18)  - Fe  -  S(Met-80)   175.05 (REF:CHARMM27.FF
axial bond angle of NR2-FE-CO is taken Ka = 418.4 KJ/mol/rad^2
with angle theta = 180 )
; Ka(gro) =
Ka(charmm)x(theta,rad,charmm)^2/(sin^2(theta,deg,gro)x(theta,ra
d,gro)^2)
;
#define ga_56    111   390.00
; CH3 -  S  - Fe(HEME)   113.6( REF: CHARMM27.FF angle CS-SS-
FE,ka=418,theta=100.6 )
; CH2 - S - Fe(HEME)     108.32
;Ka(gro) =
Ka(charmm)x(theta,rad,charmm)^2/(sin^2(theta,deg,gro)x(theta,ra
d,gro)^2)
;
;
[ bonds ]
;  ai    aj funct            c0            c1            c2
c3
    1     2     2    gb_2
    1     3     2    gb_2
    1     4     2    gb_2
    1     5     2    gb_21
    5     6     2    gb_27
    6     7     2    gb_5
    ;     ;     ;    ;
    ;     ;     ;    ;
  142  1098     2    gb_31 ; *
    ;     ;     ;    ;
```

```
    ;     ;     ;    ;
  167  1107     2    gb_32 ; *
    ;     ;     ;    ;
    ;     ;     ;    ;
  181  1074     2    gb_54 ; *
    ;     ;     ;    ;
    ;     ;     ;    ;
  816  1074     2    gb_53 ; *
    ;     ;     ;    ;
    ;     ;     ;    ;
 1117  1118     2    gb_27
 1118  1119     2    gb_6

[ pairs ]
;  ai    aj funct            c0            c1            c2
c3
    1     7     1
    1     8     1
    3     6     1
    ;     ;     ;
    ;     ;     ;
 1105  1108     1
 1112  1117     1
 1116  1120     1

[ angles ]
;  ai    aj    ak funct            c0            c1
c2              c3
    2     1     3     2    ga_10
    2     1     4     2    ga_10
    2     1     5     2    ga_11
    3     1     4     2    ga_10
    ;     ;     ;     ;    ;
    ;     ;     ;     ;    ;
  141   142  1098     2    ga_4  ; *
    ;     ;     ;     ;    ;
    ;     ;     ;     ;    ;
  163   165   168     2    ga_13
  166   167  1107     2    ga_4  ; *
  165   168   169     2    ga_30
  177   181  1074     2    ga_34 ; *
  179   181  1074     2    ga_34 ; *
    ;     ;     ;     ;    ;
    ;     ;     ;     ;    ;
  815   816  1074     2    ga_56 ; *
  817   816  1074     2    ga_56 ; *
  181  1074   816     2    ga_55 ; *
```

```
  181  1074  1075     2    ga_2  ; *
  181  1074  1076     2    ga_2  ; *
  181  1074  1077     2    ga_2  ; *
  181  1074  1078     2    ga_2  ; *
  816  1074  1075     2    ga_1  ; *
  816  1074  1076     2    ga_1  ; *
  816  1074  1077     2    ga_1  ; *
  816  1074  1078     2    ga_1  ; *
 1075  1074  1076     2    ga_2
 1075  1074  1077     2    ga_55 ; *
 1075  1074  1078     2    ga_2
 1076  1074  1077     2    ga_2
 1076  1074  1078     2    ga_55 ; *
 1077  1074  1078     2    ga_2
    ;     ;     ;     ;    ;
    ;     ;     ;     ;    ;
 1095  1096  1100     2    ga_38
  142  1098  1095     2    ga_16 ; *
  142  1098  1099     2    ga_16 ; *
 1095  1098  1099     2    ga_13 ; *
    ;     ;     ;     ;    ;
    ;     ;     ;     ;    ;
  167  1107  1104     2    ga_16 ; *
  167  1107  1108     2    ga_16 ; *
 1104  1107  1108     2    ga_13 ; *
    ;     ;     ;     ;    ;
    ;     ;     ;     ;    ;
 1117  1118  1120     2    ga_22
 1119  1118  1120     2    ga_38

[ dihedrals ]
;  ai    aj    ak    al funct            c0            c1
c2            c3            c4            c5
    2     1     5     6     1    gd_29
    1     5     6     8     1    gd_40
    5     6     8    10     1    gd_14
    ;     ;     ;     ;     ;    ;
    ;     ;     ;     ;     ;    ;
  138   140   143   145     1    gd_40
  140   141   142  1098     1    gd_26 ; *
  141   142  1098  1095     1    gd_26 ; *
  140   143   145   147     1    gd_14
    ;     ;     ;     ;     ;    ;
    ;     ;     ;     ;     ;    ;
  161   163   165   168     1    gd_39
  165   166   167  1107     1    gd_26 ; *
  166   167  1107  1104     1    gd_26 ; *
```

```
  165   168   170   172     1    gd_14
    ;     ;     ;     ;     ;    ;
    ;     ;     ;     ;     ;    ;
  177   181  1074   816     1    gd_26 ; *
  172   182   184   186     1    gd_14
    ;     ;     ;     ;     ;    ;
    ;     ;     ;     ;     ;    ;
  814   815   816   817     1    gd_26
  815   816  1074   181     1    gd_26 ; *
  813   818   820   822     1    gd_14
    ;     ;     ;     ;     ;    ;
    ;     ;     ;     ;     ;    ;
  181  1074  1075  1081     1    gd_38 ; *
  181  1074  1076  1093     1    gd_38 ; *
  181  1074  1077  1102     1    gd_38 ; *
  181  1074  1078  1111     1    gd_38 ; *
    ;     ;     ;     ;     ;    ;
    ;     ;     ;     ;     ;    ;
 1114  1079  1081  1075     1    gd_15
 1081  1079  1114  1078     1    gd_15

[ dihedrals ]
;  ai    aj    ak    al funct            c0            c1
c2            c3
    6     5     8     7     2    gi_1
    8     6    10     9     2    gi_1
    ;     ;     ;     ;     ;    ;
    ;     ;     ;     ;     ;    ;
 1109  1078  1112  1111     2    gi_1
 1111  1078  1114  1113     2    gi_1
;
; Include Position restraint file
#ifdef POSRES
#include "posre_r_cytc.itp"
#include "posre_meoh.itp"
#endif
;
; Include water topology
#include "gromos53a6.ff/spce.itp"
;
#ifdef POSRES_WATER
; Position restraint for each water oxygen
[ position_restraints ]
;  i funct       fcx        fcy        fcz
   1    1       1000       1000       1000
#endif
;
```

```
; Include topology for ions
#include "gromos53a6.ff/ions.itp"
;
[ system ]
; Name
CYTOCHROME-C in 20%MeOH in Water

[ molecules ]
; Compound        #mols
  r_cytc              1
  MeOH              718
  SOL              5222
  CL                  7
*Only in hh-CytC
================================================================
```

(this page intentionally left blank)