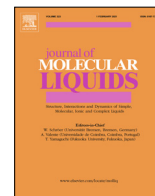




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Understanding the role of water on temperature-dependent structural modifications of SARS CoV-2 main protease binding sites



Pushyaraga P. Venugopal, Omkar Singh, Debashree Chakraborty*

Biophysical and Computational Chemistry Laboratory, Department of Chemistry, National Institute of Technology Karnataka, India

ARTICLE INFO

Article history:

Received 16 April 2022

Revised 1 July 2022

Accepted 14 July 2022

Available online 20 July 2022

Keywords:

SARS-CoV-2 main protease

Role of water

Temperature

Molecular dynamics simulation

Binding pockets

ABSTRACT

Thermally stable and labile proteases are found in microorganisms. Protease mediates the cleavage of polyproteins in the virus replication and transcription process. 6 μ s MD simulations were performed for monomer/dimer SARS CoV-2 main protease system in both SPC/E and mTIP3P water model to analyse the temperature-dependent behaviour of the protein. It is found that maximum conformational changes are observed at 348 K which is near the melting temperature. Network distribution of evolved conformations shows an increase in the number of communities with the rise in the temperature. The global conformation of the protein was found to be intact whereas a local conformational space evolved due to thermal fluctuations. The global conformational change in the free energy $\Delta\Delta G$ value for the monomer and the dimer between 278 K and 383 K is found to be 2.51 and 2.10 kJ/mol respectively. A detailed analysis was carried out on the effect of water on the temperature-dependent structural modifications of four binding pockets of SARS CoV-2 main protease namely, catalytic dyad, substrate-binding site, dimerization site and allosteric site. It is found that the water structure around the binding sites is altered with temperature. The water around the dimer sites is more ordered than the monomer sites regardless of the rise in temperature due to structural rigidity. The energy expense of binding the small molecules at substrate binding is less compared to the allosteric site. The water-water hydrogen bond lifetime is found to be more near the cavity of His41. Also, it is observed that mTIP3P water molecules have a similar effect to that of SPC/E water molecules on the main protease.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Statistics based on environmental and epidemiological studies signify the temperature dependence of COVID-19 transmission across the globe. Experimental reports suggest the inactivation of the SARS-CoV-2 virus at higher temperatures while elevated infectivity rate at lower temperatures [1]. The temperature-dependent activity of corona virus at the molecular level is still unclear and can be studied effectively using computational tools. Efforts have been done by the research communities to develop efficient drugs and vaccines against SARS CoV-2 virus using *in silico* methods [2–4]. Generally in viruses, the protease enzyme is responsible for the viral protein maturation after the translation process in the host cell. Coronavirus main protease (M^{Pro}) mediates the cleavage of polyproteins in virus replication and transcription process [5]. The inhibition of protease activity reduces the mature virus particles in the host cell; therefore, the protease enzymes are con-

sidered important drug targets. The crystal structure reveals that the structure of protease consists of two β -barrel domains (domain I and domain II) and an α -helical bundle (domain III) (Fig. 1). The substrate-binding site and the Cys-His catalytic dyad are found in the cleft between domain I and domain II. Domain III regulates the dimerization of protease essential for M^{Pro} activity [6,7]. Dimerization involves the interaction of N-terminal residues of two protomers with the Glu166 of another protomer to hold the pocket of substrate binding site [8].

Proteins are sensitive and behave differently in lower and higher temperature conditions. The temperature effect can either alter or break the non-covalent interactions thereby causing changes in the three-dimensional folding of the protein chain. Both, thermally stable, as well as labile proteases, are found in microorganisms, plants and animals. The human coronaviruses including SARS-CoV-2 showed a higher average mean lifetime in 10–60 °C temperature range and decontamination above 70 °C [9]. Studies showed that the differences in an external environment like temperature disables the outermost structural proteins, spike protein, in the coronavirus to bind ACE2 (Angiotensin Converting Enzyme 2) [10]. It is reported that the SARS-CoV-2 RBD (Receptor Binding Domain) affinity decreases at a higher tem-

* Corresponding author at: Biophysical and Computational Chemistry Laboratory, Department of Chemistry, National Institute of Technology Karnataka, Surathkal, Mangalore, Karnataka Pin-Code: 575025, India.

E-mail address: debashree@nitk.edu.in (D. Chakraborty).

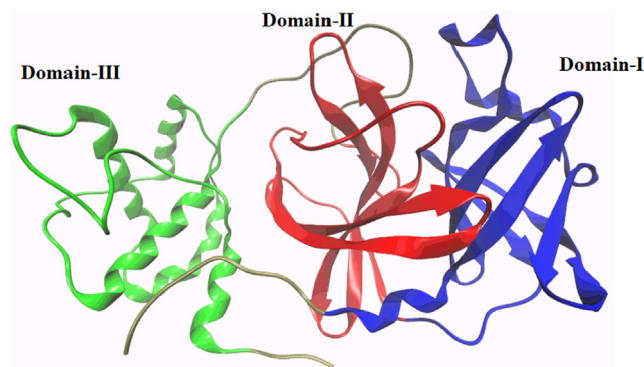


Fig. 1. Representation of SARS CoV-2 main protease (PDB ID: 6Y2E) showing Domain I (blue), Domain II (red) and Domain III (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

perature above 40 °C from both experimental and simulation studies [10–12]. As for the main protease enzyme, crystallographic studies have been carried out in cryogenic conditions, which bias enzyme conformations. Experimental evidence suggests that the protease enzyme exists in a monomer-dimer equilibrium, in which the dimer form is active whereas the monomer form is inactive [13–15]. This equilibrium can be affected by temperature and small molecule inhibitors [16]. Theoretically, Chen et. al proposed that only a protomer is active in SARS 3C-like protease dimer using MD simulations [17]. Since SARS CoV-2 belongs to similar family and genus as SARS CoV, it is essential to understand the mechanism of protease action to develop efficient small molecule inhibitors. Many researches on drug discovery focussed more on SARS CoV-2 main protease as an excellent target [2,18–21]. So it will be interesting to study the dynamic behaviour of both monomer and dimer forms as well as the temperature/ water induced changes in the system. Recently, Kordzadeh et. al showed the temperature and pH-dependent conformational fluctuations on whole SARS CoV-2 main protease using molecular dynamics simulations [22]. This study ignores the structural and dynamic changes of protein and water structure at various binding sites of protease. Even though high-resolution crystal structure of SARS-CoV-2 M^{Pro} was recorded [23], the molecular level mechanism and the solvent effect on M^{Pro} folding, stability and biological activity have not been reported for the wide temperature range.

In this work, we study the temperature effect on the structural modifications of SARS CoV-2 main protease, both, monomer and dimer using atomistic molecular dynamic simulations at four different temperatures. This allows us to understand the impact of temperature on the active dimer and inactive monomer main protease. Specifically, the role of amino acid residues in substrate binding pocket, catalytic dyad, dimerization site and allosteric sites in monomer and dimerized protease were determined using various geometric criteria of non-bonded interactions. The dynamic nature of water molecules at the interface of ligand binding sites is crucial for maintaining the activity of the protein. The interfacial water molecules may favour or disfavour the energetics of binding sites which affect the binding of chemical moieties [24–26]. Generally, this fact is overlooked while studying the degradation process of protein. Here, the effect of water molecules was studied using radial distribution functions, orientational tetrahedral order parameter, free energy $\Delta G(r)$ and hydrogen bond dynamics between key residues at the above-mentioned sites and water molecules. Dihedral analysis at these sites determines the transition between the secondary structure elements in the protein [27]. Principal Component Analysis (PCA) characterizes the essential dynamics which govern the amino acid residue fluctuations at the pocket due to temperature change. Further, the structural evo-

lution and the stability of the whole monomer and dimer during simulation are interpreted using network analysis based on the root mean square deviation for the C α atom in the protein. The global free energy of folding was determined based on the changes in the percentage of α -helices and β -sheets for each temperature. Additionally, the results of simulations were compared using two water models: SPC/E model and modified TIP3P model.

2. Materials and methods

The model of SARS COV-2 main protease (33.83 kDa) was retrieved from Protein Data Bank (PDB ID: 6Y2E, resolution: 1.75 Å) [6]. The enzyme model is devoid of small molecules and consists of 306 amino acid residues. The protein structure was solvated in a cubic box with distance between box boundary and protein structure as 1.0 nm in all directions using SPC/E water model [28] and the system is neutralized with 4 Na⁺ ions. The energy minimization of the system was done by the steepest descent algorithm with a maximum of 50,000 steps until a convergence tolerance of 1000 kJmol⁻¹nm⁻¹. All the molecular dynamics simulations were performed by GROMACS 2018.4 [29] using CHARMM27 force field for proteins [30,31]. A 10 ns NVT equilibration was performed using velocity-rescale temperature coupling method [32] with a time step of 2 fs. Next, a 10 ns NPT equilibration was carried out with a time step of 2 fs at 1 atm using Nose-Hoover thermostat [33] and Parrinello-Rahman pressure coupling scheme [34] respectively. The long-range electrostatic interactions were calculated by particle mesh Ewald method [35] with a cut-off 1.2 nm and Fourier spacing of 0.16 nm. The short-range van der Waals cut-off was fixed to 1.2 nm. LINCS constraints [36] were used to restrain the bond involving hydrogen atoms. The structural and dynamic behaviour of the protein were analyzed for a wide temperature range, from 278 K to 383 K. Finally, a production run of 200 ns was performed until the convergence of RMSD at lower temperature. The trajectories were saved at every 10 ps and GROMACS tools were used for analysis. The simulations were replicated twice (a total of 8 monomer simulations) to check the reproducibility of the results and error bars are given wherever necessary. The model retrieved from PDB is the monomer form of protease. However, the global stoichiometry of the protease is homodimer [6,37]. Therefore, for the dimer simulations, protein-protein docking was carried out using HDock server [38]. The structure retrieved from PDB (ID: 6Y2E) is considered as the target and the average structure obtained from the MD simulation of monomer at each temperature is considered as the substrate for protein-protein docking. The docking energy score obtained between the ligand and the target protease is -555.86 kcal/mol. The geometry of dimer obtained from docking is chosen as the initial configuration for the dimer simulations. We have done a total of 8 dimer simulations. Further, a set of 200 ns simulations (4 simulations each for monomer and dimer) were performed using modified TIP3P water model (m-TIP3P) to determine the effect of water model on the temperature dependent behaviour of main protease. Additional, five 200 ns simulations each for 288 K, 298 K, 308 K, 318 K, 338 K and 358 K were performed for monomer protease to determine the entropy and enthalpy contributions to the free energy. A total of 6 μ s simulations were performed.

3. Results and discussion

3.1. Structural changes in protein conformation

3.1.1. Temperature effect on M^{Pro} enzyme

The temperature-dependent stability of monomer and dimer were studied based on structural parameters and solvent effect.

200 ns molecular dynamics simulations were performed using SPC/E and m-TIP3P water model at 4 different temperatures, 278 K (cold denaturation), 310 K (physiological), 348 K (higher temperature reported for enzyme activity [39]) and 383 K (extreme thermophilic conditions [40]), to obtain the insights into the conformational evolution and stability of SARS CoV-2 main protease enzyme. The dimer is composed of two monomer proteases, which are denoted as chain A and chain B. Both the chains are showing similar results (SI, Section: Dimer results). Therefore, we selected one chain for comparing the structure and dynamics of the main protease dimer form with the monomer. For the dimer, the monomer-monomer interaction energy is calculated between the two chains of the dimer. At 310 K, the interaction energy is found to be stable than other temperatures.

The global structure of SARS CoV-2 main protease protomer and dimer is found to be similar despite the increase in temperature and difference in the water model as shown in **Figure S1**. Per residue secondary structure analysis of protease using do_dssp tool [41] at different temperatures shows no considerable changes in the average occupancy of total secondary structure components in the protein (**Table S1**). The percentage of secondary structure element obtained by do_dssp tool can be correlated with the calculated percentage from X-Ray diffraction technique (α -helix: 0.31%, β -sheet: 0.27%, unstructured region: 0.42%). However, it can be observed that β -bridges and α -helices are formed with the increase in temperature whereas native α -helices are destroyed in the structure. The secondary structure element plots reveal the degradation of the enzyme with the rise in temperature which confirms the change of the M^{Pro} conformational space (**Figure S2**). The free energy of folding for monomer and dimer-chain A were determined by the changes in α -helix and β -sheet elements.

$$\Delta G_{\text{fold}} = -k_B T \ln\left(\frac{F_N}{1 - F_N}\right) \quad (1)$$

where F_N is the fraction of protein folding [42]. The fraction of folding (F_N) is determined by the fraction of residues with α -helix, β -sheet elements and turns present in the protein as a function of time. It is found that the monomer and the dimer-chain A conformations at 310 K (human body temperature) are more stable than other temperatures with ΔG_{fold} values of -3.20 ± 0.25 and -3.43 ± 0.13 kJ/mol (**Figure S3**). The probability distribution of C_{α} -atom RMSD for both monomer and chain A of the dimer (dimer-chain A) in SPC/E water model is shown in **Fig. 2(a)** and the corresponding time profile of C_{α} RMSD is given in **Figure S4**. It is found that at 278 K and 310 K, the protein has a single probable structure for both dimer-chain A and monomer. At 348 K, the conformation of protease gradually evolves during simulation due to thermal fluctuations. The system of monomer has well equilibrated within 25 ns of the simulation time with a value of 0.187 ± 0.02 and 0.200 ± 0.02 nm at 278 K and 310 K respectively. At temperatures 348 K and 383 K, marginally higher RMSD values are observed (**Figure S4**). The dimer-chain A has achieved the rigidity after dimerization is evident from the fewer RMSD fluctuations with a value of 0.131 ± 0.01 and 0.137 ± 0.01 nm at 278 K and 310 K respectively. It can be observed that the rigidity is lost with the rise in temperature as seen in 348 K and 383 K, which is evident from the tail region. The compactness of the protease monomer decreases with rising in temperature due to the thermal fluctuations at the α -helical region of the protein [43]. At 383 K, due to random motion of the protein molecules, frequent contacts are made and broken at different stages which are not stable resulting to broader distribution. Similar trend is observed for dimer-chain A (**Fig. 2(b)**). However it is observed that the monomer is more compact than the dimer at 383 K. Further, this

observation is supported by the correlation of inter-residue distance (**Figure S5**). The correlation map shows the contact formation (orange region) and contact rupture (blue region) between amino acid residues. At 348 K, the non-bonded interactions were broken compared to the lower temperature. At 383 K, contact formation is more than the contact rupture [44]. It is evident from the RMSD and the rGyr values that the dimer-chain A has fewer fluctuations than the single-chain protease at lower temperatures. At 348 K and 383 K, both the systems behave differently compared to lower temperature. SASA provides the surface area of protein exposed to the solvent molecule. A change in SASA value indicates a change in the tertiary conformations of the protein (**Fig. 2(c)**). The main protease has the lowest SASA value at 278 K and 310 K indicates that the residues are less exposed to the environment. This also suggests that the protein undergo considerable thermal degradations at 348 K and 383 K in comparison to lower temperature. The SASA value for monomer is higher at 348 K suggesting higher solvation around the protein due to the degradation of protein. This anomaly is observed because 348 K is the melting temperature for the main protease. The dimer-chain A showed a similar SASA value as the monomer at 278 K, 310 K and 383 K which suggests that the interaction of two monomers does not affect the solvent accessibility of monomer units. At 348 K, the single peak SASA value for dimer-chain A is observed which is quite different from the monomeric form. From secondary structure analysis, it is found that at 348 K, monomer undergoes thermal fluctuations leading to the degradation of α -helix and β -sheet at domain I and domain II respectively (**Figure S6**). However in case of dimer, the interaction of two protease chains protects the secondary structure at domain II and domain III. Therefore, at 348 K, the solvent exposure is less in dimer as compared to monomer unit. Among the dimer, there is an increase in the SASA value with rise in temperature. Hydrogen bond interactions between protein-protein and protein-water are considered an important factor for the stabilization and activity of protease enzyme. The number of hydrogen bonds was calculated using gmx_hbond tool. **Fig. 2(d)** shows the average number of hydrogen bonds formed between protein/protein and protein/water. It is found that there is no significant decrease in number of intra-protein hydrogen bonds when the temperature rises from 288 K to 310 K while there is a significant reduction in the average number of hydrogen bond when the temperature rises to 348 K and 383 K. Both monomer and dimer-chain A show a decrease in the number of hydrogen bonds with temperature change. Similarly, protein-water hydrogen bonding showed a gradual decrease with the temperature rise. The number of water molecules decreases with the temperature inside the binding cavity which suggests the expansion of protein structure in both monomer and dimer-chain A. However, at higher temperatures, a tail region is observed as the water molecules can easily move in and out due to the cavity expansion. The number of water molecules in the substrate-binding site is shown in **Fig. 2(e)**. The isotropic temperature factor (B-factor) was calculated from the root mean square fluctuation (RMSF) values of the protease backbone to measure the mobility of each residue from the initial position during the simulation. The B-factor is directly related to RMSF by the following formula [45].

$$B - \text{factor} = \frac{8}{3} \pi^2 \text{RMSF}^2 \quad (2)$$

Apart from the two terminal ends of the protein, the highest fluctuations are observed for amino acid residues in the range Arg40- Glu55 and Asn180- Ile200 at higher temperatures (**Fig. 2(f)**). It is observed that from 348 K to 383 K, the amino acid residues from Lys100- Phe150 show higher fluctuations compared to the lower temperature. It is believed that residues in the regions

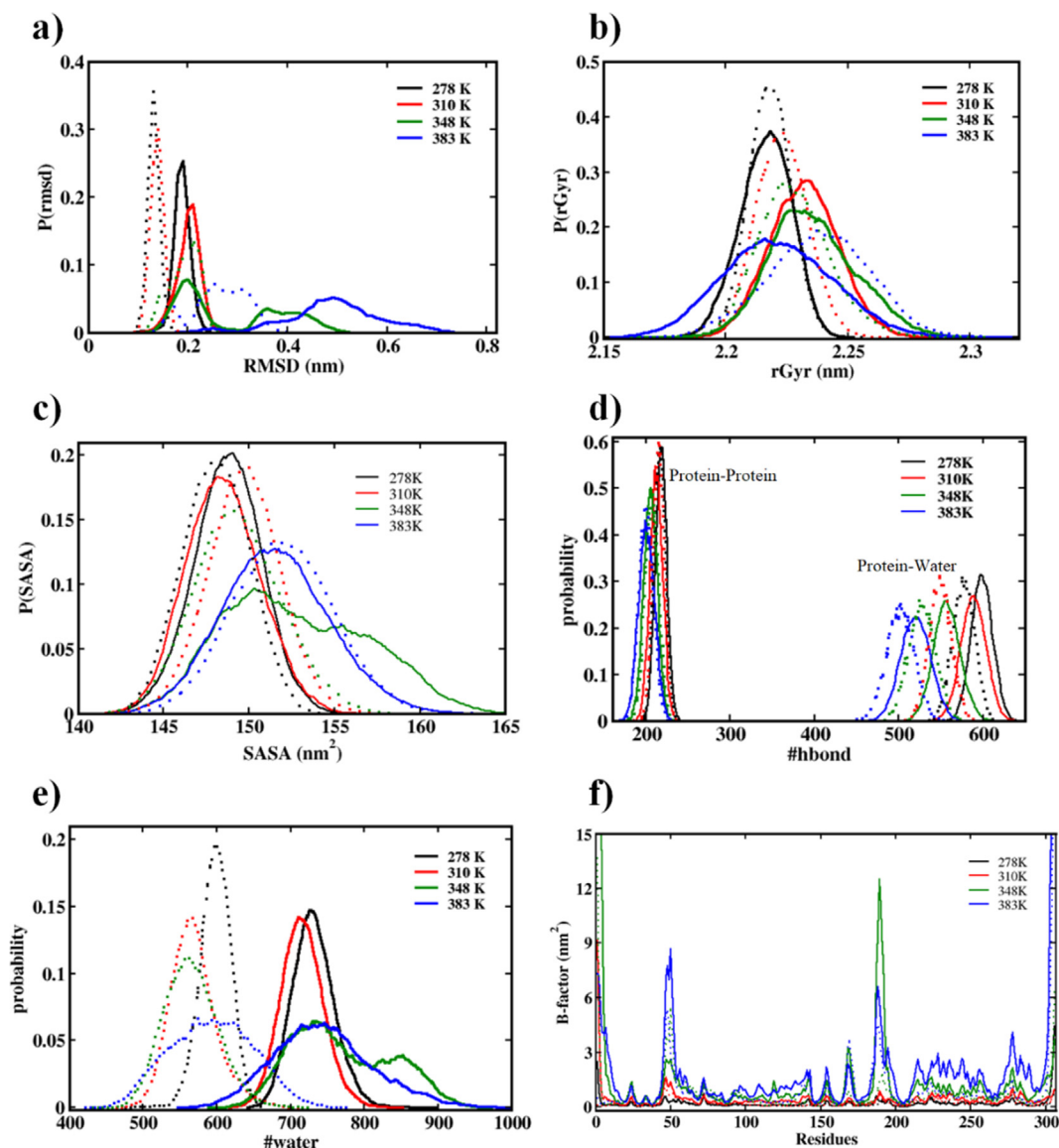


Fig. 2. Probability distribution of (a) C_{α} -RMSD (b) radius of gyration of protein (c) SASA (d) Number of M^{Pro} - M^{Pro} / M^{Pro} -water hydrogen bonds of protein (e) number of water molecules in the substrate binding site during 200 ns simulation and (f) Calculated B-factor of C_{α} atoms with amino acid residues for M^{Pro} . The solid and dotted lines represent monomer and dimer-Chain A respectively in SPC/E water model.

Arg40- Glu55 and Lys100- Phe150 contribute to the catalytic activity of the protease [46]. At higher temperatures, the fluctuations are found to be higher in the above-mentioned regions for single chain and dimerized main protease.

3.1.2. Temperature effect on various binding sites

Even though the structural analyses such as RMSD, rGyr, RMSF and SASA give information about the global changes in the protein structure (Figure S4, S7-S9), it may ignore the temperature-induced conformational changes at the ligand-binding sites and dimerization site. The structural analyses were performed selectively for important amino acid residues as given in Table 1. The catalytic dyad and the substrate-binding site are located between the cleft of domain I and domain II while the dimerization site and allosteric site is regulated mostly by domain III [7]. Domain I consists of 10–99 residues which can be regarded as ($\alpha + \beta$) regions, domain II consists of 100–182 residues which are β -sheet regions and domain III consists of 198–303 residues which are α -helix regions.

Table 1

Important amino acid residues taken for analysis and their biological function.

Residues	Function/ Role in M^{Pro}
His41 , Cys145	Catalytic dyad [47]
His41, Met49, Gly143, Ser144, His163, His164, Met165, Glu166 , Leu167, Asp187, Arg188, Gln189, Thr190, Ala191, Gln192	Substrate binding site [7]
Arg4, Ser10, Gly11, Glu14, Asn28, Cys117 , Ser139, Phe140, Ser147, Glu290, Arg298	Dimerization site [48]
Met6, Phe8, Gln127, Ser139, Asp295, Arg298 , Gln299	Allosteric site [49]

These structural properties for the amino acid residues present at substrate-binding site, dimerization site and allosteric site were determined for different temperatures rather than the whole protein for a better understanding of conformational changes occurring at the binding pockets. The basic structural analyses of

binding site amino acid residues are shown in Fig. 3. The C_{α} RMSD profile of amino acid residues present at various binding sites with respect to the domain (I, II, III) movements in monomer suggests different dynamics of protease binding sites at lower and higher temperatures. The higher RMSD values at 348 K and 383 K indicate the flexible nature of all the three domains in the protein as well as the binding site residues. Domain I and domain III motions are more flexible compared to domain II due to the α -helix regions (Figure S10). The highly flexible nature of domain III at higher temperature may have implications on the dimerization of protease which is crucial for its biological function. Similar behaviour is observed in dimer-chain A for lower temperatures. Unlike the monomer, the single chain in the dimer showed a stable movement at 348 K. At 383 K, the chain became unstable as expected. The rGyr values indicate that with temperature rise the compactness of binding site residues decreases. In case of monomer form, the compactness of binding site residue decreases with rise in temperature except 310 K. The rGyr value at 310 K is due to the stable global structure of protease as evident from ΔG_{fold} value. In dimer form, the average value of rGyr at each temperature is almost sim-

ilar, while the value of tail region is increasing which suggests the loss of rigidity with rise in temperature. This is because most of the residues lie in the α -helical region which is more flexible than the β -sheet element. An increasing trend is observed for the solvent-exposed surface area of the protease single chain and dimer. Exception at 348 K for SASA value is due to the presence of more coiled structure.

3.1.3. Conformational dynamics of the protein

Further, we evaluated the change in the local structures in the conformational space due to the thermal fluctuations by network analysis using RMSD as the criterion. Around 60–80 prominent structures which cover the major conformations of protein were extracted from 200 ns trajectory for each temperature using gromos clustering method [50]. Next, to find out the similarity/connections between two conformations, we calculated the RMSD between these conformations. Two conformations were considered to be connected if the RMSD value between them is less than 2.0 Å [51]. Since we compared the change in the conformation for all the temperatures, we preferred RMSD criteria 2 Å where we get

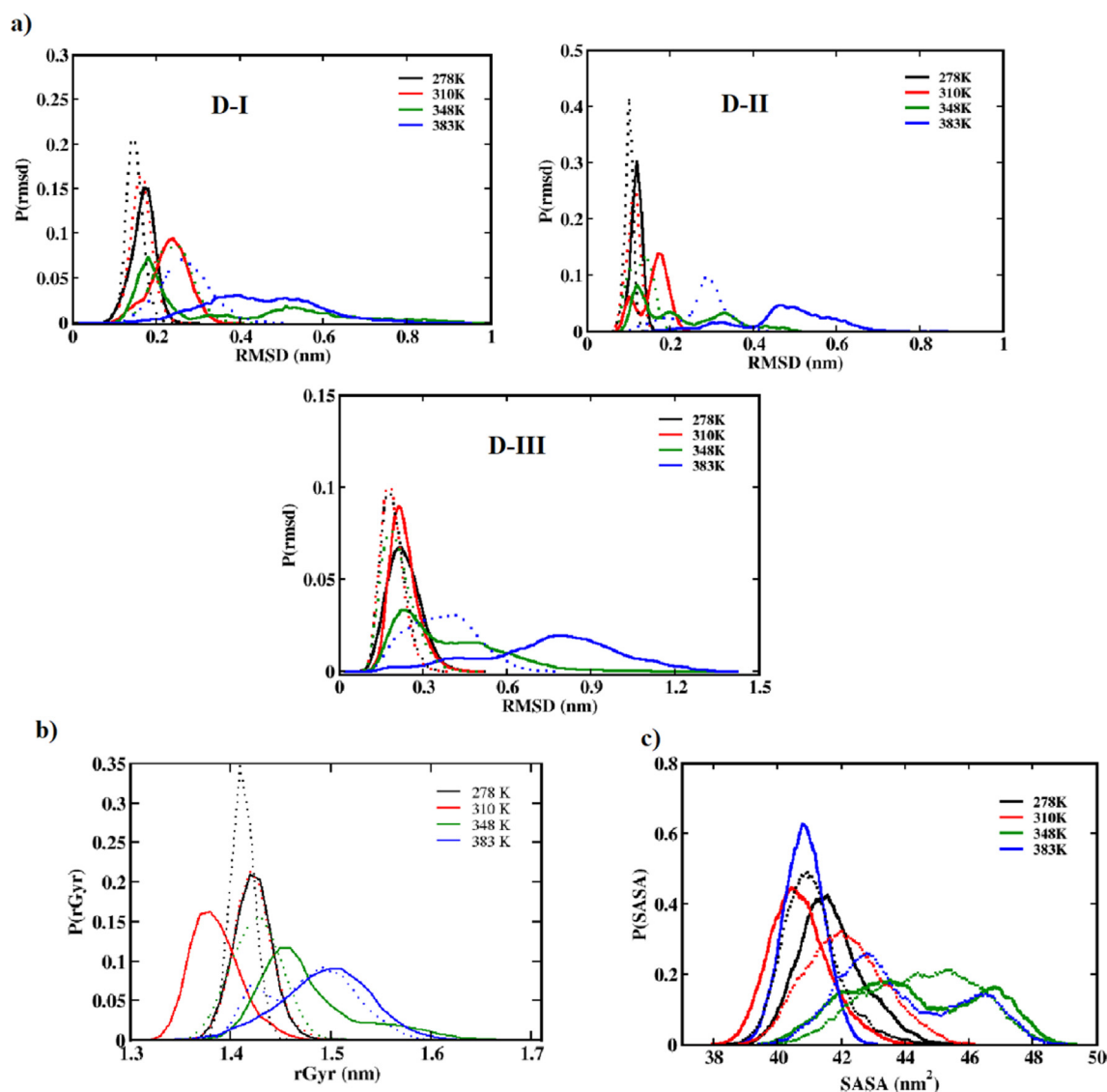


Fig. 3. Probability distribution of (a) C_{α} -RMSD of binding site residues with respect to D-I, D-II, D-III of protein (b) radius of gyration of binding site residues (c) SASA of binding site residues. The binding site residues are considered as per Table 1. The solid and dotted lines represent monomer and dimer-Chain A respectively in SPC/E water model.

a single prominent conformation for 278 K. Conformations obtained from other temperatures are compared with respect to 278 K. The details of the network distribution can be found in our previous work [52]. The network distribution for monomer and dimer-A is given in Fig. 4 with communities represented in different colours. It can be seen that only a single community is present at 278 K and are densely connected to each other. At 310 K, two communities are mainly found where the violet community corresponds to evolved conformations and is densely connected. At higher temperatures, many communities with fewer connections are observed which explains the conformational evolution due to thermal fluctuations. This trend is further confirmed by

RMSD based PMF (potential of mean force) calculation [53] which is given by the formula

$$\Delta G = -k_B T \ln P(\text{rmsd}) \tag{3}$$

where k_B is the Boltzmann constant (8.3145×10^{-3} kJ/mol) and $P(\text{rmsd})$ is the probability of conformations with particular RMSD. This examines the energy change of the evolved conformations as a function of RMSD. PMF profiles of the monomer and dimer-Chain A provides the information about 1-D free energy landscape which depicts the most stable conformation at a particular RMSD value (Fig. 4). It can be seen that at 278 K and 310 K, the monomer has similar free energy minima around -7.5 - -7.7 kJ/mol with

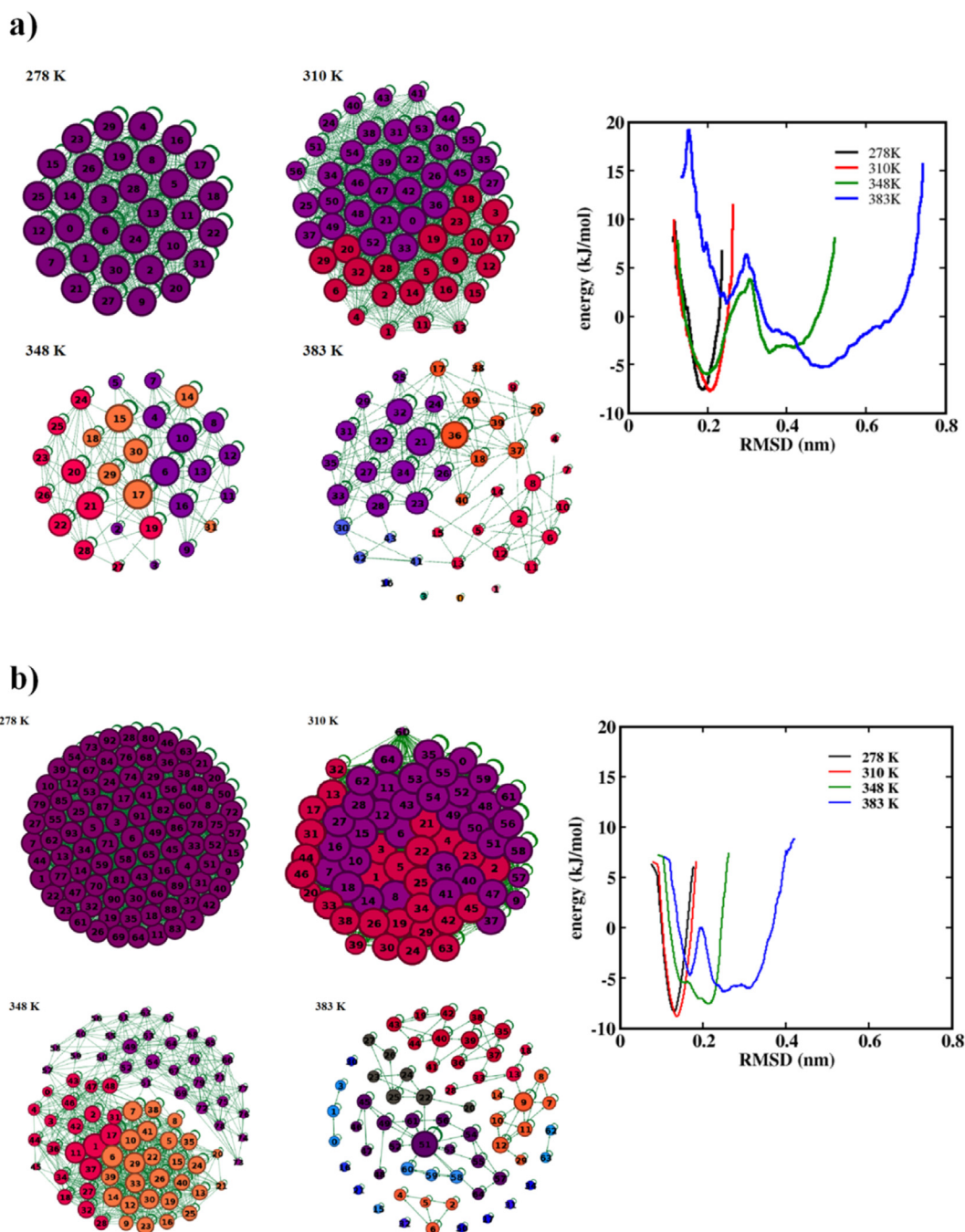


Fig. 4. Network distributions of most prominent conformations of main protease from 200 ns trajectory at different temperatures in SPC/E water model. PMF profile for SARS CoV-2 main protease as a function of RMSD (Right) (a) monomer (b) dimer-Chain A.

different RMSD range at 0.187 and 0.205 nm respectively. For dimer-chain A, a single minimum in the range of -8.4 – -9.0 kJ/mol with RMSD 0.13 and 0.14 nm is observed at 278 K and 310 K respectively. With the increase in the temperature, the free energy curve becomes shallow having multiple global minima. This explains the presence of structures with different energy values. These results correlate well with the network distribution results. The above results suggest that main protease possesses unique structural and dynamic properties at lower and higher temperatures. The energy levels for the monomer are found to be much shallower than the dimer at higher temperatures. The unique structural and dynamic properties observed at different temperatures are further analyzed by PCA analysis and deviation in dihedral angles considering the various binding sites (Fig. 5).

3.1.4. Principal component analysis

PCA is a statistical method to describe the dynamics, which reduces the degrees of freedom of a molecule through decomposition process that spans from largest to smallest observed motions. PCA is based on the correlated fluctuations inside a protein and provides orthogonal eigenvectors and corresponding eigenvalues. The eigenvectors indicate the directions of concerted motions whereas the eigenvalues represent the magnitude along the direction of motion. PCA was performed for backbone C_{α} atoms of key amino acid residues in both the monomer and dimer-Chain A to characterize the essential dynamics which govern the conformational changes throughout 200 ns simulation (Fig. 5). The first eigenvectors correspond to all the possible motions induced by the temperature. The eigenvalue plot indicates the highly flexible nature of the binding site residues. For better representation, the

first two eigenvectors are plotted in 2D projection plot for both monomer and dimer enzyme binding pocket. In the case of the monomer, the scatter plot spread over the phase space more than the dimer. Also, it is observed that at 278 K and 310 K, the conformational changes are not very different but at 348 K and 383 K, the plots deviated to a great extent indicating the temperature-induced flexibility in the protein. The binding site of dimer-Chain A is less flexible due to the structural stability achieved during the dimerization. The PCA analysis shows the flexibility of amino acid residues at the binding pocket of the monomer, agreeing with the structural analysis. Even though the dimer-Chain A is less flexible compared to the monomer, it is found that the flexibility of the binding site increased with temperature rise. The structural flexibility observed at different temperatures is characterized by deviation in dihedral angles considering the various binding sites (Fig. 6).

3.1.5. Scatter plot of dihedral angles at various binding sites

The changes in the backbone of protein residues can be characterized with the help of dihedral angles (ϕ , ψ). The deviations in ϕ , ψ angles provide insight into the changes in the secondary structure elements due to the temperature-dependent fluctuations at these binding sites. The time evolution of dihedral angle deviations for the key residues at four different binding sites is presented in Fig. 6. Each dot in the scatter plots shows the combination of ϕ , ψ angles of residues and the colour indicates the time scale at which the current combination is present. It can be seen that broader distribution of ϕ , ψ angles are observed at 348 K and 383 K except for amino acid residues at the catalytic dyad. At lower temperatures, the catalytic dyad is found to be intact with fewer

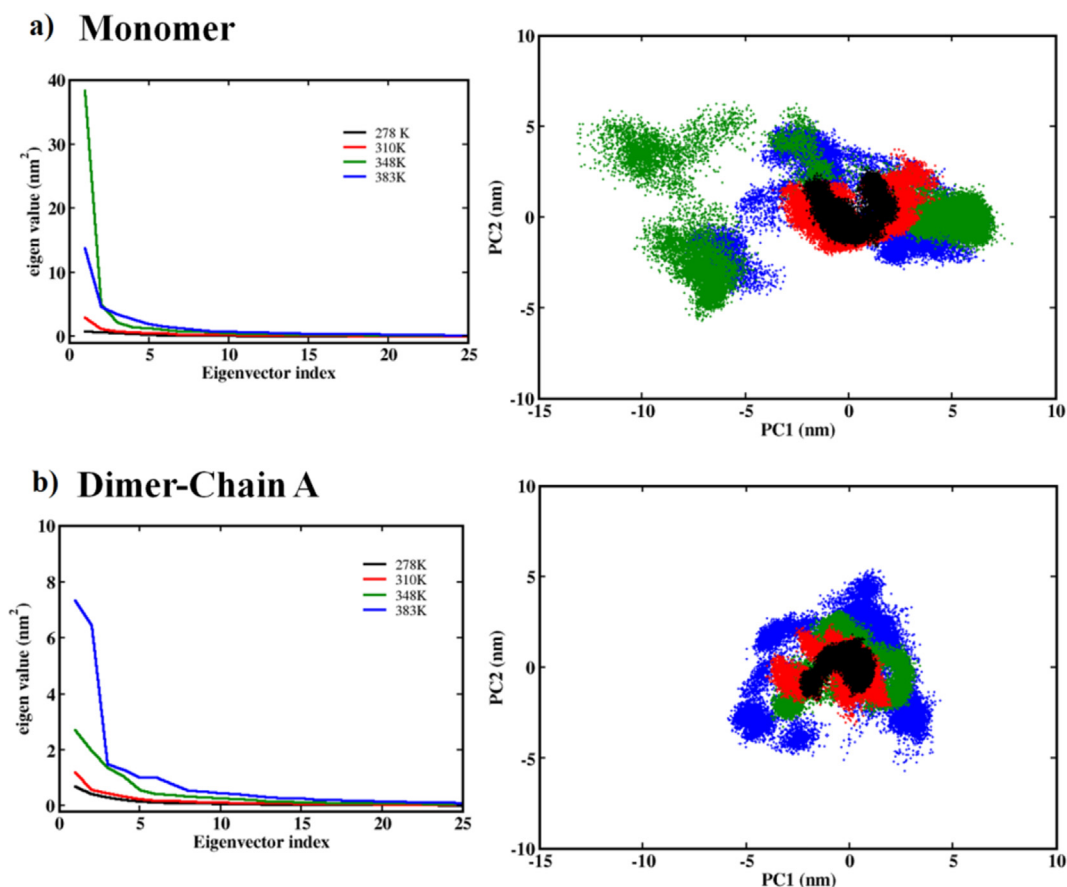


Fig. 5. Plot of eigenvalue vs first 25 eigenvector index derived from PCA over 200 ns MD trajectory for (a) M^{Pro} monomer and (b) M^{Pro} dimer-Chain A. 2-D projection of first two principal motions (PC1 and PC2) of main protease (Right).

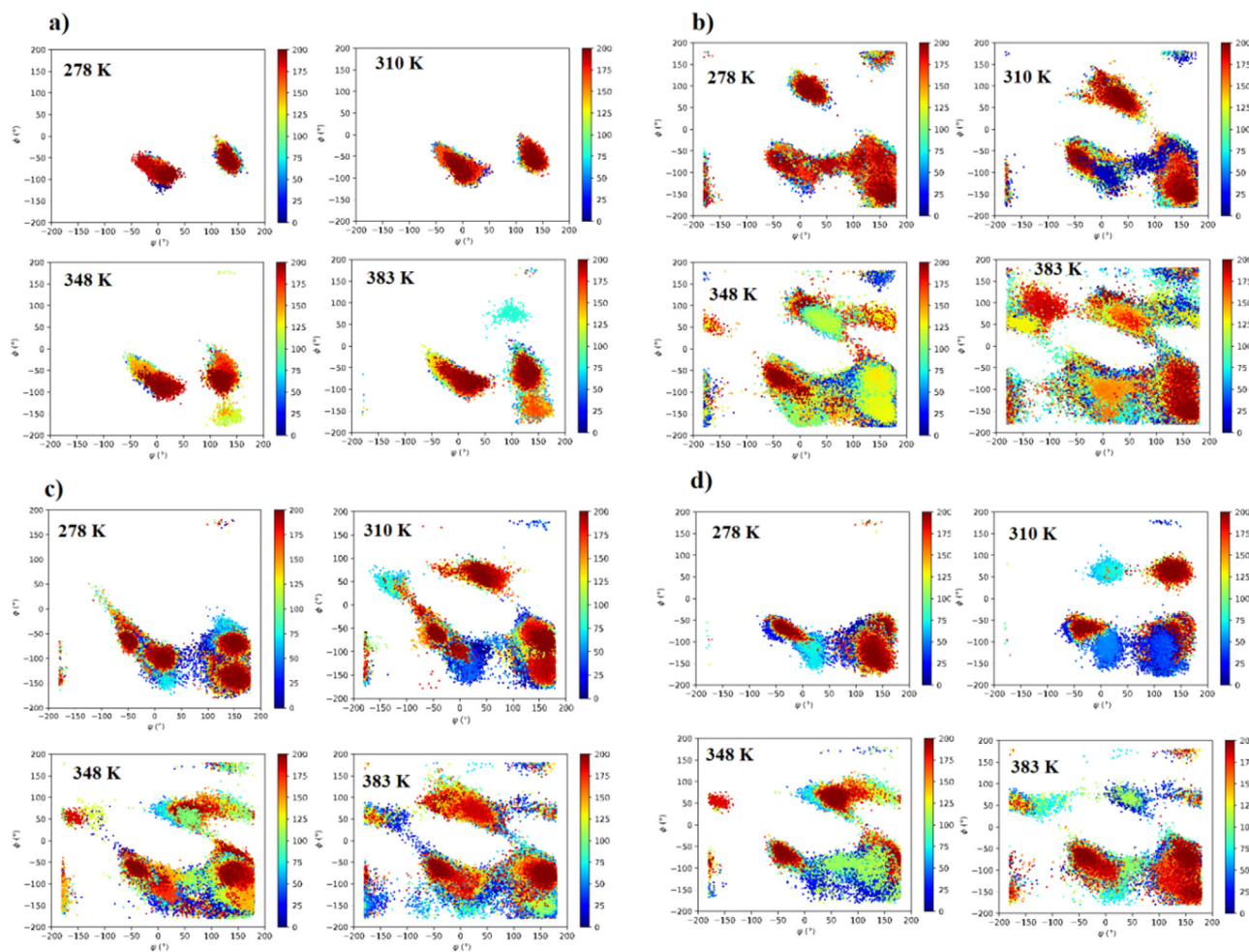


Fig. 6. Dihedral angle (ϕ , ψ) transitions of amino acid residues at various sites of monomer during the course of simulations (a) Catalytic dyad (b) Substrate binding site (c) Dimerization site (d) Allosteric site. The colour scale represents the simulation time in ns.

fluctuations providing better stability and biological activity to the protein. Even though deviations are observed for 348 K and 383 K, the residues showed a tendency to revert to the initial conformation during simulation. At 278 K, the residues at the substrate-binding site, dimerization site and allosteric site tend to remain in the initial structure during the simulation. With temperature rise, it can be seen that there are more deviations in ϕ , ψ angles which might be due to the changes in the local and non-local interactions. This includes the making and breaking of non-covalent interactions such as hydrogen bonds, salt bridges, van der Waals interactions and aromatic interactions. This also suggests the local structural changes at the secondary structure elements of protein binding sites which might get neglected in global conformational analysis. The results for dimer-chain A show a similar trend as monomer (Figure S11). Therefore, it became more important to see the water role at the binding site, as there are many deviations at the protein binding sites with rise in temperature.

3.2. Role of water at binding sites

The structure and dynamics properties of water molecules were calculated to understand the role of water in the previously mentioned binding sites as given in (Table 1). The amino acid residues, His41, Cys117, Glu166 and Arg298 were selected from the catalytic dyad, dimerization site, substrate-binding site and allosteric site respectively to analyze the role of water. These residues are

selected based on their interaction with the organic moieties/inhibitors at four different binding sites.

3.2.1. Radial distribution function

The radial distribution functions (RDF) between the pair of atoms are used to analyse the structural properties of the system. For this, we have plotted the RDF between oxygen atoms and hydrogen atom of water molecules $g_{ow-ow}(r)$, $g_{Hw-ow}(r)$ (Fig. 7) and $g_{C\alpha-ow}$ pair correlation functions of His41, Cys117, Glu166 and Arg298 amino acids (Fig. 8) at four different temperatures. It can be seen that there are three distinct peaks in Fig. 7(a). The first peak is the highest and sharp, which is found at the separation of about 0.345 nm from centred atom's position. The second and third peaks are relatively shorter and wider, which are located approximately at positions 0.57 nm and 0.68 nm, respectively. It can be seen that the peak height of the RDFs decreases with the increase in temperature and higher solvation shell peaks are not well defined. Therefore, it can be concluded that there is a change in the hydration shell of the water molecules with the rise in temperature which becomes less compact.

Further, to study the detailed overview of the arrangement of water molecules around C_{α} atom of the amino acids mentioned at the binding sites (Table 1), we plotted the C_{α} -Ow RDF around His41, Cys117, Glu166 and Arg298. Therefore, we calculated the C_{α} -Ow RDF of four selected amino acid residues at different temperatures in Fig. 8. Here also, it is found that the solvation of the

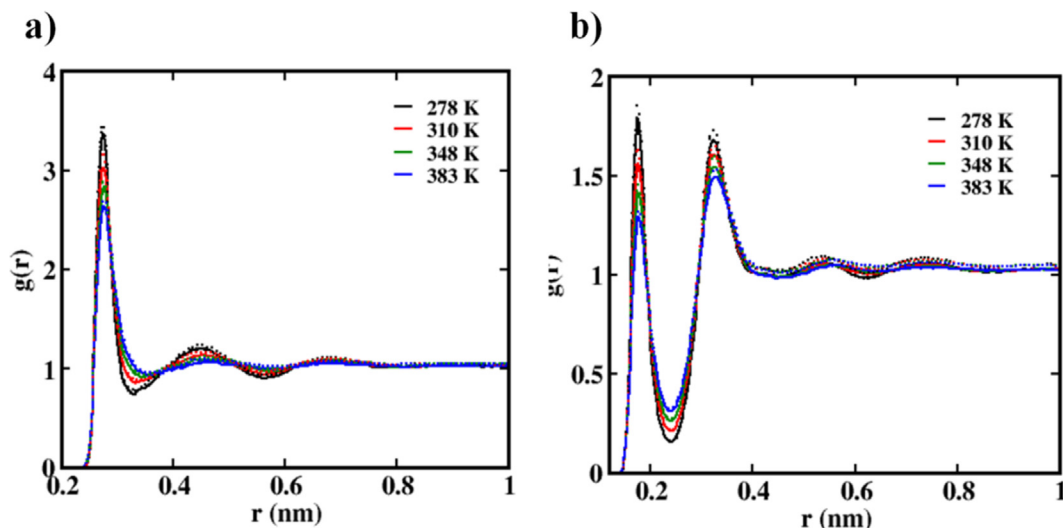


Fig. 7. RDF of (a) Oxygen–Oxygen and (b) Hydrogen–Oxygen of water molecules at different temperatures.

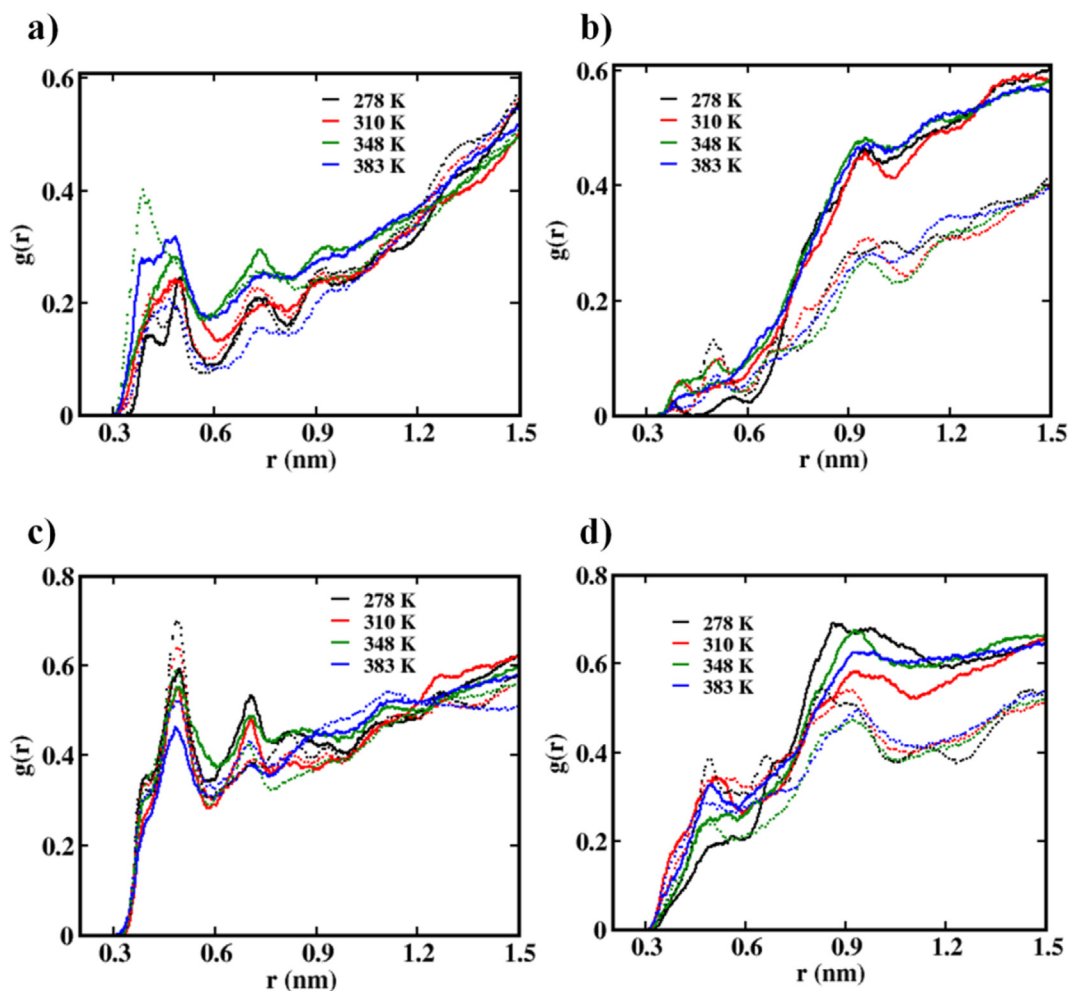


Fig. 8. Radial distribution function of C α -Ow of key amino acid residues at various binding site of protease at different temperatures in SPC/E water model. Solid line and dash line represent monomer and dimer-Chain A respectively. (a) His41 (b) Cys117 (c) Glu166 (d) Arg298.

water molecules becomes broader with the increase in the temperature. In the case of the monomer, the water molecules are more altered around C α atom of His41 at lower temperatures and showing the sharp peak but at higher temperatures the peak is broad. It

is observed that Cys117 have fewer neighbouring water molecules while Glu166 showed a narrow and sharp peak (Fig. 8 (c)) which suggests that the water molecules present here are more ordered around the binding site at lower temperature. In the monomer,

the C_{α} -Ow RDF of Arg298 indicates the low number of water molecules is present in the first solvation shell, thereby less interaction with the water. In the case of the dimer, the C_{α} -Ow RDF of all selected amino acids has a narrow and sharp peak which indicates the homogeneity of the water molecule compared to the monomer. Further, it would be interesting to check the tetrahedral arrangement of water molecules around the C_{α} atom of key amino acid residues in the solvation shell by orientational order parameter which is discussed in a later section.

3.2.2. Number of hydrogen-bonded water molecules

Further, the distorted structure of water molecules near the interface of His41, Cys117, Glu166 and Arg298 residues was confirmed by calculating the fraction (f_n) of oxygen atoms of water molecules that contains n number of water-water hydrogen bonds. The hydrogen bond distance criteria between oxygen-oxygen are considered as 3.25 Å, to incorporate the flexible hydrogen bonds [26]. A cut-off value of 6.0 Å was selected to define the interfacial water molecules present around C_{α} of selected amino acid residues of protease based on the first solvation shell in the RDF. The distance criteria will select the neighbouring water molecules with in this cut-off value obtained from RDF.

In all the cases, it has been observed that the probability of lower coordinated (mono-coordinated) water molecules increases at higher temperatures (Fig. 9). At the lower temperature, the fraction of higher coordinated water molecules is higher as compared to 348 K and 383 K. At the higher temperature the probability of higher coordinated water molecules is less. Further, we have calculated the tetrahedral order parameter to see the orientation of the water molecules around the binding site residues (Fig. 10).

3.2.3. Orientational tetrahedral order parameter

The structure of water molecules near the protein surface can be characterized based on the tetrahedral order parameter. The orientation order parameter is defined as the normalized sum of the squares of the differences between the cosines of the inter-bond angles and the cosine of the angles that would have been made if the bonds are arranged tetrahedrally [54]. Since, all the angles between bonds are the same in a tetrahedral arrangement,

$$S_g = \frac{3}{32} \sum_{j=1}^3 \sum_{k=j+1}^4 (\cos\theta_{jk} + \frac{1}{3})^2 \quad (4)$$

where θ_j , k is the angle between the central oxygen atom and the j^{th} and k^{th} bonds. The factor 3/32 normalizes the S_g values to the range $0 < S_g < 1$. The squaring ensures the contribution from each inter-bond angle is always greater than or equal to zero. If all the angles are arranged in perfect tetrahedral order, the value of S_g will be zero and 1 for extreme non-tetrahedral cases. If the arrangement is random, then S_g has a value 0.25.

The probability distribution of S_g is plotted for the water molecules which are present at a cut-off of 6.0 Å from C_{α} atoms of His41, Cys117, Glu166 and Arg298 residues selected from various ligand binding sites (Fig. 10). It can be seen that the peaks corresponding to the monomer are broader compared to the dimer, which suggests more heterogeneity of the water molecules in the monomer compared to dimer-chain A. The sharp and narrow distribution of S_g for dimer-Chain A indicates less heterogeneity of the water molecules. This homogeneous distribution of water molecules in dimer systems are observed irrespective of the temperature. In the case of monomer, a tail region towards the higher S_g value is observed indicating that the water molecules have a more distorted tetrahedral arrangement. Hence, the water structure in the first solvation shell around the binding site residues of the monomer has a more disrupted and distorted tetrahedral structure with a lower density as compared to dimer-Chain A. Further, with the

increment of the temperature this tail region decreases both for the dimer and the monomer which suggests that the water molecules are randomly and homogeneously distributed at the interface of the protein at higher temperature as compared to the lower temperature. This can be explained based on the interstitial water molecules present in the cavity formed from the binding site residues. The interstitial water molecules can lead to non-tetrahedral arrangement of water molecules called high density water [55]. With the rise in temperature, the interstitial water molecules moves out of the cavity due to thermal expansion, which may leads to more tetrahedral like water at higher temperatures. Next, we focused on the thermodynamic properties in terms of free energy, entropy and enthalpy of the water molecule with the binding site amino acid residues.

3.2.4. Free energy vs. temperature

The hydration shell formed by the water molecules around the C_{α} of binding site amino acids residues at 288 K, 308 K and 348 K is shown in Figure S12 and Table S2. The hydrogen-bonding network between the water molecules is perturbed due to the interaction of water molecules with these amino acids residues with the rise in temperatures which in turn changes the free energy. The free energy, $\Delta G(r)$, between these residues and water was determined with the pair correlation function $g(r)$ given by equation

$$\Delta G(r) = -k_B T \ln g(r) \quad (5)$$

where r is the inter-atomic distance, k_B is the Boltzmann constant and T is the temperature. Additional 5 MD simulations were performed to calculate the entropy-enthalpy contributions as shown in Table S3. The thermodynamic contributions to the free energy have been calculated from the finite difference temperature derivative of the PMF at each temperature from the $g(r)$. The RDFs for C_{α} -Ow pairs for additional temperatures (298 K, 318 K, 338 K and 358 K) are given in Figure S13. The entropy contributions, $-\Delta S(r)$ in free energy profile for 288 K, 308 K and 348 K is calculated as follows

$$-\Delta S(T) = \frac{\Delta G(T + \Delta T) - \Delta G(T - \Delta T)}{2\Delta T} \quad (6)$$

The enthalpy contribution, $\Delta H(r)$ is computed using equation (7) [56,57].

$$\Delta H(r) = \Delta G(r) + T\Delta S(r) \quad (7)$$

The entropy and enthalpy contributions to the free energy for each binding site residue are given in Figure S14. The free energy of C_{α} -Ow pairs calculated as a function of pair distance for monomer residues at different temperatures are presented in Fig. 11. With the rise in temperature, the free energy of the binding pocket increases and become more positive as compared to the lower temperature. This can be explained based on the number of hydrogen bonded water molecules present inside the cavity formed by the binding sites. With the rise in temperature, there is presence of more fractions of broken hydrogen bonds in water molecules (Fig. 9) and the water-water hydrogen bond strength also decreases. Therefore, the protein-solvent interaction becomes weaker resulting in a positive free energy. It is found that the free energy of the allosteric site and substrate binding site is lower at 288 K and higher at 348 K. The $\Delta\Delta G$ value for the water interaction is 0.3 and 0.55 kJ/mol for the substrate-binding site and allosteric site, respectively, between 288 K and 348 K. The low $\Delta\Delta G$ value (0.3 kJ/mol) indicates that the water can be easily replaced at the substrate-binding site by small organic molecules. The higher value of $\Delta\Delta G$ (0.55 kJ/mol) for the allosteric site indicates the more stability of the amino acid residues due to the water molecules. It may be concluded that the energy expense of binding the small molecules at substrate binding is less compared to the allosteric

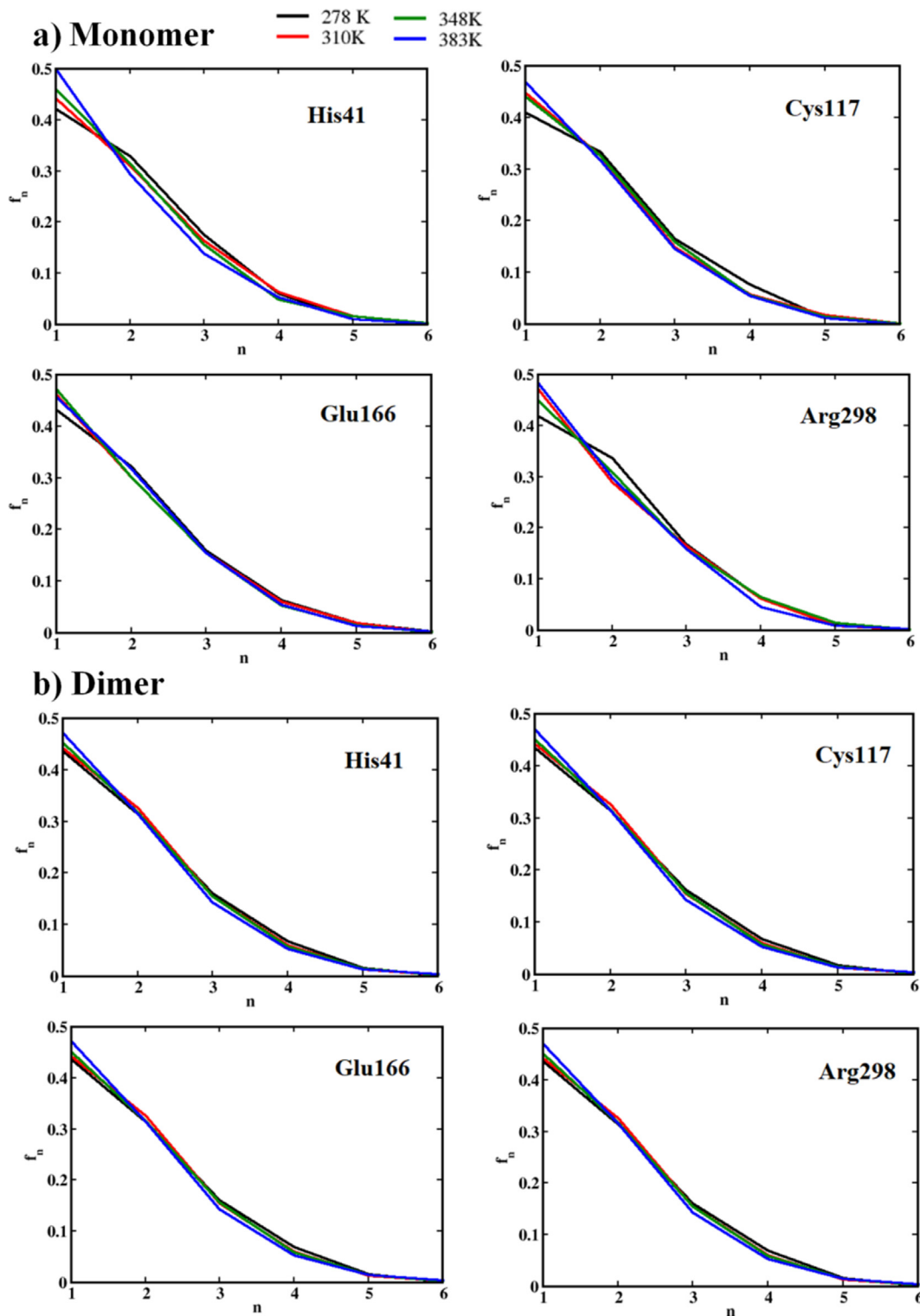


Fig. 9. Fraction of water molecules having n number of hydrogen bonds within a distance of 6.0 \AA from C_{α} of amino acid residues of (a) M^{Pro} monomer and (b) dimer-Chain A for SPC/E water model at different temperatures.

site. Even though the dimerization site residues show an increase in free energy at higher temperatures, ΔG values at 288 K and 308 K show a similar trend. At 348 K higher value for free energy can be attributed due to the thermal fluctuations of protein and the free energy differences between the 288 K and 348 K are found

to be $\Delta\Delta G = 0.51 \text{ kJ/mol}$. In the catalytic dyad, the free energy is almost similar at all temperatures. It is observed that there is the presence of stable conformation of the protein at lower temperatures which are unstable at the higher temperature. It can be found that the interaction of free energy is governed by the arrangement

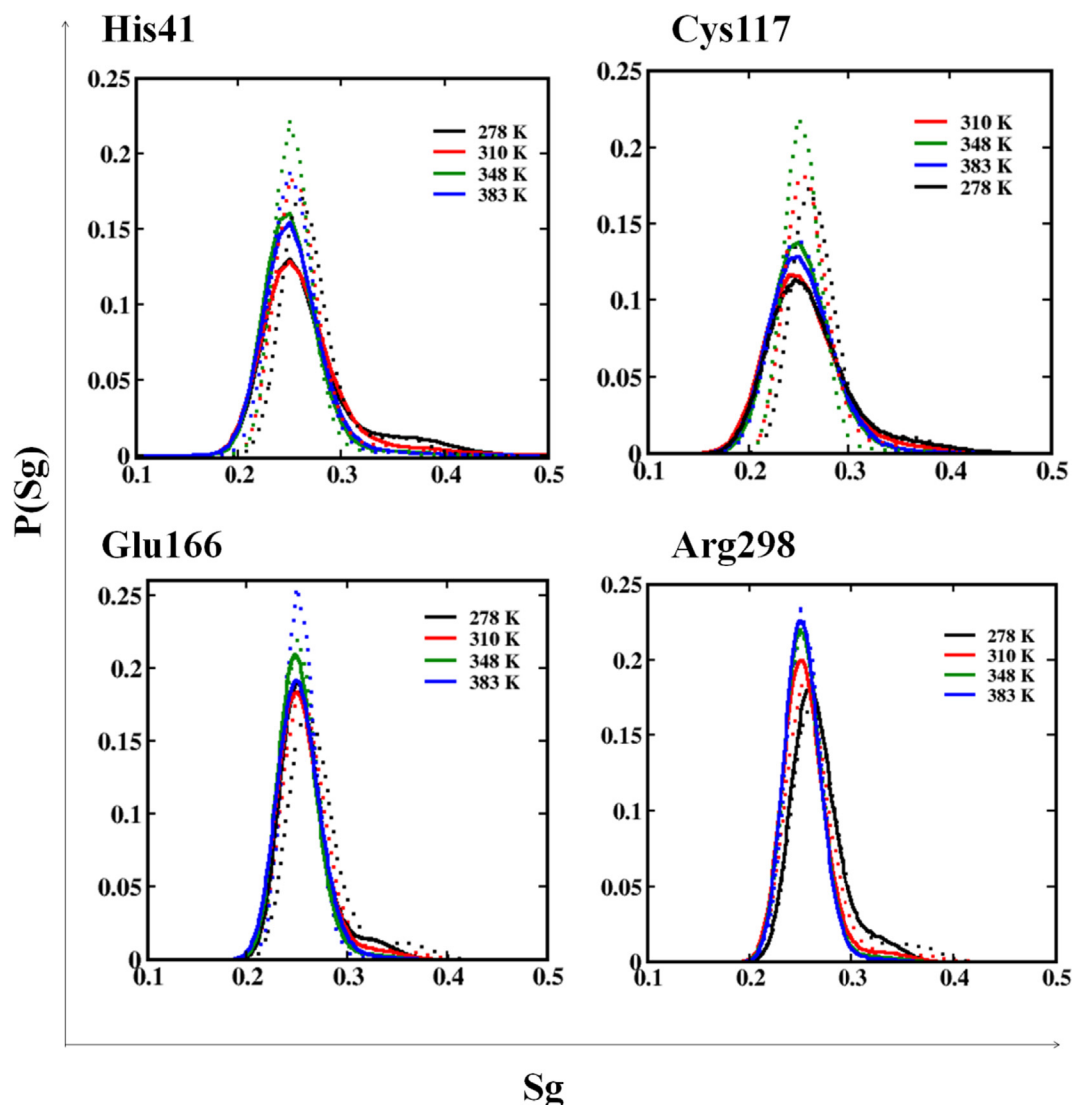


Fig. 10. Orientational tetrahedral order parameter of water molecules within 6.0 Å from C_{α} of key amino acid residues at various binding site of monomer (solid line) and dimer-Chain A (dash line) in SPC/E water model at different temperatures.

of water molecules and the fraction of hydrogen-bonded water molecules around the C_{α} atom of key amino acid residues in the solvation shell.

3.2.5. Hydrogen bond dynamics

The hydrogen bond population variables $h(t)$ and $H(t)$ were defined to calculate the hydrogen bond dynamics. $h(t) = 1$, when a particular pair of water molecules are hydrogen bonded at time t according to the definition of hydrogen bond. If two molecules remain continuously hydrogen bonded from $t = 0$ to $t = t$, then the function $H(t) = 1$ or it is 0 otherwise. The two molecules are considered to be hydrogen-bonded if the interatomic distance between hydrogen of donor and acceptor atom (H---O) is < 2.5 Å. Also, the HO-O distance is < 3.5 Å [58] and the OH---O angle criterion is considered as 130° [59]. The continuous hydrogen bond autocorrelation functions $S_{HB}(t)$ are defined as [60,61],

$$S_{HB}(t) = \frac{\langle h(0), H(t) \rangle}{\langle h(0)^2 \rangle} \quad (8)$$

where $\langle \dots \rangle$ describes an average over all the pairs of a given type. The function describes the probability of hydrogen bond pair at $t = 0$ continuously bonded up to time t . The hydrogen bond life-

time of water molecule around His41 and Glu166 residues is comparatively higher as compared to Cys117 and Arg298 residues for M^{Pro} monomer at 278 K and 310 K. The lifetime is even more at cold denaturation temperature, 278 K than 310 K. The His41 and Glu166 in catalytic dyad and substrate binding site are buried between the cleft of domain I and domain II of the protease. This forms a cavity that holds the water molecules firmly leading to strong hydrogen bond interaction between water molecules. At the high temperatures 348 K and 383 K the hydrogen bond lifetime decreases around the residues of the binding site. The higher hydrogen bond lifetime of water molecules indicates that the binding sites in the protease are stable. This is supported by the observation that the number of water molecules decreases with a rise in temperature (Fig. 2(e)) due to the expansion of binding sites. The hydrogen bond lifetime in ps of water molecules around the amino acids has been shown in Table 2. In the case of the dimer, a similar trend for water-water hydrogen bond lifetime is observed with a lower value due to its protein structural rigidity. The two monomers interact in such a way that the binding sites are buried inside the homodimer. The number of water molecules residing in the binding sites of dimer is lesser than the monomer. Therefore, the water-water hydrogen bond lifetime is lower in case of dimer as

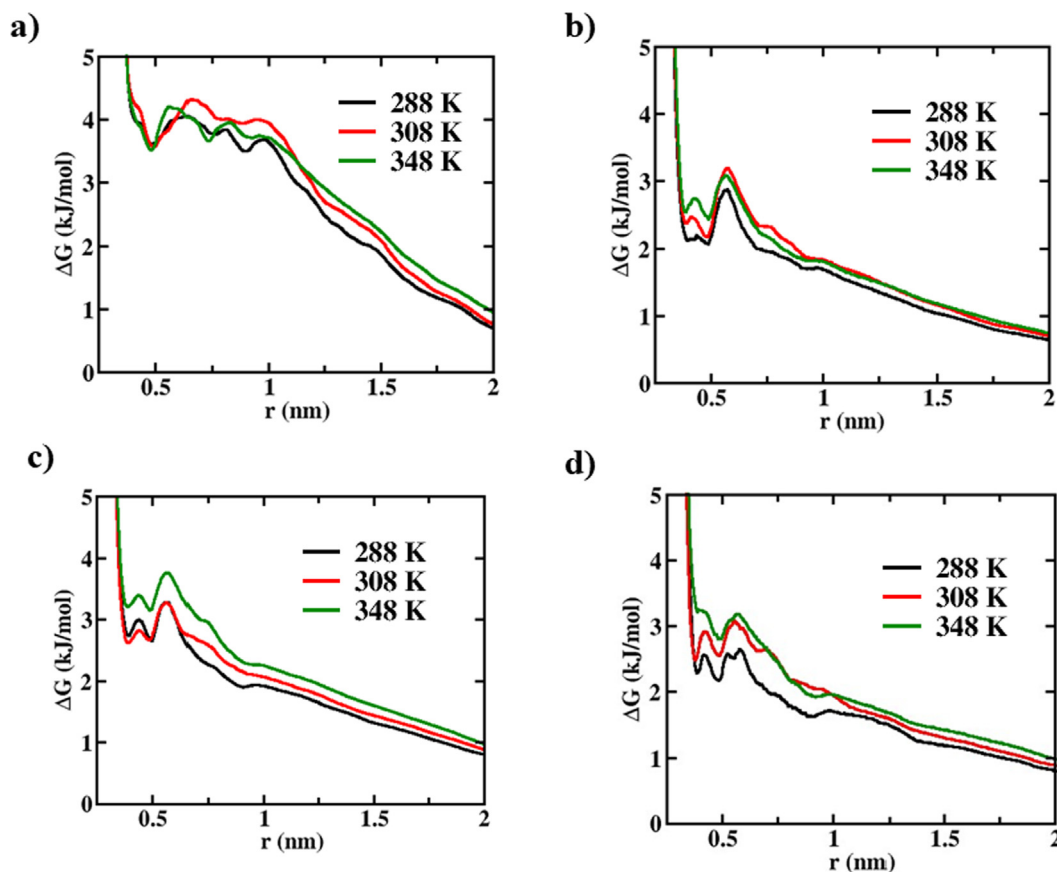


Fig. 11. Free energy function of C_{α} - O_w of amino acid residues at various binding site of protease at different temperatures for monomer protease in SPC/E water model. (a) Catalytic dyad (b) Substrate binding site (c) Dimerization site (d) Allosteric site.

Table 2

Hydrogen bond lifetime (ps) of water molecules around the key amino acid residues at the various sites of main protease in SPC/E water model.

Temperature	Monomer				Dimer			
	His41	Cys117	Glu166	Arg298	His41	Cys117	Glu166	Arg298
278 K	4.67	2.21	3.07	2.33	0.92	0.78	0.81	0.98
310 K	1.29	1.24	1.83	1.22	0.60	0.56	0.59	0.54
348 K	0.85	0.72	0.74	0.76	0.40	0.39	0.42	0.39
383 K	0.57	0.51	0.52	0.60	0.32	0.30	0.31	0.30

compared to the monomer, even though the trend is similar. Thus, it can be concluded that the hydrogen-bonded water molecules stabilize the binding sites of the protein.

3.3. Structure and dynamics of protease in mTIP3P water model

The temperature-dependent stability of monomer and dimer were studied based on structural parameters and solvent effect using mTIP3P water model. It is observed that the global conformation of the main protease monomer and dimer remains similar despite of the rise in temperature (Figure S1). The local structural evolutions are found due to certain changes in the secondary structure elements in the protease. Unlike SPC/E model, structural analyses like C_{α} -RMSD, r_{Gyr} and SASA suggest that the protease monomer undergoes thermal fluctuations starting from 310 K. The dimer-Chain A is found to be more rigid with fewer conformational evolutions. Like SPC/E water model, the solvent exposure is higher in the monomer than the dimer with the rise in temperature as expected. In the case of monomer and dimer, the water molecules present in the first solvation shell of Arg298 is less as

compared to other binding site residues in the mTIP3P water model which shows a similar trend to SPC/E water model. Like SPC/E water model, it has been observed that the probability of the lower coordinated water molecules is found at higher temperatures around all residues in both monomer and dimer. The probability distribution of order parameter around the binding site residues shows that the TIP3P water molecules are more heterogeneous in monomer compared to dimer chain-A which shows a similar trend to SPC/E water model. All the results are given in Supplementary Material, Section: mTIP3P results (Figure S18-S21).

4. Conclusion

Temperature dependence on the conformational evolution of SARS CoV-2 M^{Pro} using classical molecular dynamics simulations reveals that the global structure of the enzyme is intact with minimum changes while significant changes are observed at the various binding sites of the protease. The various binding sites of the main protease can be divided into catalytic dyad, substrate-

binding site, dimerization site and allosteric site. The basic structural changes evaluated by RMSD, rGyr and RMSF describe the thermal fluctuations in the protein with the rise in the temperature. Even though the dimerization of protease helps in the structural stability and function, the dimer turns to be unstable with rise in the temperature. We studied two different water models, namely the CHARMM-SPC/E and the CHARMM-TIP3P water model system. Network analysis shows the conformational evolution of the protease at higher temperatures. The structural analyses performed selectively for important amino acid residues at the various binding site showed the flexible nature of domain III at higher temperatures may affect the dimerization of the main protease which is crucial for its biological function. The dihedral analysis and PCA analysis around the binding sites concluded the flexible behaviour of both monomer and dimer-Chain A with an increase in temperature. The water structural properties were analysed in terms of radial distribution functions, tetrahedral order parameter (S_g) values and fraction of the number of hydrogen bonds around various binding sites of the main protease. The water structure in the solvation shell is found to be more heterogeneous at lower temperatures around the binding site amino acid residues. The higher hydrogen bond lifetime of the water molecules shows that the water molecules have a considerable effect on the stability of the binding sites. It is found that the $\Delta\Delta G$ value of interaction between water and the binding site amino acids is around 0.5 kJ/mol for both dimerization and allosteric binding site and 0.3 kJ/mol for substrate binding site. This explains the exposure of amino acid residues to the water molecules. The results of the two different water models showed a similar trend; however, the SPC/E water model showed a more pronounced effect compared to the TIP3P water model.

CRediT authorship contribution statement

Pushyraga P. Venugopal: Software, Methodology, Validation, Formal analysis, Visualization, Investigation, Data curation, Writing – original draft. **Omkar Singh:** Software, Methodology, Validation, Formal analysis, Visualization, Data curation, Writing – original draft. **Debashree Chakraborty:** Conceptualization, Methodology, Formal analysis, Resources, Investigation, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Funding from Department of Science and Technology, Science and Engineering Research Board, India (CRG/2019/000578) is highly acknowledged. We would like to thank Schrödinger Inc. and Department of Chemistry NITK Surathkal for their constant support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.molliq.2022.119867>.

References

- [1] H.A. Aboubakr, T.A. Sharafeldin, S.M. Goyal, Stability of SARS-CoV-2 and Other Coronaviruses in the Environment and on Common Touch Surfaces and the Influence of Climatic Conditions: A Review, *Transbound Emerg. Dis.* 68 (2) (2021) 296–312, <https://doi.org/10.1111/tbed.13707>.
- [2] I. Maffucci, A. Contini, In Silico Drug Repurposing for SARS-CoV-2 Main Proteinase and Spike Proteins, *J. Proteome Res.* 19 (11) (2020) 4637–4648,

- <https://doi.org/10.1021/acs.jproteome.0c00383.1021/acs.jproteome.0c00383.s001>.
- [3] B.K. Das, D. Chakraborty, Epitope-Based Potential Vaccine Candidate for Humoral and Cell-Mediated Immunity to Combat Severe Acute Respiratory Syndrome Coronavirus 2 Pandemic, *J. Phys. Chem. Lett.* 11 (22) (2020) 9920–9930, <https://doi.org/10.1021/acs.jpcllett.0c02846>.
- [4] R. Alexpandii, J.F. De Mesquita, S.K. Pandian, A.V. Ravi, Quinolines-Based SARS-CoV-2 3CLpro and RdRp Inhibitors and Spike-RBD-ACE2 Inhibitor for Drug-Repurposing Against COVID-19: An in Silico Analysis, *Front. Microbiol.* 11 (2020) 1796, <https://doi.org/10.3389/fmicb.2020.01796>.
- [5] B. Meyer, J. Chiaravalli, S. Gellenoncourt, P. Brownridge, D.P. Bryne, L.A. Daly, A. Grauslys, M. Walter, F. Agou, L.A. Chakrabarti, C.S. Craik, C.E. Eyers, P.A. Eyers, Y. Gambin, A.R. Jones, E. Sierecki, E. Verdin, M. Vignuzzi, E. Emmott, Characterising Proteolysis during SARS-CoV-2 Infection Identifies Viral Cleavage Sites and Cellular Targets with Therapeutic Potential, *Nat. Commun.* 12 (1) (2021) 5553, <https://doi.org/10.1038/s41467-021-25796-w>.
- [6] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved α -Ketoamide Inhibitors, *Science* 368 (6489) (2020) 409–412, <https://doi.org/10.1126/science.abb3405>.
- [7] H.M. Mengist, T. Dilnessa, T. Jin, Structural Basis of Potential Inhibitors Targeting SARS-CoV-2 Main Protease, *Front. Chem.* 9 (2021) 622898, <https://doi.org/10.3389/fchem.2021.622898>.
- [8] K. Anand, Structure of Coronavirus Main Proteinase Reveals Combination of a Chymotrypsin Fold with an Extra Alpha-Helical Domain, *EMBO J.* 21 (13) (2002) 3213–3224, <https://doi.org/10.1093/emboj/cdf327>.
- [9] T.F. Yap, Z. Liu, R.A. Shveda, D.J. Preston, A Predictive Model of the Temperature-Dependent Inactivation of Coronaviruses, *Appl. Phys. Lett.* 117 (6) (2020) 060601, <https://doi.org/10.1063/5.0020782>.
- [10] S.L. Rath, K. Kumar, Investigation of the Effect of Temperature on the Structure of SARS-CoV-2 Spike Protein by Molecular Dynamics Simulations, *Front. Mol. Biosci.* 7 (2020) 583523, <https://doi.org/10.3389/fmolb.2020.583523>.
- [11] Z. Zhou, Z. Yang, J. Ou, H. Zhang, Q. Zhang, M. Dong, G. Zhang, Temperature Dependence of the SARS-CoV-2 Affinity to Human ACE2 Determines COVID-19 Progression and Clinical Outcome, *Comput. Struct. Biotechnol. J.* 19 (2021) 161–167, <https://doi.org/10.1016/j.csbj.2020.12.005>.
- [12] J. He, H. Tao, Y. Yan, S.-Y. Huang, Y. Xiao, Molecular Mechanism of Evolution and Human Infection with SARS-CoV-2, *Viruses* 12 (4) (2020) 428, <https://doi.org/10.3390/v12040428>.
- [13] N. Zhong, S. Zhang, P. Zou, J. Chen, X. Kang, Z. Li, C. Liang, C. Jin, B. Xia, Without Its N-Finger, the Main Protease of Severe Acute Respiratory Syndrome Coronavirus Can Form a Novel Dimer through Its C-Terminal Domain, *J. Virol.* 82 (9) (2008) 4227–4234, <https://doi.org/10.1128/JVI.02612-07>.
- [14] P.L. Darke, J.L. Cole, L. Waxman, D.L. Hall, M.K. Sardana, L.C. Kuo, Active Human Cytomegalovirus Protease Is a Dimer, *J. Biol. Chem.* 271 (13) (1996) 7445–7449, <https://doi.org/10.1074/jbc.271.13.7445>.
- [15] R. Khayat, R. Batra, G.A. Bebermiz, M.W. Olson, L. Tong, Characterization of the Monomer–Dimer Equilibrium of Human Cytomegalovirus Protease by Kinetic Methods, *Biochemistry* 43 (2) (2004) 316–322, <https://doi.org/10.1021/bi035170d>.
- [16] L. Silvestrini, N. Belhaj, L. Comez, Y. Gerelli, A. Lauria, V. Libera, P. Mariani, P. Marzullo, M.G. Ortore, A. Palumbo Piccionello, C. Petrillo, L. Savini, A. Paciaroni, F. Spinazzi, The Dimer-Monomer Equilibrium of SARS-CoV-2 Main Protease Is Affected by Small Molecule Inhibitors, *Sci. Rep.* 11 (1) (2021) 9283, <https://doi.org/10.1038/s41598-021-88630-9>.
- [17] H. Chen, P. Wei, C. Huang, L. Tan, Y. Liu, L. Lai, Only One Protomer Is Active in the Dimer of SARS 3C-like Proteinase, *J. Biol. Chem.* 281 (20) (2006) 13894–13898, <https://doi.org/10.1074/jbc.M510745200>.
- [18] A.S. Achutha, V.L. Pushpa, S. Suchitra, Theoretical Insights into the Anti-SARS-CoV-2 Activity of Chloroquine and Its Analogs and In Silico Screening of Main Protease Inhibitors, *J. Proteome Res.* 19 (11) (2020) 4706–4717, <https://doi.org/10.1021/acs.jproteome.0c00683>.
- [19] S. Huff, I.R. Kummetha, S.K. Tiwari, M.B. Huante, A.E. Clark, S. Wang, W. Bray, D. Smith, A.F. Carlin, M. Endsley, T.M. Rana, Discovery and Mechanism of SARS-CoV-2 Main Protease Inhibitors, *J. Med. Chem.* 65 (4) (2022) 2866–2879, <https://doi.org/10.1021/acs.jmedchem.1c00566>.
- [20] P.P. Venugopal, D. Chakraborty, Molecular Mechanism of Inhibition of COVID-19 Main Protease by β -Adrenoreceptor Agonists and Adenosine Deaminase Inhibitors Using in Silico Methods, *J. Biomol. Struct. Dyn.* 40 (11) (2022) 5112–5127, <https://doi.org/10.1080/07391102.2020.1868337>.
- [21] T. Huynh, H. Wang, B. Luan, In Silico Exploration of the Molecular Mechanism of Clinically Oriented Drugs for Possibly Inhibiting SARS-CoV-2's Main Protease, *J. Phys. Chem. Lett.* 11 (11) (2020) 4413–4420, <https://doi.org/10.1021/acs.jpcllett.0c00994>.
- [22] A. Kordzadeh, A.R. Saadatabadi, Effects of the Temperature and the PH on the Main Protease of SARS-CoV-2: A Molecular Dynamics Simulation Study, *Biointerf. Res. Appl. Chem.* 12 (6) (2021) 7239–7248, <https://doi.org/10.33263/BRIAC126.72397248>.
- [23] A. Ebrahim, B.T. Riley, D. Kumaran, B. Andi, M.R. Fuchs, S. McSweeney, D.A. Keedy, The Temperature-Dependent Conformational Ensemble of SARS-CoV-2 Main Protease (M^{PRO}); preprint, *Biophysics* (2021), <https://doi.org/10.1101/2021.05.03.437411>.
- [24] D. Cui, B.W. Zhang, N. Matubayasi, R.M. Levy, The Role of Interfacial Water in Protein-Ligand Binding: Insights from the Indirect Solvent Mediated Potential of Mean Force, *J. Chem. Theory Comput.* 14 (2) (2018) 512–526, <https://doi.org/10.1021/acs.jctc.7b01076>.

- [25] N. Morningstar-Kywi, K. Wang, T.R. Asbell, Z. Wang, J.B. Giles, J. Lai, D. Brill, B.T. Sutch, I.S. Haworth, Prediction of Water Distributions and Displacement at Protein-Ligand Interfaces, *J. Chem. Inf. Model.* 62 (6) (2022) 1489–1497, <https://doi.org/10.1021/acs.jcim.1c01266>.
- [26] O. Singh, D. Chakraborty, Preferential Binding Affinity of Ions and Their Effect on Structure and Dynamics of Water near Antimicrobial Peptide, *J. Mol. Liq.* 344 (2021) 117789, <https://doi.org/10.1016/j.molliq.2021.117789>.
- [27] S. Dayalan, N.D. Gooneratne, S. Bevinakoppa, H. Schroder, Dihedral Angle and Secondary Structure Database of Short Amino Acid Fragments, *Bioinformation* 1 (1) (2006) 78–80, <https://doi.org/10.6026/97320630001078>.
- [28] B. Hess, N.F.A. van der Vegt, Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models, *J. Phys. Chem. B* 110 (35) (2006) 17616–17626, <https://doi.org/10.1021/jp0641029>.
- [29] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers, *SoftwareX* 1–2 (2015) 19–25, <https://doi.org/10.1016/j.softx.2015.06.001>.
- [30] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkerich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins, *J. Phys. Chem. B* 102 (18) (1998) 3586–3616, <https://doi.org/10.1021/jp973084f>.
- [31] A.D. MacKerell, M. Feig, C.L. Brooks, Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations, *J. Comput. Chem.* 25 (11) (2004) 1400–1415, <https://doi.org/10.1002/jcc.20065>.
- [32] G. Bussi, D. Donadio, M. Parrinello, Canonical Sampling through Velocity Rescaling, *J. Chem. Phys.* 126 (1) (2007) 014101, <https://doi.org/10.1063/1.2408420>.
- [33] G.J. Martyna, M.L. Klein, M. Tuckerman, Nosé-Hoover Chains: The Canonical Ensemble via Continuous Dynamics, *J. Chem. Phys.* 97 (4) (1992) 2635–2643, <https://doi.org/10.1063/1.463940>.
- [34] M. Parrinello, A. Rahman, Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method, *J. Appl. Phys.* 52 (12) (1981) 7182–7190, <https://doi.org/10.1063/1.328693>.
- [35] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A Smooth Particle Mesh Ewald Method, *J. Chem. Phys.* 103 (19) (1995) 8577–8593, <https://doi.org/10.1063/1.470117>.
- [36] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A Linear Constraint Solver for Molecular Simulations, *J. Comput. Chem.* 18 (12) (1997) 1463–1472.
- [37] Z. Sun, L. Wang, X. Li, C. Fan, J. Xu, Z. Shi, H. Qiao, Z. Lan, X. Zhang, L. Li, X. Zhou, Y. Geng, An Extended Conformation of SARS-CoV-2 Main Protease Reveals Allosteric Targets, *Proc. Natl. Acad. Sci. U.S.A.* 119 (15) (2022), <https://doi.org/10.1073/pnas.2120913119> e2120913119.
- [38] Y. Yan, H. Tao, J. He, S.-Y. Huang, The HDock Server for Integrated Protein-Protein Docking, *Nat. Protoc.* 15 (5) (2020) 1829–1852, <https://doi.org/10.1038/s41596-020-0312-x>.
- [39] F.C. Lawyer, S. Stoffel, R.K. Saiki, S.Y. Chang, P.A. Landre, R.D. Abramson, D.H. Gelfand, High-Level Expression, Purification, and Enzymatic Characterization of Full-Length Thermus Aquaticus DNA Polymerase and a Truncated Form Deficient in 5' to 3' Exonuclease Activity, *Genome Res.* 2 (4) (1993) 275–287, <https://doi.org/10.1101/gr.2.4.275>.
- [40] C. Vieille, G.J. Zeikus, Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability, *Microbiol. Mol. Biol. Rev.* 65 (1) (2001) 1–43, <https://doi.org/10.1128/MMBR.65.1.1-43.2001>.
- [41] W. Kabsch, C. Sander, Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features, *Biopolymers* 22 (12) (1983) 2577–2637, <https://doi.org/10.1002/bip.360221211>.
- [42] C. Tanford, Protein Denaturation, in: *Advances in Protein Chemistry*; Elsevier, 1968; Vol. 23, pp 121–282. [https://doi.org/10.1016/S0065-3233\(08\)60401-5](https://doi.org/10.1016/S0065-3233(08)60401-5).
- [43] M.Y. Lobanov, N.S. Bogatyreva, O.V. Galzitskaya, Radius of Gyration as an Indicator of Protein Structure Compactness, *Mol. Biol.* 42 (4) (2008) 623–628, <https://doi.org/10.1134/S0026893308040195>.
- [44] D. Mercadante, F. Gräter, C. Daday, CONAN: A Tool to Decode Dynamical Information from Molecular Interaction Maps, *Biophys. J.* 114 (6) (2018) 1267–1273, <https://doi.org/10.1016/j.bpj.2018.01.033>.
- [45] O. Caldararu, R. Kumar, E. Oksanen, D.T. Logan, U. Ryde, Are Crystallographic B-Factors Suitable for Calculating Protein Conformational Entropy?, *Phys. Chem. Chem. Phys.* 21 (33) (2019) 18149–18160, <https://doi.org/10.1039/C9CP02504A>.
- [46] J.C. Ferreira, S. Fadl, A.J. Villanueva, W.M. Rabeh, Catalytic Dyad Residues His41 and Cys145 Impact the Catalytic Activity and Overall Conformational Fold of the Main SARS-CoV-2 Protease 3-Chymotrypsin-Like Protease, *Front. Chem.* 9 (2021) 692168, <https://doi.org/10.3389/fchem.2021.692168>.
- [47] C. Huang, P. Wei, K. Fan, Y. Liu, L. Lai, 3C-like Proteinase from SARS Coronavirus Catalyzes Substrate Hydrolysis by a General Base Mechanism, *Biochemistry* 43 (15) (2004) 4568–4574, <https://doi.org/10.1021/bi036022q>.
- [48] B. Goyal, D. Goyal, Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy, *ACS Comb. Sci.* 22 (6) (2020) 297–305, <https://doi.org/10.1021/acscombsci.0c00058>.
- [49] T.J. El-Baba, C.A. Lutowski, A.L. Kantsadi, T.R. Malla, T. John, V. Mikhailov, J.R. Bolla, C.J. Schofield, N. Zitzmann, I. Vakonakis, C.V. Robinson, Allosteric Inhibition of the SARS-CoV-2 Main Protease: Insights from Mass Spectrometry Based Assays**, *Angew. Chem. Int. Ed.* 59 (52) (2020) 23544–23548, <https://doi.org/10.1002/anie.202010316>.
- [50] X. Daura, K. Gademann, B. Jaun, W.F. van Gunsteren, A.E. Mark, Peptide Folding: When Simulation Meets Experiment. 5.
- [51] O. Carugo, S. Pongor, A Normalized Root-Mean-Square Distance for Comparing Protein Three-Dimensional Structures, *Protein Sci.* 10 (7) (2008) 1470–1473, <https://doi.org/10.1110/ps.690101>.
- [52] O. Singh, P.P. Venugopal, A. Mathur, D. Chakraborty, Temperature-Dependent Conformational Evolution of SARS CoV-2 RNA Genome Using Network Analysis, *J. Phys. Chem. B* 125 (38) (2021) 10672–10681, <https://doi.org/10.1021/acs.jpcc.1c05795>.
- [53] R. Anandakrishnan, S. Izadi, A.V. Onufriev, Why Computed Protein Folding Landscapes Are Sensitive to the Water Model, *J. Chem. Theory Comput.* 15 (1) (2019) 625–636, <https://doi.org/10.1021/acs.jctc.8b00485>.
- [54] P.-L. Chau, A.J. Hardwick, A New Order Parameter for Tetrahedral Configurations, *Mol. Phys.* 93 (3) (1998) 511–518, <https://doi.org/10.1080/002689798169195>.
- [55] S. Biswas, D. Chakraborty, B.S. Mallik, Interstitial Voids and Resultant Density of Liquid Water: A First-Principles Molecular Dynamics Study, *ACS Omega* 3 (2) (2018) 2010–2017, <https://doi.org/10.1021/acsomega.7b01996>.
- [56] B.M. Pettitt, P.J. Rossky, Alkali Halides in Water: Ion-Solvent Correlations and Ion-Ion Potentials of Mean Force at Infinite Dilution, *J. Chem. Phys.* 84 (10) (1986) 5836–5844, <https://doi.org/10.1063/1.449894>.
- [57] D.E. Smith, A.D.J. Haymet, Free Energy, Entropy, and Internal Energy of Hydrophobic Interactions: Computer Simulations, *J. Chem. Phys.* 98 (8) (1993) 6445–6454, <https://doi.org/10.1063/1.464809>.
- [58] E.D. Alves, D.X. de Andrade, A.R. de Almeida, G. Colherinhas, Atomistic Molecular Dynamics Study on the Influence of High Temperatures on the Structure of Peptide Nanomembranes Candidates for Organic Supercapacitor Electrode, *J. Mol. Liq.* 334 (2021) 116126, <https://doi.org/10.1016/j.molliq.2021.116126>.
- [59] L. Schaeffer, The Role of Functional Groups in Drug-Receptor Interactions, in: *The Practice of Medicinal Chemistry*; Elsevier, 2008; pp 359–378. <https://doi.org/10.1016/B978-0-12-417205-0.00014-6>.
- [60] A. Luzar, D. Chandler, Hydrogen-Bond Kinetics in Liquid Water, *Nature* 379 (6560) (1996) 55–57, <https://doi.org/10.1038/379055a0>.
- [61] S. Paul, A. Chandra, Hydrogen Bond Properties and Dynamics of Liquid-Vapor Interfaces of Aqueous Methanol Solutions, *J. Chem. Theory Comput.* 1 (6) (2005) 1221–1231, <https://doi.org/10.1021/ct050098d>.