

# **Supplementary Information for: Molecular Simulations Explain The Exceptional Thermal Stability, Solvent Tolerance And Solubility of Protein-Polymer Surfactant Bioconjugates in Ionic Liquids**

Sudarshan Behera and Sundaram Balasubramanian\*

*Chemistry and Physics of Materials Unit*

*Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064, India*

E-mail: [bala@jncasr.ac.in](mailto:bala@jncasr.ac.in)

# Contents

<b>List of Tables</b>	<b>S3</b>
<b>List of Figures</b>	<b>S3</b>
<b>S1 Computational Details</b>	<b>S7</b>
S1.1 Choosing an ionic liquid . . . . .	S7
S1.2 Initial structure of the surfactant . . . . .	S7
S1.3 Simulation of bulk liquid [EMIM][NTf <sub>2</sub> ] . . . . .	S8
S1.4 Simulation of P-Water . . . . .	S8
S1.5 Simulation of P-IL . . . . .	S9
S1.6 Simulation of cationized protein in water . . . . .	S9
S1.7 Simulation of SFPL . . . . .	S10
S1.8 Simulation of PS-Water . . . . .	S11
S1.9 Simulation of PS-IL . . . . .	S11
S1.10 Simulation of PS-IL and PS-Water at high temperatures . . . . .	S11
S1.11 Force Field Parameters of HEM to simulate oxy-myoglobin . . . . .	S12
S1.12 Additional Computational Details . . . . .	S12
S1.13 Number Density Profiles . . . . .	S13
<b>S2 Results</b>	<b>S15</b>
S2.1 Validation of Force field parameters of oxy-myoglobin . . . . .	S15
<b>S3 Movies</b>	<b>S15</b>
S3.1 Aggregation of the surfactant molecules on LipA surface in its PS-Water form (SM1.mp4) . . . . .	S15
S3.2 Dynamics of the surfactant molecules in the PS-IL form of LipA (SM2.mp4)	S16

S3.3	Insertion of the alkyl parts of the surfactants into the hydrophobic core of Myoglobin in its PS-Water form (SM3.mp4) . . . . .	S16
<b>S4</b>	<b>Tables</b>	<b>S17</b>
<b>S5</b>	<b>Figures</b>	<b>S22</b>
	<b>References</b>	<b>S48</b>

## List of Tables

S1	Additional information about the SFPLs of the three proteins	S17
S2	Summary of all the simulations involving Lysozyme	S18
S3	Summary of all the simulations involving Myoglobin	S19
S4	Number of intra-protein H-bonds	S20
S5	Number of EMIM and NTf <sub>2</sub> ions present within 5 Å of proteins in P-IL and PS-IL systems	S20
S6	Semi-axes lengths (radii) and shapes of the three proteins	S21

## List of Figures

S1	The crystal structures of LipA, Lysozyme and Myoglobin	S22
S2	Validation of HEM force field in Myoglobin	S23
S3	The protocol to simulate SFPL of Lysozyme and Myoglobin	S24

S4	The structures of SFPLs of LipA, Lysozyme and Myoglobin	S25
S5	SASA of protein and exposure of SFT alkyl tail	S26
S6	RMSDs of different configurations of LipA in P-Water system simulated at different temperatures	S27
S7	RMSDs of different configurations of LipA in PS-Water system simulated at different temperatures	S27
S8	RMSDs of different configurations of LipA in P-IL and PS-IL systems simulated at 300 K	S28
S9	RMSDs of different configurations of LipA in the PS-IL system simulated at higher temperatures	S28
S10	RMSDs of different configurations of Lysozyme in P-Water system simulated at different temperatures	S29
S11	RMSDs of different configurations of Lysozyme in PS-Water system simulated at different temperatures	S29
S12	RMSDs of different configurations of Lysozyme in P-IL and PS-IL systems simulated at 300 K	S30
S13	RMSDs of different configurations of Lysozyme in the PS-IL system simulated at higher temperatures	S30
S14	RMSDs of different configurations of Myoglobin in P-Water system simulated at different temperatures	S31
S15	RMSDs of different configurations of Myoglobin in PS-Water system simulated at different temperatures	S31

S16	RMSDs of different configurations of Myoglobin in P-IL and PS-IL systems simulated at 300 K	S32
S17	RMSDs of different configurations of Myoglobin in the PS-IL system simulated at higher temperatures	S32
S18	The initial and final structures of the surfactant, and molecular structure of the IL	S33
S19	Initial and final structures of PS-complexes of all the three proteins from their PS-Water and PS-IL systems	S34
S20	Residue pairs which form H-bonds in the PS-IL but not in the P-Water, and involve atleast one polar residue	S35
S21	Residue pairs which form H-bonds in the PS-IL but not in the P-Water, and involve one polar residue and one cationized residue	S36
S22	The shapes of LipA, Lysozyme and Myoglobin	S37
S23	Electrostatic potential map of WT LipA and PS complex in PS-IL of LipA	S38
S24	Electrostatic potential map of WT Lysozyme and PS complex in PS-IL of Lysozyme	S39
S25	Electrostatic potential map of WT Myoglobin and PS complex in PS-IL of Myoglobin	S40
S26	Evolution of secondary structures in the P-Water and PS-Water systems	S41
S27	Time evolution of the number density profile of SFT tail in PS-Water	S41

S28	Distance between $C\alpha$ of Glu6 and Ala130 in P-Water and PS-Water forms of Myoglobin	S42
S29	Unfolding of LipA and Lysozyme in the PS-Water systems at higher temperatures	S43
S30	The Electrostatic potential maps of cationized LipA, Lysozyme and Myoglobin	S44
S31	Fraction of surface residues with positively charged side chains ( $F_{SRP}$ )	S45
S32	Number density profile of PEG part and alkyl part of the surfactant in PS-IL systems	S46
S33	Various RDF and coordination number plots	S47

# S1 Computational Details

## S1.1 Choosing an ionic liquid

Solubility and thermostability of proteins are shown to be higher when their SFPLs are dissolved in various kinds of ionic liquid such as [bmpy][OTf], [bmpy][NTf<sub>2</sub>], [bmpyrr][OAc], [bmpyrr][OTf], [bmpyrr][NTf<sub>2</sub>], and [emim][EtSO<sub>4</sub>].<sup>1,2</sup> Unavailability of good force field parameters for these ionic liquids, which reproduce their transport properties quantitatively, is an issue. However, excellent force field parameters of [EMIM][NTf<sub>2</sub>] are available, which express both the structural as well as transport properties quite well,<sup>3</sup> and the [EMIM].[NTf<sub>2</sub>] has also been used to dissolve proteins with covalent attachment of surfactants.<sup>4,5</sup> Hence, [EMIM].[NTf<sub>2</sub>] was chosen as the ionic liquid to carry out this work.

## S1.2 Initial structure of the surfactant

A random elongated structure of the surfactant (SI Figure S18C), which is used in the preparations of solvent-free liquid of Lysozyme and Myoglobin (SI Figure S18A, was constructed employing GaussView.<sup>6</sup> It was geometry optimized using Gaussian09<sup>7</sup> at HF/6-31G\* level of theory, followed by RESP<sup>8</sup> atomic charges and GAFF<sup>9</sup> parameters generations using Antechamber<sup>10</sup> and Acypype.<sup>11</sup> The optimized structure was then soaked in water, energy minimized, followed by slow heating till 300 K. It was subjected to NPT equilibration at T = 300 K and P = 1 bar for 2 ns, and finally simulated at NVT ensemble at T = 300 K for 30ns. Gromos clustering algorithm<sup>12</sup> was employed to cluster the conformations of the surfactant using an rmsd cutoff of 6 Å and the structure representing the major cluster (SI Figure S18D) was used as an initial configuration for SFPL construction. The RESP and GAFF parameters of the surfactant for the structure representing the major cluster was generated and used for further simulations.

### S1.3 Simulation of bulk liquid [EMIM][NTf<sub>2</sub>]

512 molecules each of [EMIM] cation and [NTf<sub>2</sub>] anion were packed in a cubic box of length 70 Å, using Packmol.<sup>13</sup> The packed configuration was then subjected to energy minimization, followed by a gradual increase of temperature from 0 to 303.15 K during the first 500 ps of NVT equilibration. Another 500 ps of equilibration at the NVT ensemble (at 303.15 K) was carried out, followed by equilibration of 5 ns in the NPT ensemble at T = 303.15 K and P = 1 bar. A final production simulation of 30 ns was performed in the NPT ensemble. The same simulation protocol was followed to generate two more independent production trajectories of 30 ns each.

The diffusion coefficient of ions in this ionic liquid have been experimentally determined at a temperature of 353.15 K.<sup>14</sup> Thus, we performed a MD simulation of the neat liquid at the same temperature. The configuration obtained after NVT equilibration, described above, was again subjected to gradual heating from 303.15 to 353.15 K for 500 ps, and another 500 ps NVT equilibration at T=353.15 K. This was followed by 5ns of another equilibration in the NPT ensemble at T = 353.15 K and P = 1 bar. The converged box volume was used for the final production run of 100 ns at the NVT ensemble. The temperature of 353.15 K was employed to compare the diffusivities calculated from the simulations with the experimentally reported values, which is at 353.15 K.<sup>14</sup> Two more independent trajectories each of 100ns duration were generated following the same protocol.

### S1.4 Simulation of P-Water

The crystal structure of *Bacillus subtilis* Lipase A (LipA, SI Figure S1A), Hen egg-white Lysozyme (HEWL, SI Figure S1B), and Horse heart Myoglobin (SI Figure S1C) were taken from protein data bank with PDB code 1i6w,<sup>15</sup> 2vb1,<sup>16</sup> 2vly,<sup>17</sup> respectively. The protonation state of each residue in all these enzymes at pH 7.4 was determined using the ProteinPrepare web server.<sup>18</sup> These crystal structures, along with the crystal waters present were soaked, individually, in a large cubic water box to avoid interaction of a protein with its periodic

image. Appropriate number of counter ions were added to achieve charge neutrality. The aqueous systems were then energy minimized, followed by slow thermal heating till 300 K in the course of the first 500 ps and then another 500 ps of NVT equilibration at 300 K. These were then equilibrated in the NPT ensemble at  $T = 300$  K and  $P = 1$  bar for 10 ns. The heavy atoms of the proteins were position restrained during the energy minimization and NVT equilibration stages, and were released during the NPT equilibration. The converged box volume was used for the final production runs in the NVT ensemble at 300 K. Two more trajectories for each system were generated, each thermally equilibrated with reinitialized velocities. The trajectory length of one production run is 500ns and 200ns for the other two. The aqueous WT protein will be referred to as *P-Water* system henceforth. Three independent configurations of the aqueous WT proteins were also simulated for 200ns at higher temperatures (333 K and 353 K), employing the same simulation protocol.

## S1.5 Simulation of P-IL

The crystal structure of a protein (only protein residues without crystal water and other molecules/ions), 2000 pairs of [EMIM].[NTf<sub>2</sub>], and a fixed number of [EMIM] molecules (equal to the number of counter ions needed to charge neutralize the system) were packed in a large cubic box of length 14 nm using Packmol.<sup>13</sup> The additional [EMIM] molecules were replaced by appropriate counter ions using *genion* module of GROMACS. The system was energy minimized and simulated using the same procedure as described for the simulation of WT proteins in water, to generate three independent trajectories of 500 ns each for all three proteins.

## S1.6 Simulation of cationized protein in water

Cationization is one of the crucial steps in the preparation of SFPL.<sup>19,20</sup> Refer to SI Table S1 for the number of cationized sites present for each protein. Cationizations of the proteins are done in multiple steps for better equilibration of the protein structure. Each surface-exposed

aspartate and glutamate residues were modified to their respective cationized residues, one by one, each time the structure was soaked in water followed by energy minimization and thermal annealing till 300 K. For example, there are 21 surface-exposed aspartate and glutamate residues in the case of myoglobin, and hence these residues were modified in 21 steps with 21 times of soaking in water, energy minimization and slow thermal heating till 300 K. After the cationization, the simulation protocol is the same as that of the simulation of WT proteins in water.

## S1.7 Simulation of SFPL

The previously constructed SFPL of LipA was used,<sup>21</sup> whereas, for Lysozyme and Myoglobin, SFPLs were freshly constructed. A required number of surfactants (26 and 42 for Lysozyme and Myoglobin, respectively,<sup>22</sup> refer to SI Table S1) along with the cationized protein (refer to “Simulation of cationized protein” in SI section - S1.6 for more detail) were packed in a cubic box of large size using Packmol. Refer to SI section - S1.2 for the force field parameters and initial configuration of the surfactants. The cationized proteins also contain few water molecules per protein, which is three for Lysozyme<sup>23</sup> and four for myoglobin<sup>24</sup> as reported in the experiments. The water molecules are randomly chosen from the buried water molecules of the crystal structure and retained in the construction of SFPL. All the surfactant molecules were then translated and rotated manually so that the negatively charged carboxylate group of the surfactant is in close contact with the positive charge site of the protein and ensuring that there was no hard contact between any two atoms. The cationized protein, along with the surfactants and the few water molecules (3 for Lysozyme and 4 for Myoglobin), will be referred to as *PS-complex* henceforth. The PS-complex was soaked in water and energy minimized with the protein being position restrained. 27 copies of the energy minimized PS-complex along with counter ions (0 and 2 per protein in the case of Lysozyme and Myoglobin, respectively) was then packed in a large cubic box of 80 nm length using packmol. The same equilibration protocol for the simulation of SFPL was followed as mentioned in

our earlier report<sup>21</sup> and as shown in SI Figure S3. A production trajectory of 100 ns was generated for each SFPL. Snapshots of SFPLs of LipA, Lysozyme and Myoglobin at the last MD frames are shown in SI Figure S4.

### **S1.8 Simulation of PS-Water**

Three randomly chosen PS-complexes from the SFPL were extracted, and each complex was soaked individually, in a cubic box of water of large enough box length to avoid periodic contacts. Counter ions were added whenever necessary. The system was then subjected to energy minimization with the positions of protein heavy atoms being restrained. The same equilibration protocol as followed in the simulation of WT proteins in water was employed. Final production runs for each configuration were simulated in the NVT ensemble at  $T = 300$  K.

### **S1.9 Simulation of PS-IL**

The same three chosen PS-complexes, as in the simulation of PS-Water case were packed, individually, in a large cubic box of length 14 nm, along with 3000 pairs of [EMIM] and [NTf<sub>2</sub>] ions. Two additional [EMIM] ions were introduced in the case of the SFPL of Myoglobin (as there are two more surfactants present than the number of surfactant binding sites per protein for Myoglobin). The two additional [EMIM] in the Myoglobin case were later replaced by counter ions to generate charge neutral systems. The equilibration protocol is the same as that in the simulation of WT proteins in IL. Each PS-complex in IL was subjected to production run at the NVT ensemble at  $T = 300$  K.

### **S1.10 Simulation of PS-IL and PS-Water at high temperatures**

The PS-IL systems were simulated at three high temperatures; 333 K, 353 K and 383 K, whereas for the PS-Water systems, the simulations were carried out at only two high

temperatures; 333 K and 353 K, as the proteins are found to distort in PS-Water at these temperatures. The previously thermalized PS-IL and PS-Water systems at 300 K were slowly heated to the respective high temperature, followed by NPT equilibration at that particular temperature and  $P = 1$  bar. Final production runs for 300ns were simulated at NVT ensemble at the high temperature with the converged volume taken from the NPT equilibration.

### **S1.11 Force Field Parameters of HEM to simulate oxy-myoglobin**

The all-atom AMBER force field parameters for HEM with hexa-coordinated Iron of myoglobin is available at the <http://amber.manchester.ac.uk/>. The parameters are in AMBER format (prep and frcmod files). These parameters were converted to GROMACS format using tleap and Acpype.<sup>11</sup> The HEM was then added as a separate residue in the *aminoacid.rtp* file, and other required files were modified as well.

### **S1.12 Additional Computational Details**

The force field parameter set employed in the earlier work<sup>21</sup> was used for the cationized aspartate and glutamate residues, whereas the AMBER99SB-ILDN<sup>25</sup> force field was used for the standard amino acid residues. The force field parameter for the surfactant, used for SFPL of LipA, was taken from the earlier report,<sup>21</sup> whereas for the surfactant employed for the construction of SFPLs of Lysozyme and Myoglobin, force field parameters were freshly generated and explained in the “Initial structure of the surfactant” section. Conjugate gradient algorithm<sup>26</sup> and Leap-frog integrator<sup>27</sup> were employed for energy minimization and molecular dynamics simulation, respectively, using GROMACS-2018.3<sup>28</sup> code having three-dimensional periodic boundary condition employed. TIP3P water model<sup>29</sup> was used. A time step of 2fs was employed, and coordinates were dumped every 10 ps. PME<sup>30</sup> was used for electrostatic calculations, and all bonds were constrained using LINCS algorithm.<sup>31</sup> A van der Waals and electrostatic cutoff of 12 Å was used, and long-range corrections were added to both energy and pressure. For temperature and pressure coupling, Bussi-Donadio-Parrinello velocity

rescaling thermostat<sup>32</sup> and Berendsen barostat<sup>33</sup> were used, respectively. Matplotlib<sup>34</sup> was used for plotting, whereas visualization was done using VMD.<sup>35</sup> Movies were made with the help of Molywood.<sup>36</sup> Python scripting using MDAnalysis package<sup>37</sup> and in-built modules of GROMACS were used for various analyses. Unless otherwise stated, the analyses were averaged over the last 100ns of the production run, all the independent simulations, and three different protein systems. The summaries of various simulated systems of LipA, Lysozyme and Myoglobin are presented in the Table 2 of manuscript, SI Table S2 and S3, respectively.

### S1.13 Number Density Profiles

In order to generate the density profiles of various components of the systems around proteins, the shapes of LipA, Lysozyme and Myoglobin were first defined. To characterize the semi axes that define an ellipsoid, the gyration tensor (S) was constructed as follows:

$$S_{ij} = \frac{1}{N} \sum_{k=1}^N (r_i^{(k)} - r_i^{(COM)})(r_j^{(k)} - r_j^{(COM)}); S = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix}$$

N is the number of atoms, i and j are the indices that take the values of x, y and z coordinates. The square roots of the principal moments (eigenvalues) of S are the radii of the ellipsoid that describe the shape of a protein. The eigenvectors of S characterize the orientation of the ellipsoid.

The lengths of the semi-axes of the ellipsoid describing the shape of each protein are tabulated in SI Table S6. The shapes of LipA, Lysozyme and Myoglobin are spherical, cylindrical and disc-shaped, respectively (SI Figure S22 and SI Table S6). The density profile for each protein was generated as follows:

- For LipA, the density profile is defined as the number of heavy atoms (for protein and surfactant) or molecules (for EMIM and NTf<sub>2</sub>) present per Å<sup>3</sup> as a function of the radial distance from the COM of LipA.
- For Myoglobin, density profiles were obtained as follows:
  1. The COM of the protein was translated to the center of the simulation box, and all the other molecules were translated accordingly. Then the COM of the protein was translated to the origin of the laboratory axes, and the positions of other molecules were readjusted by the same translation vector.
  2. The protein molecule was rotated such that the eigenvector corresponding to the smallest eigenvalue (5.76 Å, refer to SI Table S6) was aligned with the positive Z-axis. The same rotation was also performed on every other molecule.
  3. Once the above two steps were done for the first MD frame, all the successive frames were aligned to the protein of the first frame and other molecules were translated and rotated accordingly.
  4. The Cartesian coordinate system was then transformed to a cylindrical coordinate system using the following formula:

$$\rho = \sqrt{x^2 + y^2}$$

$$\theta = \tan^{-1}\left(\frac{y}{x}\right)$$

$$z = z$$

where  $\rho \in [0, \infty]$ ,  $\theta \in [0, 2\pi]$  and  $z \in [-\infty, \infty]$ .

5. The nDP was calculated only along the z, restricting  $\rho$  values to be less than 11 Å (length of the major axis). This is because of the surface area along z is larger than along  $\rho$ .

- For Lysozyme,
  1. The procedure till step-4 as narrated for Myoglobin was followed for Lysozyme as well except in step-2, where the major axis was aligned with the positive Z-axis in the Lysozyme case.
  2. The nDP was calculated only along  $\rho$ , restricting z values to be less than 6 Å (length of the minor axis). A larger surface area is present along  $\rho$  than along z.

## S2 Results

### S2.1 Validation of Force field parameters of oxy-myoglobin

The simulation of the aqueous oxy-myoglobin (wildtype) at ambient conditions shows the structural stability of the native state of the protein. The root-mean-squared-deviation (RMSD) of the backbone atomic positions with respect to the crystal structure of oxy-myoglobin converges within 2 Å (SI Figure S2A). The core structure of HEM (shown by ball representation in SI Figure S2B) is stable during the course of the simulation, with the RMSD with respect to crystal structure being just 0.2 Å.

## S3 Movies

### S3.1 Aggregation of the surfactant molecules on LipA surface in its PS-Water form (SM1.mp4)

The alkyl parts of the surfactant (green spheres) stay away from water and form micelle-like clusters on the protein surface (blue colour) in the PS-Water form of LipA at 300 K. As a consequence, the enzyme's surface becomes more exposed to the solvent. The PEG parts of the surfactants are shown as red spheres. The carboxylate groups of the surfactants invert

themselves in the process and interact with water. Water molecules are not shown. The length of trajectory shown in the movie is 1  $\mu$ s.

### **S3.2 Dynamics of the surfactant molecules in the PS-IL form of LipA (SM2.mp4)**

The surfactant molecules (PEG part: red, alkyl part: green) remain covering the enzyme surface (blue) in the PS-IL form of LipA at 300 K (1  $\mu$ s simulation) and do not form aggregates like the PS-Water case. Ionic liquid components are not shown. The movie is generated from a 1  $\mu$ s MD trajectory.

### **S3.3 Insertion of the alkyl parts of the surfactants into the hydrophobic core of Myoglobin in its PS-Water form (SM3.mp4)**

The alkyl tails (green licorice representations) of the surfactants insert themselves into the hydrophobic core of Myoglobin (blue cartoon) in its PS-Water system at 300 K (1  $\mu$ s simulation), leading to the unfolding of Myoglobin. Only a few surfactants whose alkyl groups insert into the protein are shown. The PEG parts of the surfactants and water molecules are not displayed. The movie presented here is a 1  $\mu$ s MD trajectory.

## S4 Tables

Table S1: Additional information about the SFPLs of the three proteins.

Proteins	No. of waters per protein	No. of cationized sites	No. of SFT binding sites	No. of PS complexes in the SFPL	No. of counter ions added per protein
LipA	40	11	27	64	0
Lysozyme	3	9	26	27	0
Myoglobin	4	21	42	27	2

Table S2: Summary of all the simulations involving Lysozyme at 1 bar pressure. Abbreviations used are the same as that in the Table 2 of manuscript.

System	Temp (K)	$N_{H_2O}$	$N_{SFT}$	$N_P$	No. of [EMIM] [NTf <sub>2</sub> ]	No. of counter ions	Cubic box length (nm)	Total no. of atoms	$N_{runs}$	Trajectory length of each run (ns)
P-Water	300	10,492	0	1	0	8 Cl <sup>-</sup>	6.99	33,954	2, 1	200, 500
	333	10,492	0	1	0	8 Cl <sup>-</sup>	7.06	33,954	3	200
	353	10,492	0	1	0	8 Cl <sup>-</sup>	7.11	33,954	3	200
P-IL	300	0	0	1	2000	8 Cl <sup>-</sup>	9.539	69,968	3	500
SFPL	333	81	702	27	0	0	11.10	147,177	1	100
PS - Water	300	18,610	26	1	0	0	8.52	61,272	3	1000
	333	18,610	26	1	0	0	8.59	61,272	3	300
	353	18,610	26	1	0	0	8.65	61,272	3	300
PS-IL	300	3	26	1	3000	0	10.97	107,451	3	1500
	333	3	26	1	3000	0	11.07	107,451	3	300
	353	3	26	1	3000	0	11.13	107,451	3	300
	383	3	26	1	3000	0	11.22	107,451	3	300

Table S3: Summary of all the simulations involving Myoglobin at 1 bar pressure. Abbreviations used are the same as that in the Table 2 of manuscript.

System	Temp (K)	$N_{H_2O}$	$N_{SFT}$	$N_P$	No. of [EMIM] [NTf <sub>2</sub> ]	No. of counter ions	Cubic box length (nm)	Total no. of atoms	$N_{runs}$	Trajectory length of each run (ns)
P-Water	300	11,233	0	0	0	2 Na <sup>+</sup>	7.16	36,670	2, 1	200, 500
	333	11,233	0	0	0	2 Na <sup>+</sup>	7.24	36,670	3	200
	353	11,233	0	0	0	2 Na <sup>+</sup>	7.29	36,670	3	200
P-IL	300	0	0	0	2000	2 Na <sup>+</sup>	9.55	70,488	3	500
SFPL	333	108	1,134	27	0	27x2=54 Na <sup>+</sup>	12.74	222,858	1	100
PS - Water	300	69,023	27	1	0	2 Na <sup>+</sup>	10.59	118,282	2, 1	500, 2000
	333	69,023	27	1	0	2 Na <sup>+</sup>	10.70	118,282	3	300
	353	69,023	27	1	0	2 Na <sup>+</sup>	10.78	118,282	3	300
PS-IL	300	4	27	1	3000	2 Na <sup>+</sup>	11.05	110,254	3	1000
	333	4	27	1	3000	2 Na <sup>+</sup>	11.14	110,254	3	300
	353	4	27	1	3000	2 Na <sup>+</sup>	11.20	110,254	3	300
	383	4	27	1	3000	2 Na <sup>+</sup>	11.29	110,254	3	300

Table S4: Number of intra-protein H-bonds. These values are from one arbitrarily chosen MD frame.

Protein	System	No. of intra-protein H-bonds involving at least one residue being			Total no. of intra-protein H-bonds
		non-polar	polar	charged	
LipA	P-W	83 (62%)	<b>81 (60%)</b>	49 (37%)	134
	PS-IL	86 (58%)	<b>101 (68%)</b>	54 (36%)	148
Lysozyme	P-W	65 (58%)	72 (64%)	51 (46%)	112
	PS-IL	63 (57%)	72 (65%)	48 (44%)	110
Myoglobin	P-W	74 (60%)	69 (56%)	55 (48%)	123
	PS-IL	76 (61%)	68 (55%)	58 (47%)	124

Table S5: Number of EMIM and NTf<sub>2</sub> present within 5 Å of proteins in P-IL and PS-IL systems. An EMIM (or NTf<sub>2</sub>) is defined to be present within 5 Å of protein if its COM falls within 5 Å of any heavy atom of the protein.

Systems		LipA	Lysozyme	Myoglobin
P-IL	no. of EMIM	66 ± 3	49 ± 3	70 ± 3
	no. of NTf <sub>2</sub>	61 ± 3	60 ± 3	56 ± 3
PS-IL	no. of EMIM	24 ± 2	27 ± 2	19 ± 2
	no. of NTf <sub>2</sub>	18 ± 2	17 ± 2	12 ± 2

Table S6: Semi-axes lengths (radii) and shapes of the three proteins.

Proteins	Axis-1 (Å)	Axis-2 (Å)	Axis-3 (Å)	Shape
LipA	7.8	8.4	9.2	Spherical
Lysozyme	6.36	6.85	10.44	Cylindrical
Myoglobin	5.76	9.61	10.36	Disc Shaped

## S5 Figures

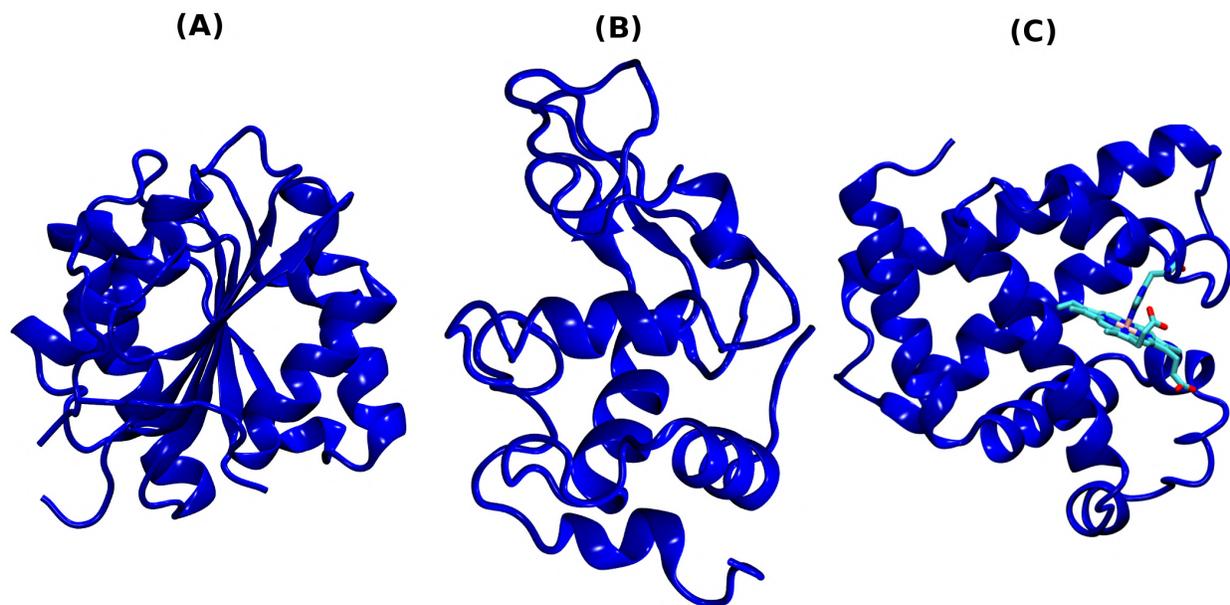


Figure S1: The various proteins studied here. (A) *Bacillus Subtilis* Lipase A (LipA), (B) Hen Egg-White Lysozyme (HEWL) and (C) Horse heart Myoglobin. Myoglobin contains HEM ligand (licorice representation), which is covalently bound to the protein via HIS93 (licorice representation).

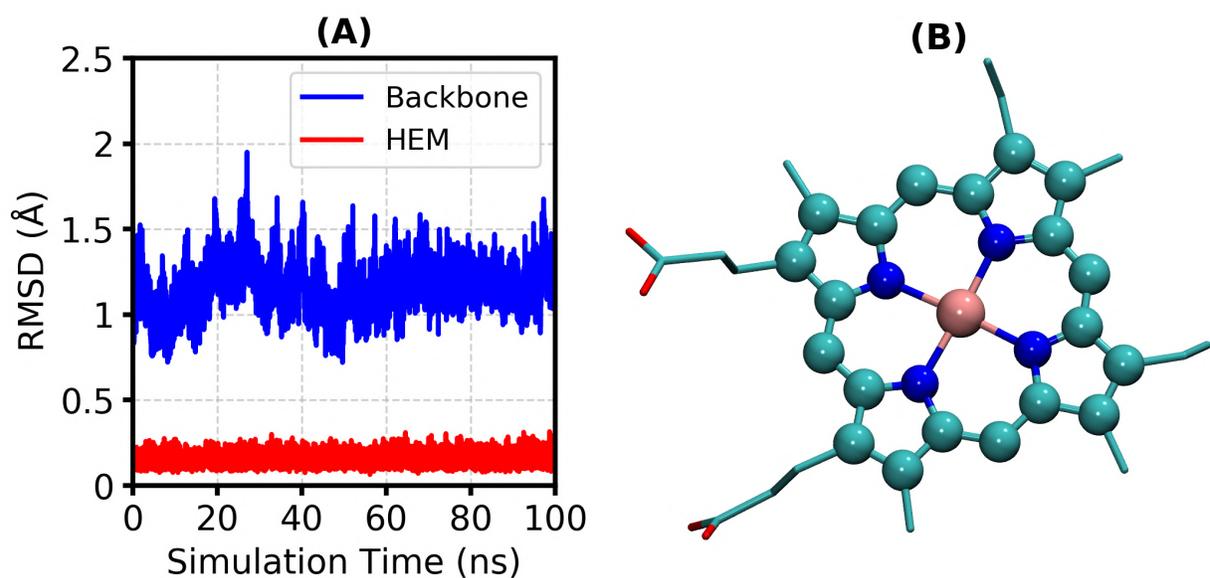


Figure S2: (A) The root-mean-squared-deviation (RMSD) of the backbone atomic positions of Myoglobin (blue) and the heavy atoms of HEM (red), which form a planar structure (shown as ball structure in panel B), with respect to the crystal structure. RMSDs of backbone atomic positions of Myoglobin in water simulations are also presented for two more configurations in SI Figure S14A, (B) Molecular structure of HEM with Carbon, Oxygen, Nitrogen and Iron atoms represented in cyan, red, blue and orange colors, respectively. The in the plane atoms are shown in ball representation.

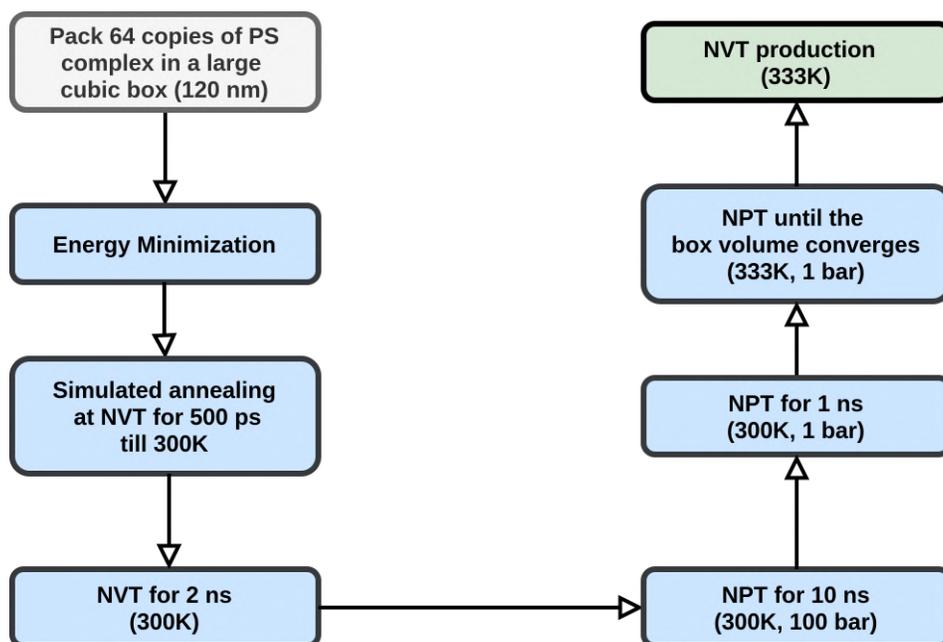


Figure S3: The protocol followed for the preparation of solvent-free liquids of Lysozyme and Myoglobin. The presence of vacuum spaces in the equilibrated systems before the final production runs was checked manually; they were not found for both Lysozyme and Myoglobin cases.

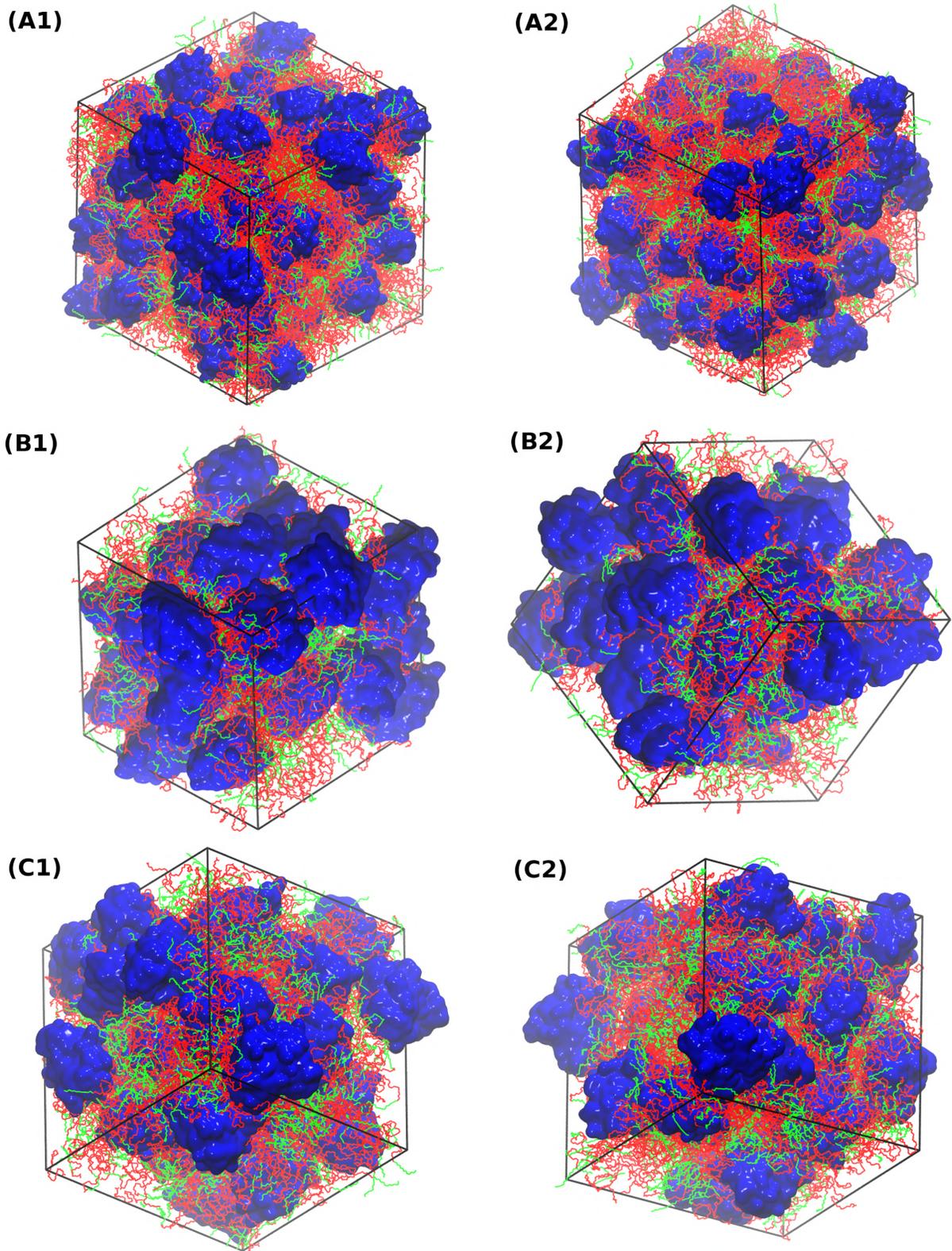


Figure S4: Two views of SFPL of LipA (64 proteins, A1 & A2), Lysozyme (27 proteins, B1 & B2) and Myoglobin (27 proteins, C1 & C2). Proteins are represented as blue surfaces, whereas surfactant PEG and alkyl tail parts are represented as red and green lines.

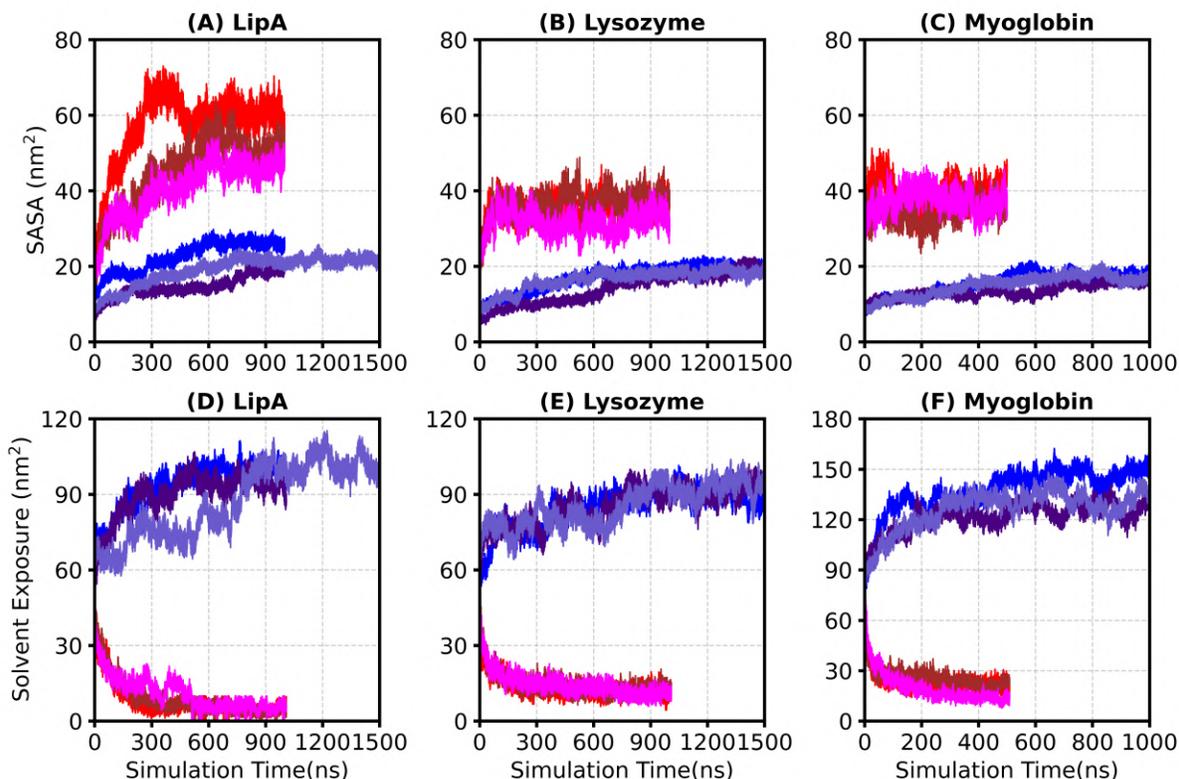


Figure S5: Time evolution of SASA (in  $\text{nm}^2$ ) of proteins (top panels; A, B, & C) and the exposure of alkyl part of surfactant to the solvent (bottom panels; D, E, & F). Red, magenta and brown colors are for the three independent simulation runs of PS-Water systems, whereas blue, indigo and slate-blue are for the three independent runs of PS-IL systems. The SASA of the protein and the solvent (IL &/or Water) exposure of the alkyl tail of SFT for all the three proteins (LipA (A & D), Lysozyme (B & E) and Myoglobin (C & F)) are presented. The SASA of a protein is the exposure of the protein surface to a sphere of radius  $1.4 \text{ \AA}$ , which is rolled over it. The exposure of the alkyl part of the surfactant to the solvent was also calculated following the same protocol. The *sasa* module of the GROMACS code was employed to compute the SASA. The index group corresponding to the protein (or alkyl tail) was fed into the *'-output'* flag of the module, whereas the index group for the PS complex (i.e., protein and surfactant atoms) was given for the *'-surface'* flag. The results of the *'-output'* flag are the SASA of protein (or the solvent exposure of alkyl tail). The values of the quantity at the time point  $t=0\text{ns}$  are that for the SFPL (as these PS complexes are extracted from the SFPL).

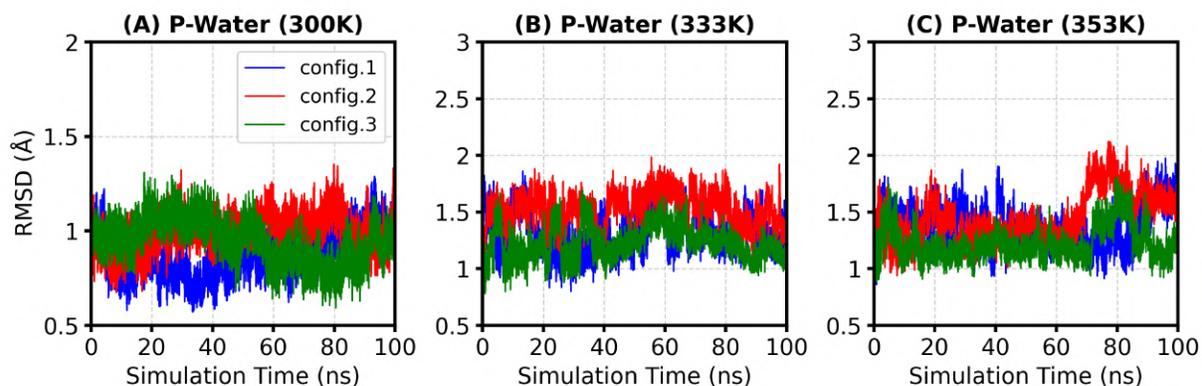


Figure S6: RMSDs of the backbone atoms of LipA with respect to the crystal structure of LipA in three different configurations of P-Water system simulated at (A) 300 K, (B) 333 K, and (C) 353 K.

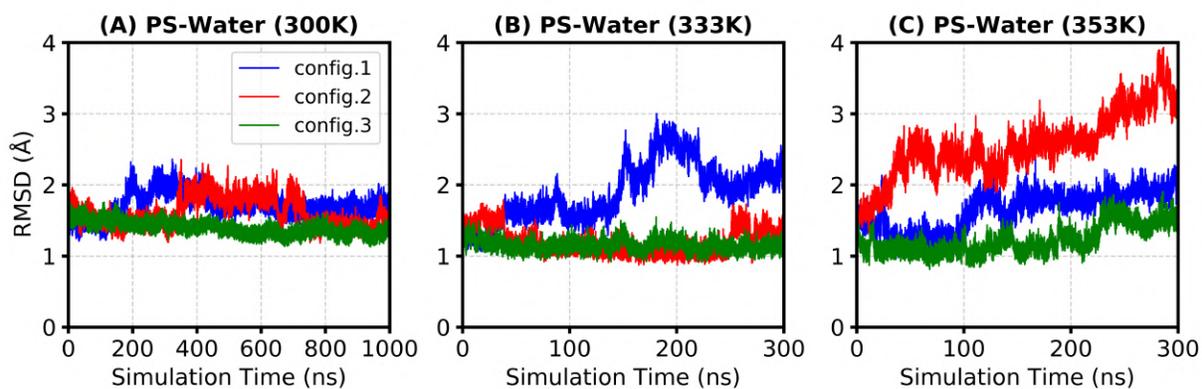


Figure S7: RMSDs of the backbone atoms of LipA with respect to the crystal structure of LipA in three different configurations of PS-Water system simulated at (A) 300 K, (B) 333 K, and (C) 353 K.

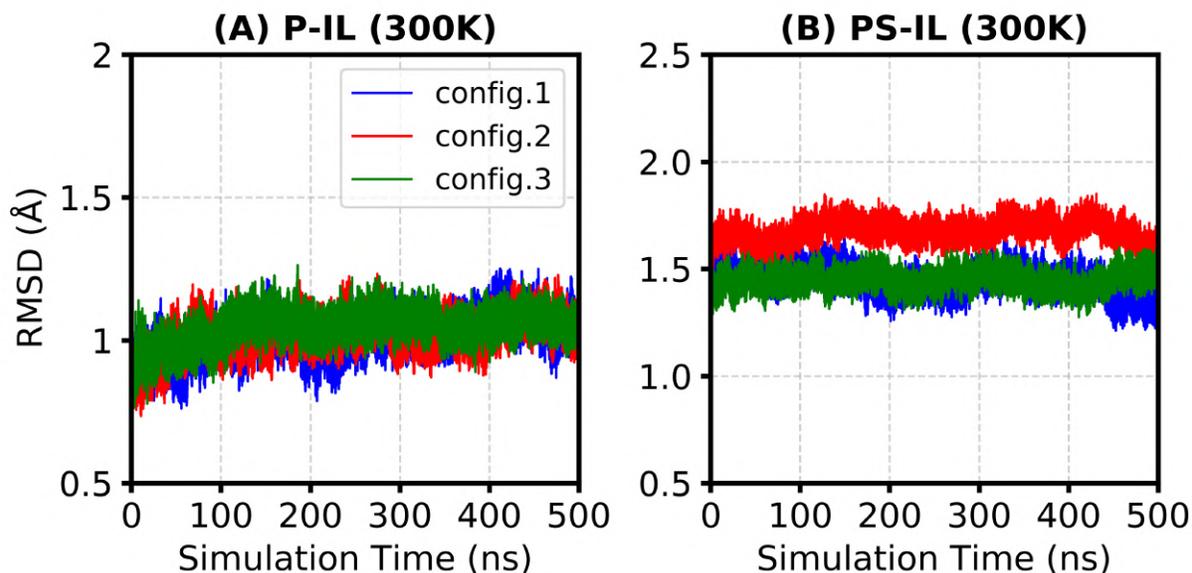


Figure S8: RMSDs of the backbone atoms of LipA with respect to the crystal structure of LipA in (A) P-IL and (B) PS-IL systems at 300 K.

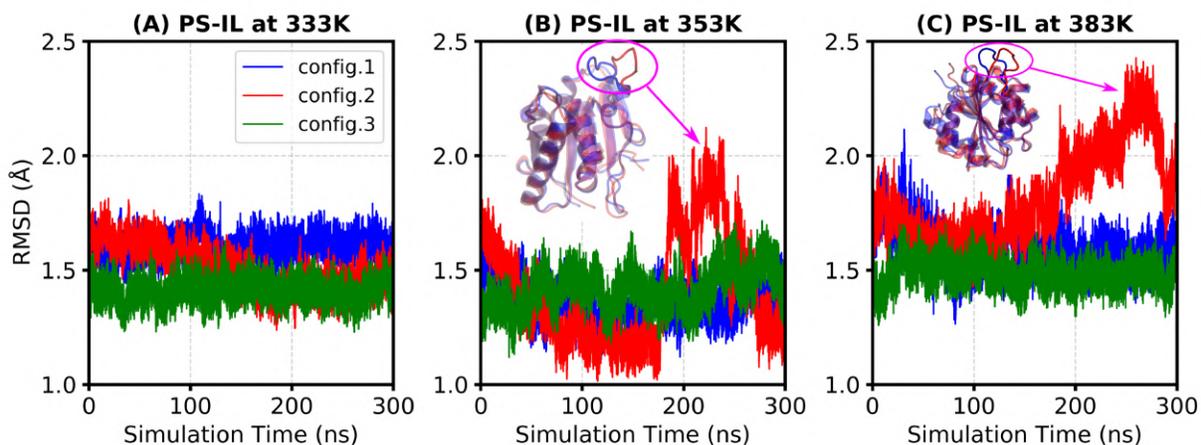


Figure S9: RMSDs of the backbone atoms of LipA with respect to the crystal structure of LipA, in PS-IL system at higher temperature; (A) at 333 K, (B) at 353 K and (C) at 383 K. Three configurations for each temperature have been shown in blue, red and green. The sudden change in RMSD for configuration-2 (red) at 353 K and 383 K are due to the loop motion of residue 131-140 and residue 114-123, respectively. The loops are highlighted in a magenta ellipse, where blue and red color cartoons represent the crystal structure and the structure in the PS-IL simulations showing high RMSD values, respectively.

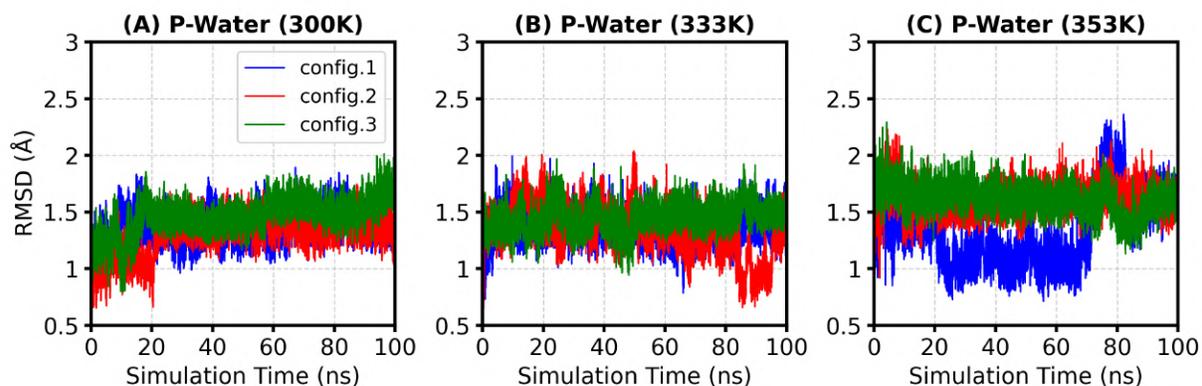


Figure S10: RMSDs of the backbone atoms of Lysozyme with respect to the crystal structure of Lysozyme in three different configurations of P-Water system simulated at (A) 300 K, (B) 333 K, and (C) 353 K.

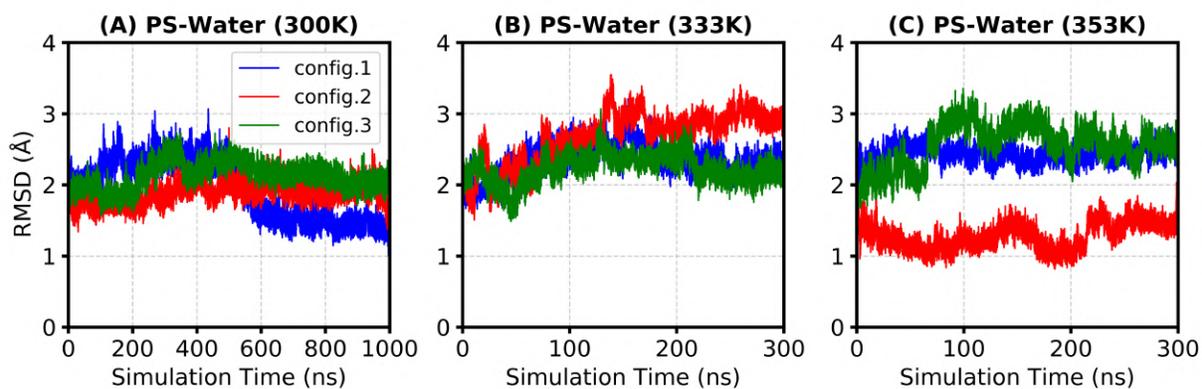


Figure S11: RMSDs of the backbone atoms of Lysozyme with respect to the crystal structure of Lysozyme in three different configurations of PS-Water system simulated at (A) 300 K, (B) 333 K, and (C) 353 K.

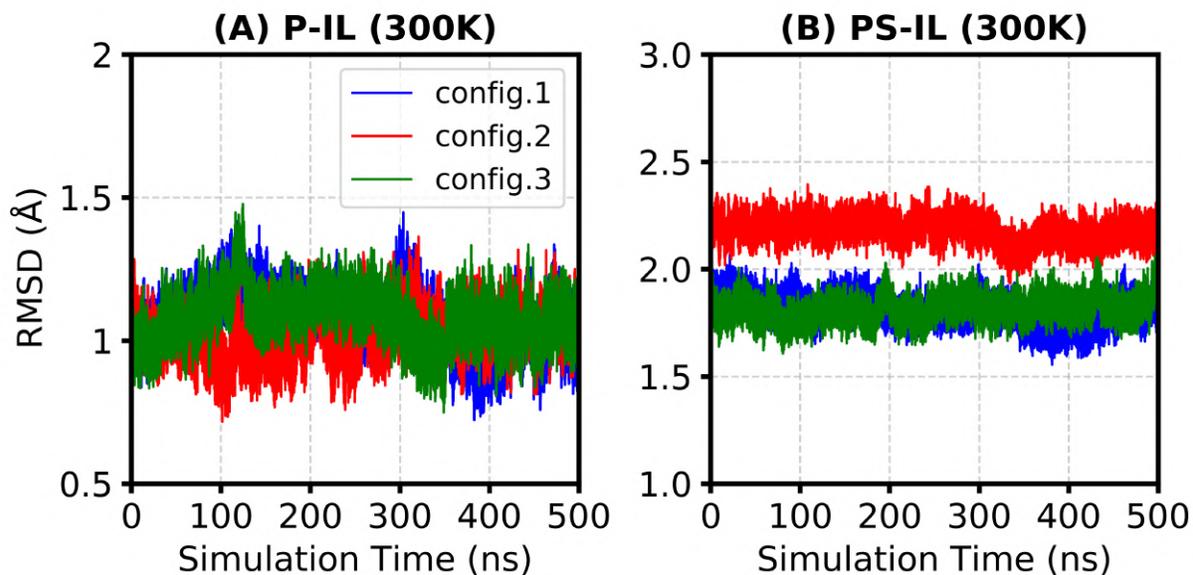


Figure S12: RMSDs of the backbone atoms of Lysozyme with respect to the crystal structure of Lysozyme in (A) P-IL and (B) PS-IL systems at 300 K.

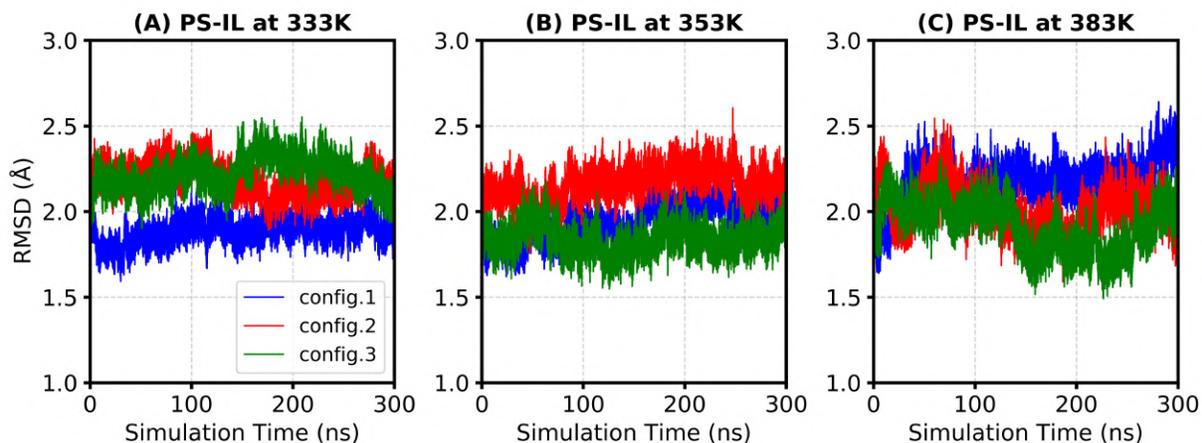


Figure S13: RMSDs of the backbone atoms of Lysozyme with respect to the crystal structure of Lysozyme in the PS-IL system at (A) 333 K, (B) 353 K, and (C) 383 K.

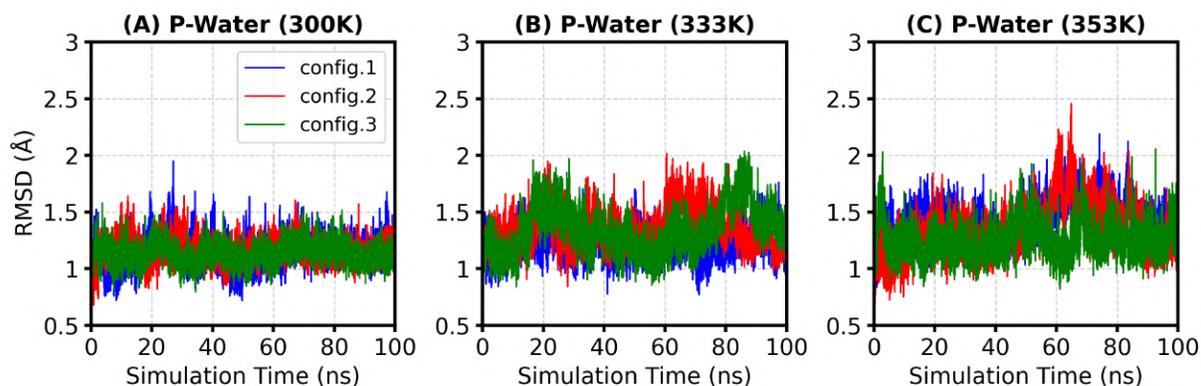


Figure S14: RMSDs of the backbone atoms of Myoglobin with respect to the crystal structure of Myoglobin in three different configurations of P-Water system simulated at (A) 300 K, (B) 333 K, and (C) 353 K.

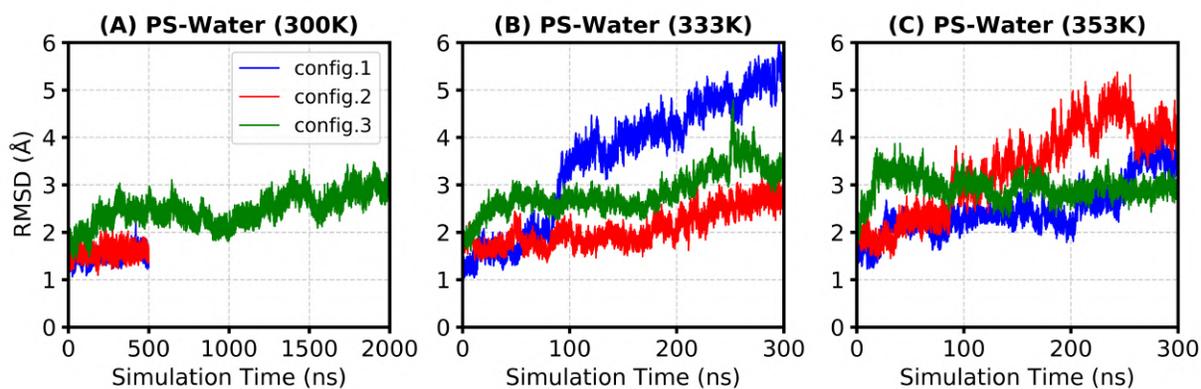


Figure S15: RMSDs of the backbone atoms of Myoglobin with respect to the crystal structure of Myoglobin in three different configurations of PS-Water system simulated at (A) 300 K, (B) 333 K, and (C) 353 K.

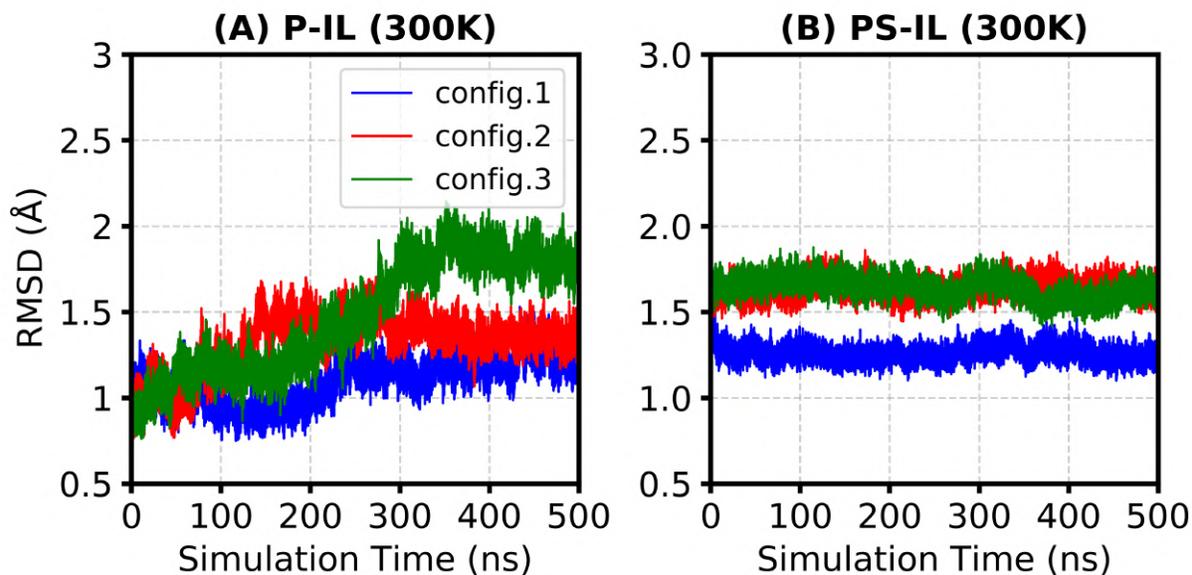


Figure S16: RMSDs of the backbone atoms of Myoglobin with respect to the crystal structure of Myoglobin in (A) P-IL and (B) PS-IL systems at 300 K.

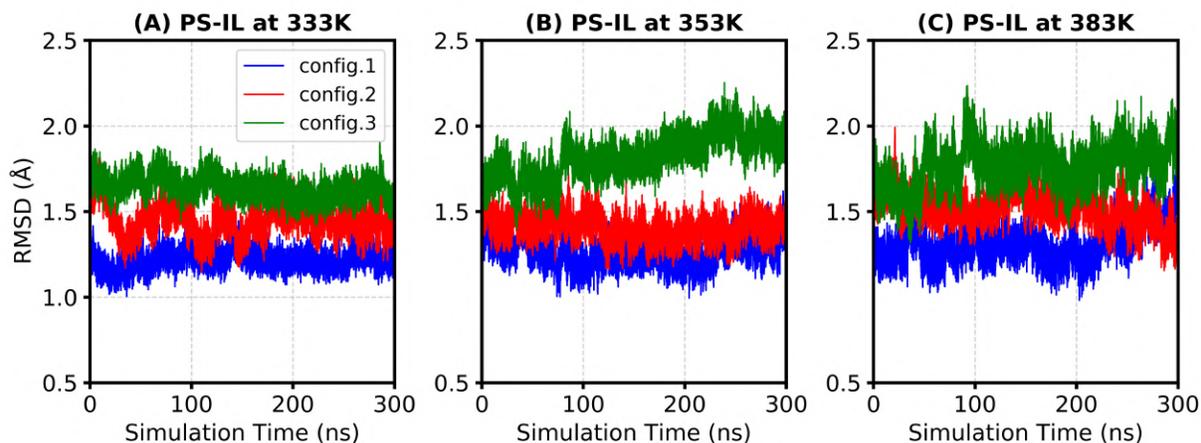


Figure S17: RMSDs of the backbone atoms of Myoglobin with respect to the crystal structure of Lysozyme in the PS-IL system at (A) 333 K, (B) 353 K, and (C) 383 K.

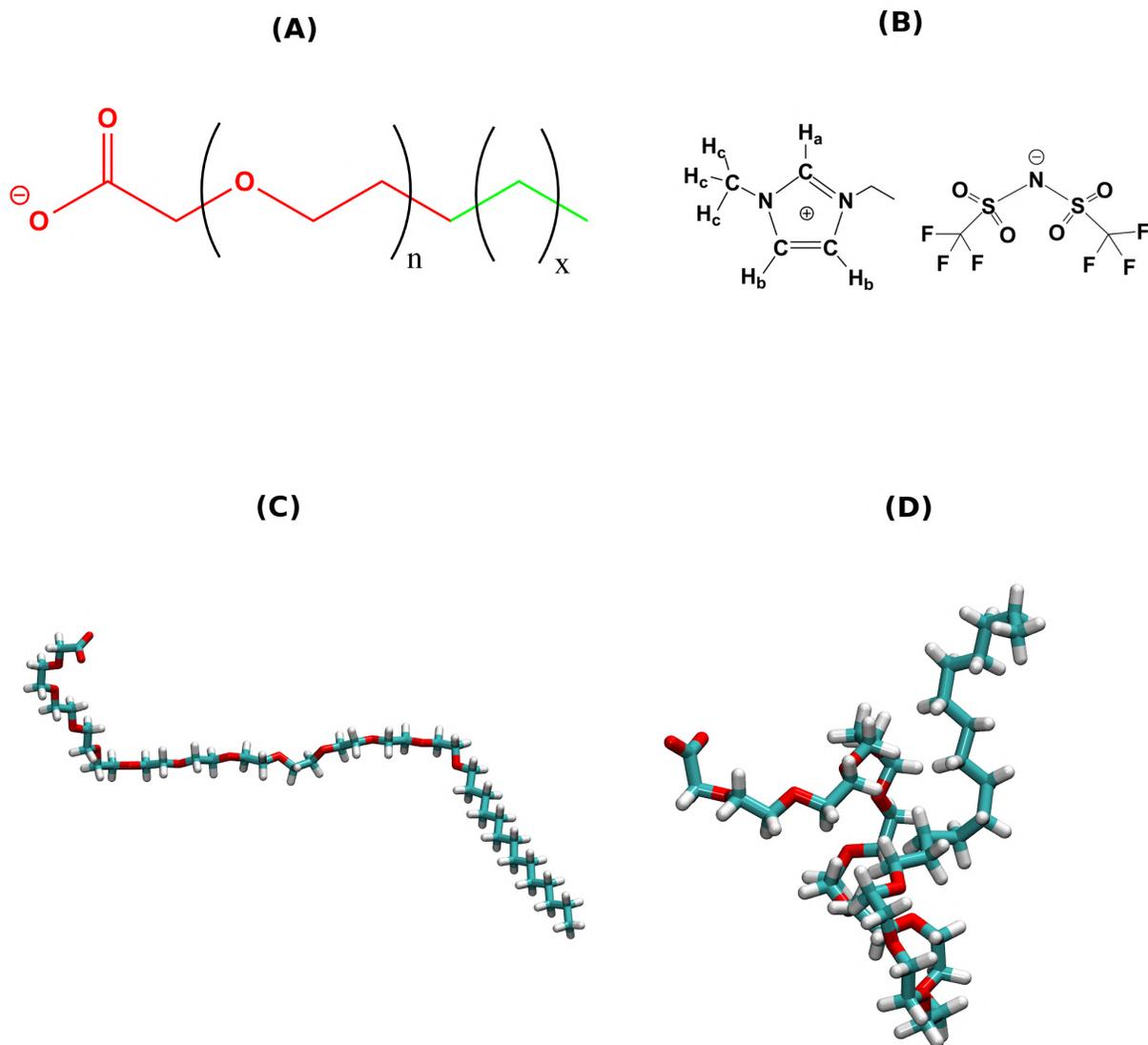


Figure S18: (A) The surfactants used for the construction of SFPLs of the three different proteins studied here. For LipA SFPL,  $n=23$  and  $x=12$ , whereas  $n=12$  and  $x=10$  for Lysozyme and Myoglobin SFPLs. The surfactants are the same as used in the experimental preparations of the SFPLs.<sup>22,38</sup> The red and green colors represent the hydrophilic and hydrophobic parts of the surfactant. (B) The molecular structure and naming convention of [EMIM].[NTf<sub>2</sub>]. (C) Randomly constructed elongated initial structure of the surfactant used for SFPLs of Lysozyme and Myoglobin ( $n=12$  and  $x=10$  in Panel A), (D) the structure representing the major cluster of the surfactant extracted from a 30 ns MD simulation of aqueous surfactant.

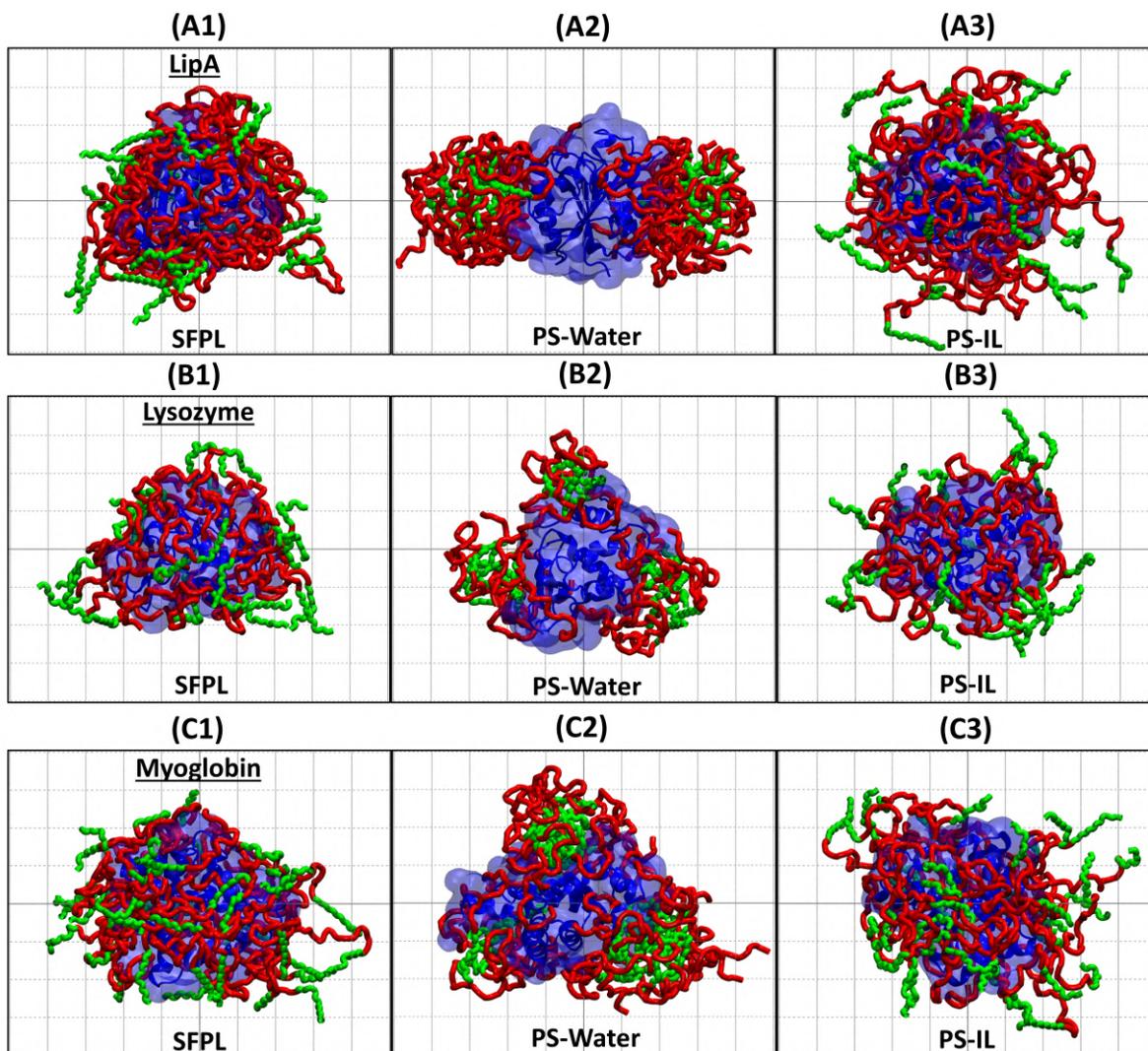


Figure S19: Another view of all the structures presented in Figure 1 and Figure 2 of the main manuscript. (A1) Initial structure of one PS complex of LipA extracted from its SFPL, (A2) the structure of PS complex of LipA in the last MD frame in the PS-Water simulation, (A3) the structure of PS complex of LipA in the last MD frame in the PS-IL simulation. Protein is shown as transparent surface and cartoon representations with blue colors, whereas the PEG and alkyl part of the surfactant are shown as red and green tubes, respectively. Water molecules and solvents are not shown. Panels B1, B2 & B3 are the same as A1, A2 & A3, but for Lysozyme whereas, C1, C2 & C3 are for Myoglobin. All the structures are presented on the same scale ( $100 \times 80 \text{ \AA}$ ), where each small grid box is a  $10 \times 10 \text{ \AA}$  square. The structures are from only one independent simulation out of the three performed.

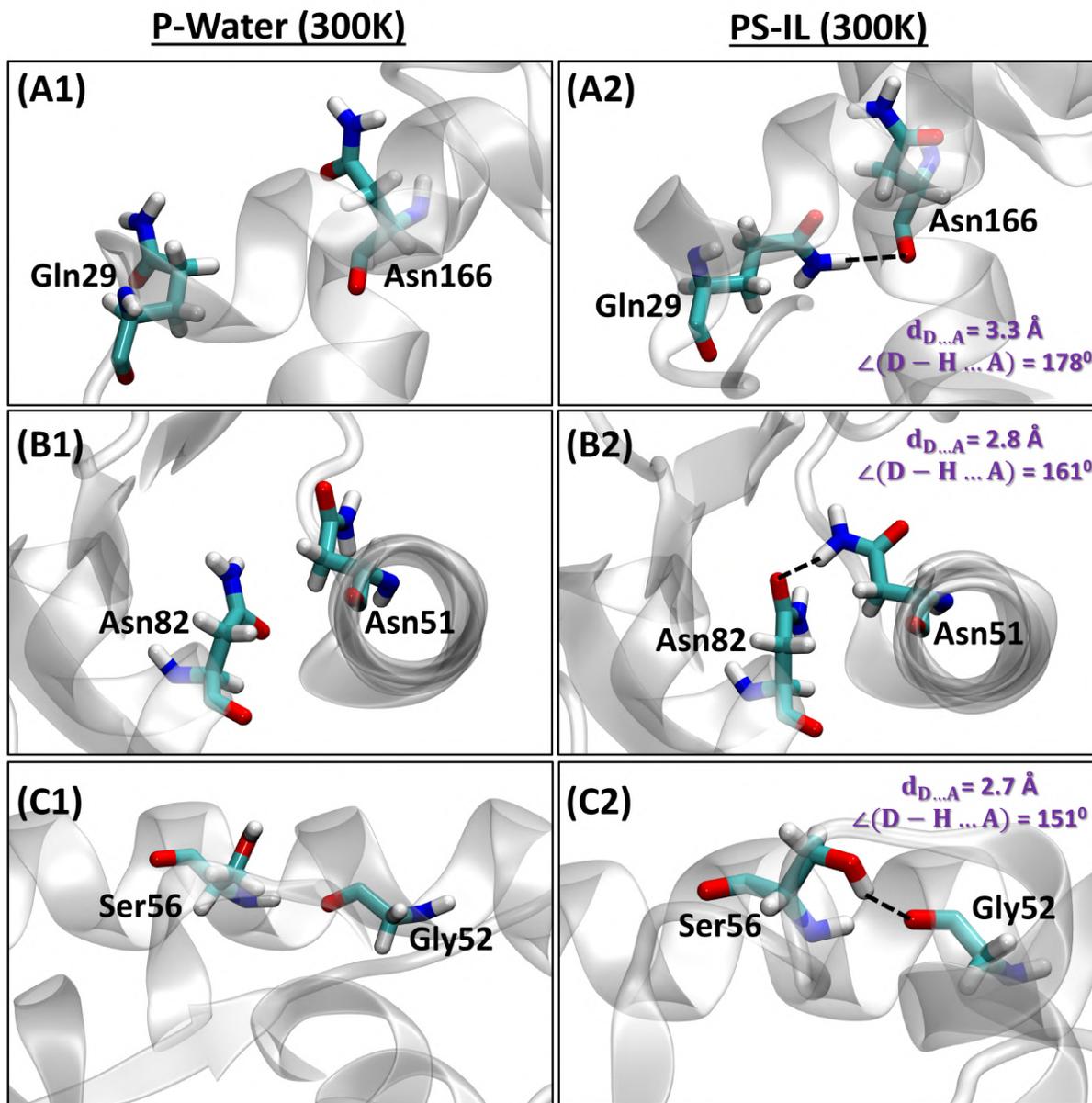


Figure S20: Snapshots of three residue pairs of LipA are shown (A, B, and C), which form H-bonds in the PS-IL but not in the P-Water, and involve at least one polar residue. One of the polar residues of a pair (Gln29, Asn51, and Ser56 in panels A2, B2, and C2, respectively) reorient itself towards the protein in the PS-IL to form intra-protein H-bonds.

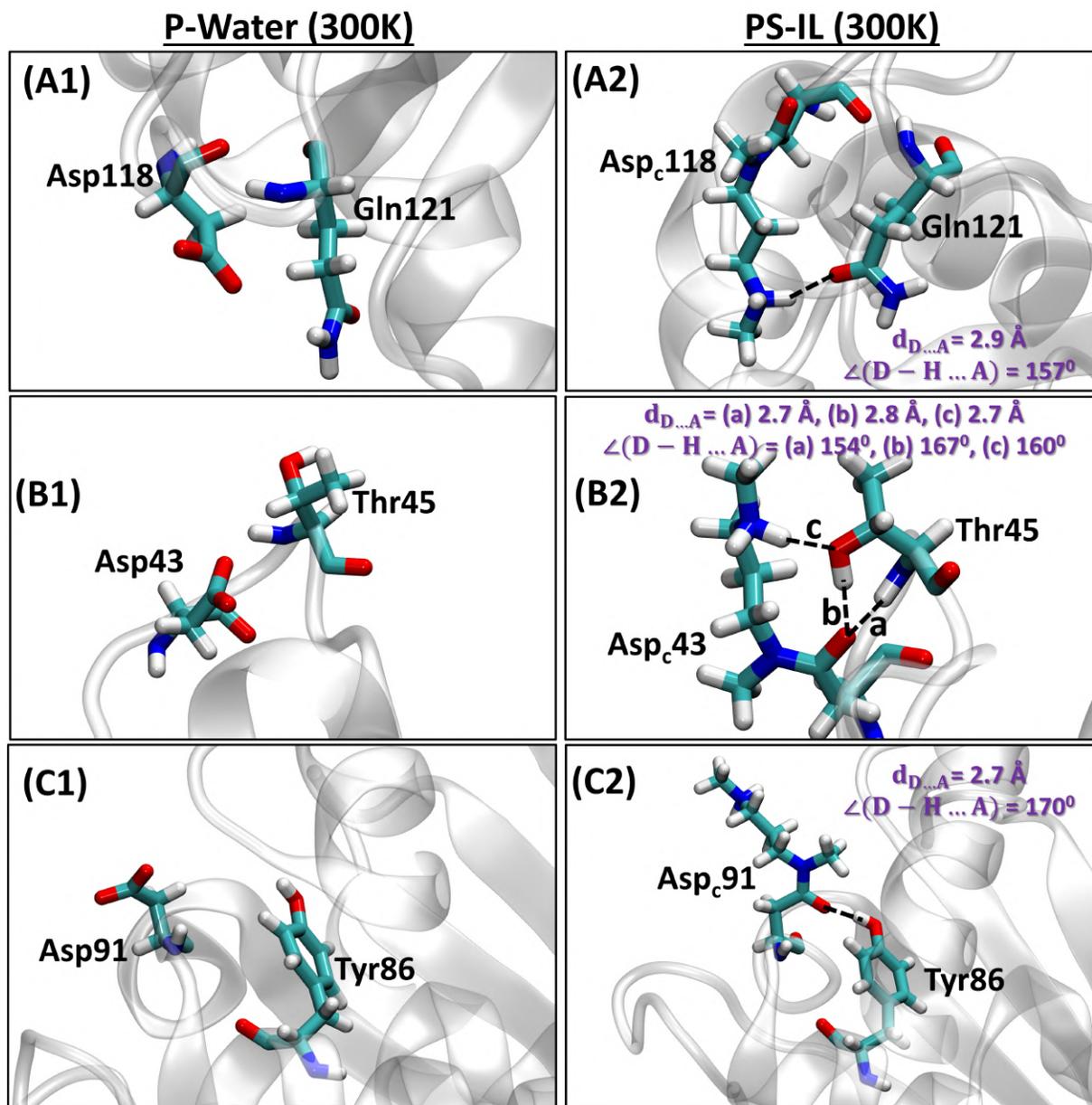


Figure S21: Snapshots of three residue pairs of LipA are shown (A, B, and C), which form H-bonds in the PS-IL but not in the P-Water, and involve one polar residue and one cationized residue (negatively charged residue (Asp/Glu) in P-Water). The three letter code 'Asp<sub>c</sub>' is used for the cationized Aspartate. These H-bonds show the formation of additional hydrogen bonding partners for the polar residues as a consequence of cationization and surfactant coating in the PS-IL of LipA.

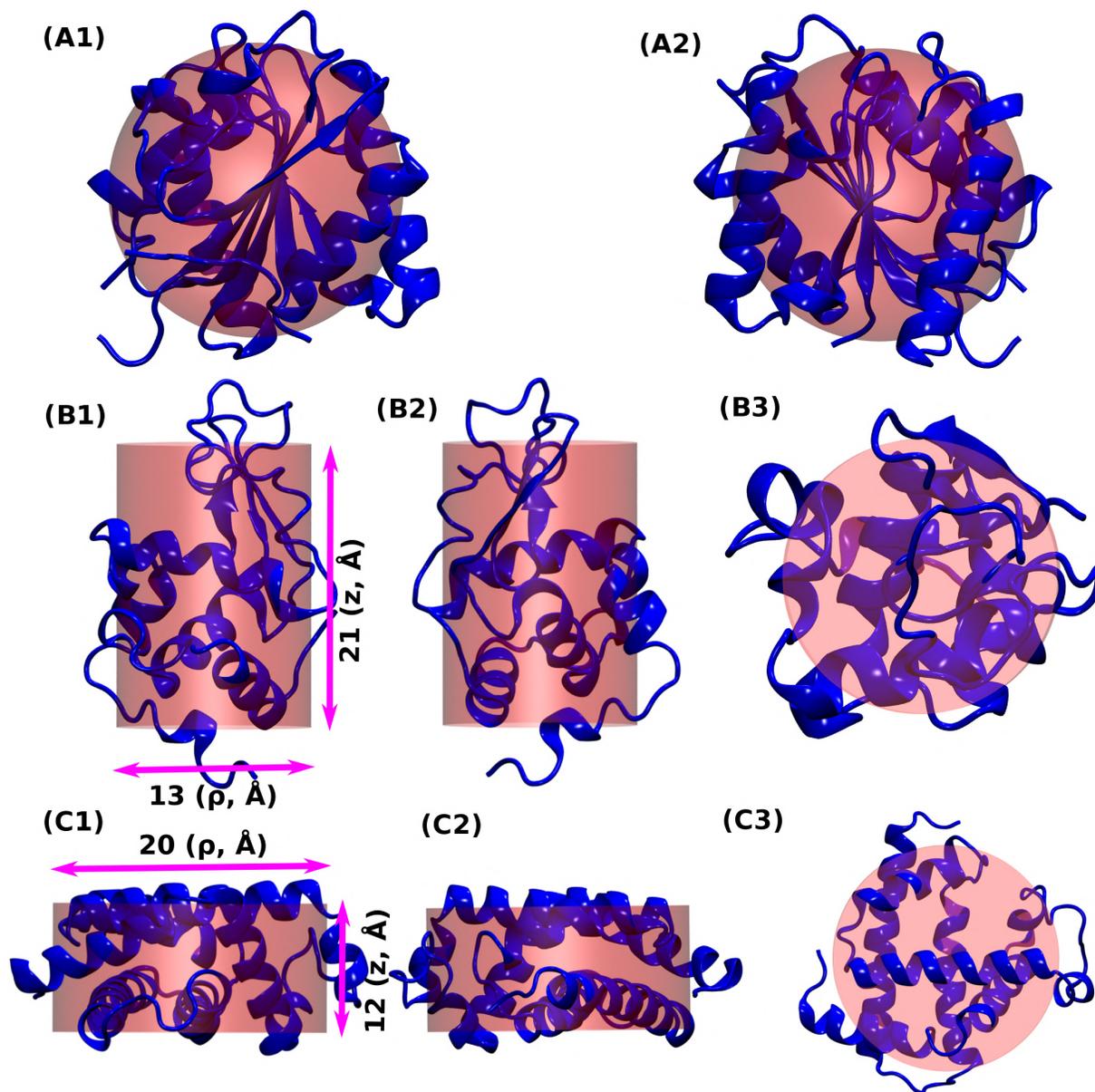


Figure S22: The shapes of the three proteins studied here. Two views of the spherical shape of LipA (A1 & A2), two side views (B1 & B2) and one top view (B3) of the cylindrical shape and disc shape of Lysozyme and Myoglobin, respectively. Values of the semi-major and semi-minor axes of all the proteins are tabulated in SI Table S6. The distances shown with magenta arrows in panels B1 and C1 are the major and minor axes lengths of Lysozyme and Myoglobin.

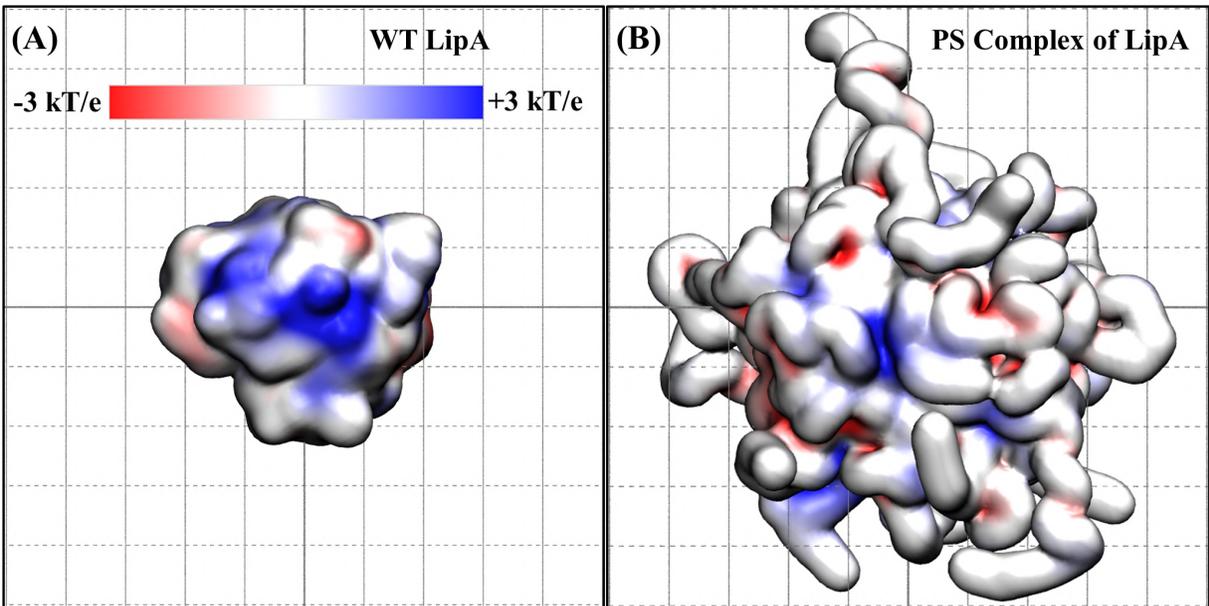


Figure S23: Another view of Figure 6 of the main manuscript with the same color scale.

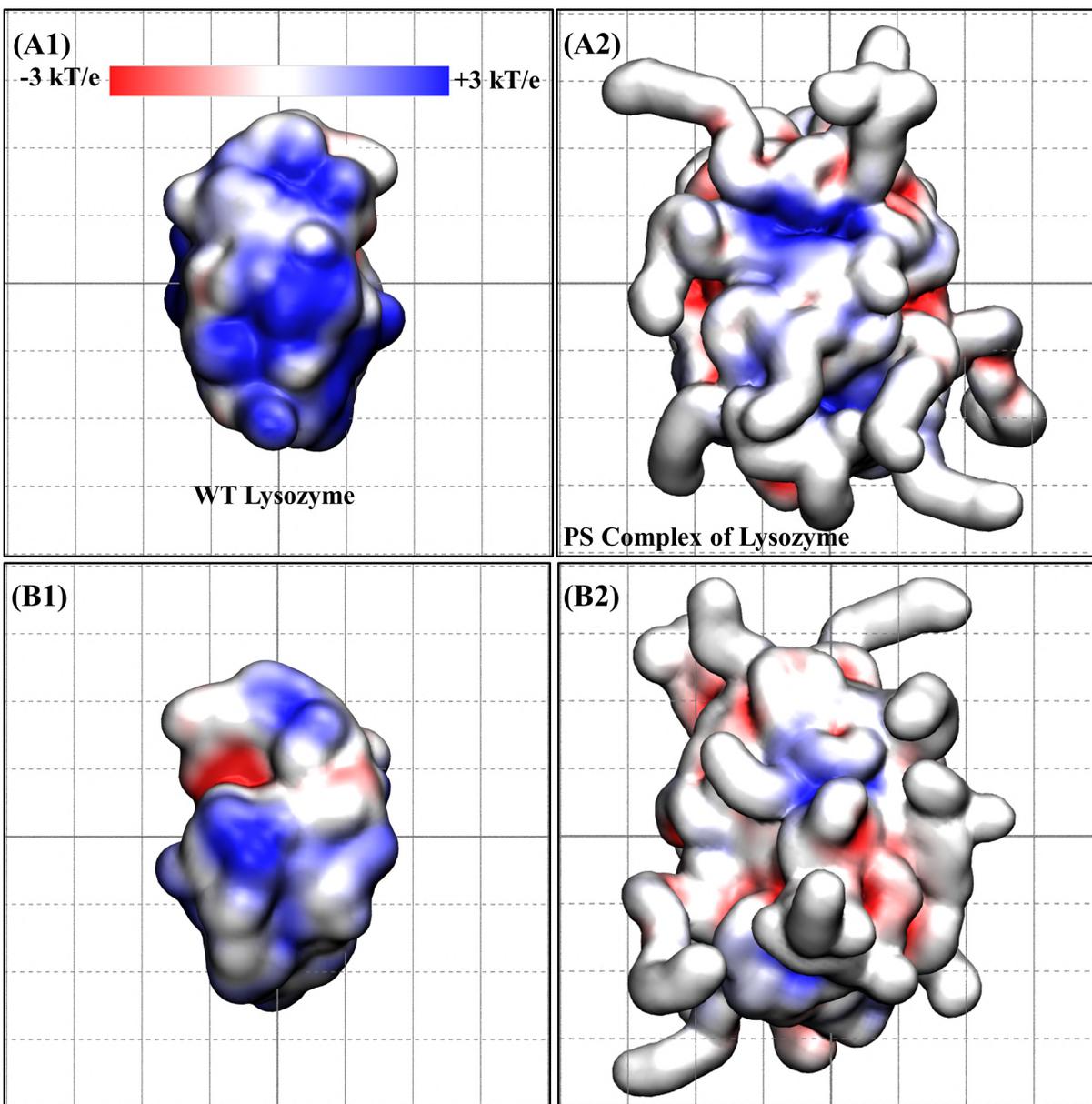


Figure S24: Two views of the electrostatic potential map of (A1 & B1) WT Lysozyme and (A2 & B2) Protein-Surfactant complex from PS-IL of Lysozyme. The colormap used here is Red-White-Blue with red and blue being negative and positive potential, respectively. It is clear from the two panels that Protein-Surfactant complex (B) has more white surface than WT protein (A).

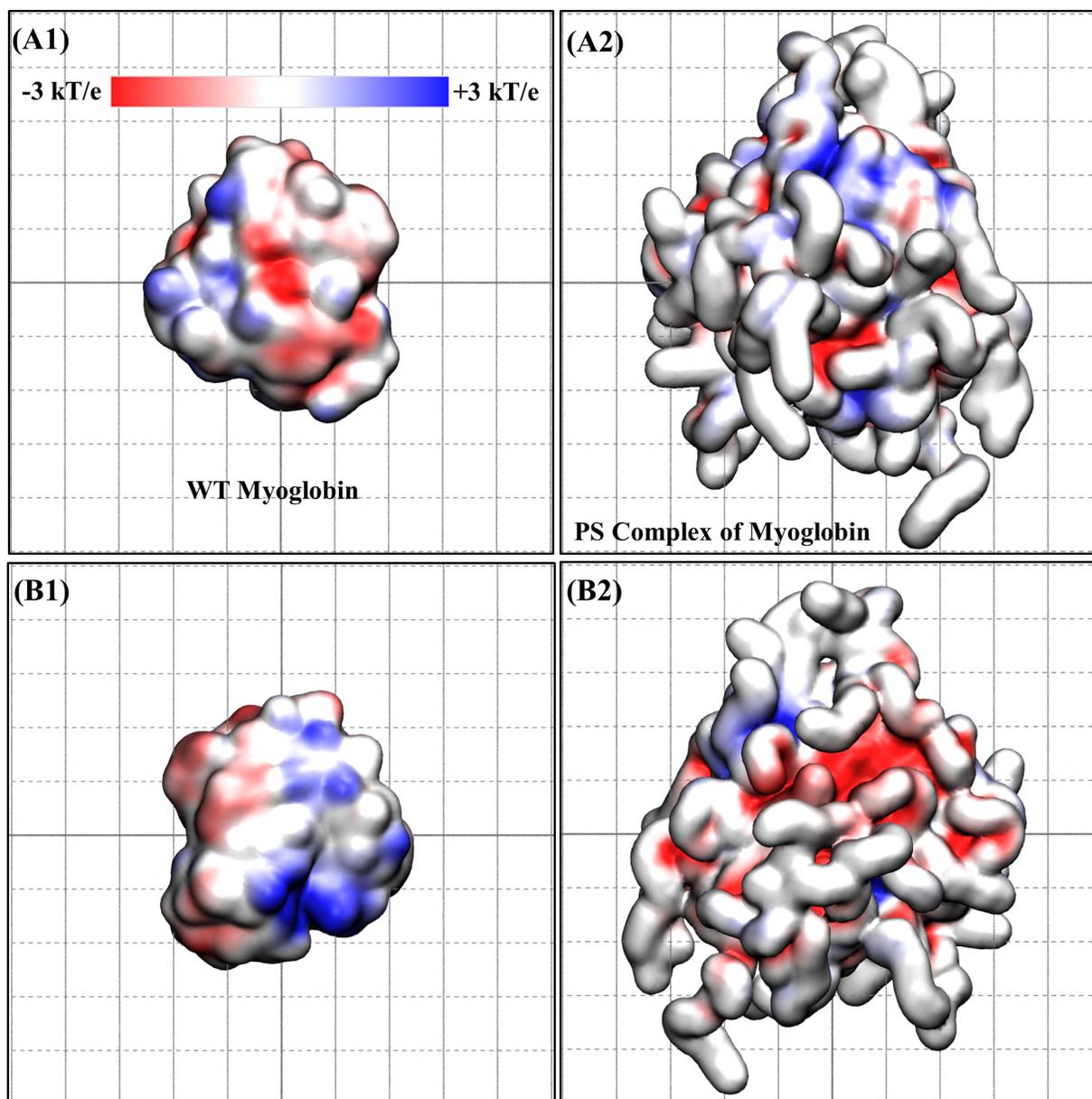


Figure S25: Two views of the electrostatic potential map of (A1 & B1) WT Myoglobin and (A2 & B2) Protein-Surfactant complex from PS-IL of Myoglobin. The colormap used here is Red-White-Blue, with red and blue being negative and positive potential, respectively. It is clear from the two panels that the Protein-Surfactant complex (B) has a more white surface than WT protein (A).

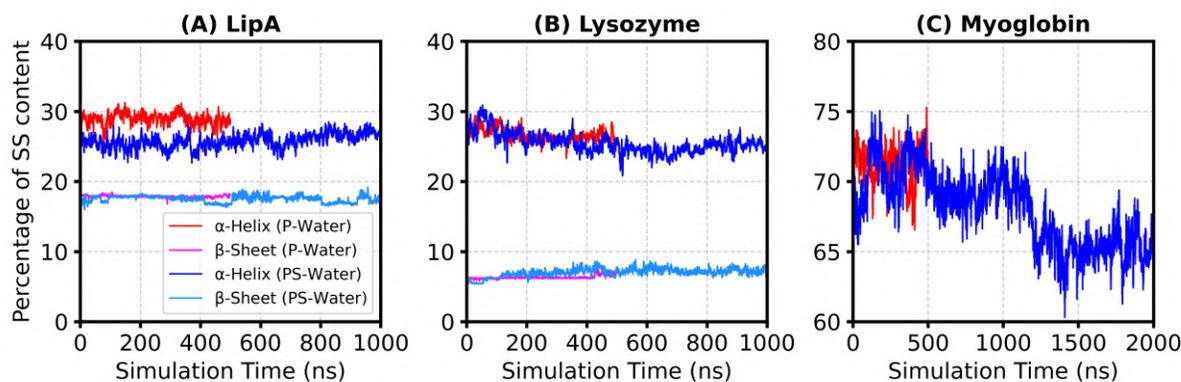


Figure S26: Evolution of the secondary structure contents in the P-Water and PS-Water forms of (A) LipA, (B) Lysozyme and (C) Myoglobin, at 300 K. The secondary structure contents in the PS-Water do not differ much from the P-Water in the case of LipA and Lysozyme. In contrast, there is a decrease in the  $\alpha$ -helical content in the PS-Water form of Myoglobin.

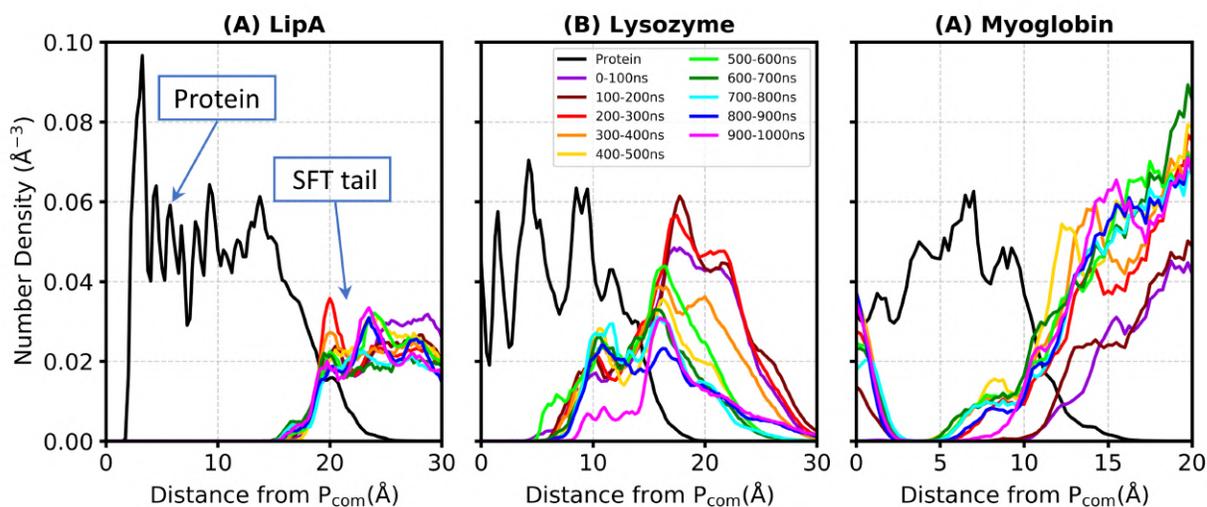


Figure S27: Time evolution of the nDP of the alkyl tail (heavy atoms) of surfactants shown for ten 100ns blocks (total 1000ns) for the PS-Water forms of (A) LipA, (B) Lysozyme and (C) Myoglobin. The nDPs of the heavy atoms of proteins are also shown in black color. The tail parts of the surfactants are observed to insert themselves into the protein for the Myoglobin case. The number density of SFT tail, close to the  $P_{com}$ , increases with simulation time (insertion starts in the 100-200ns range, maroon color curve). Such insertion of the alkyl tail of the surfactants are not seen in the PS-water cases of LipA and Lysozyme.

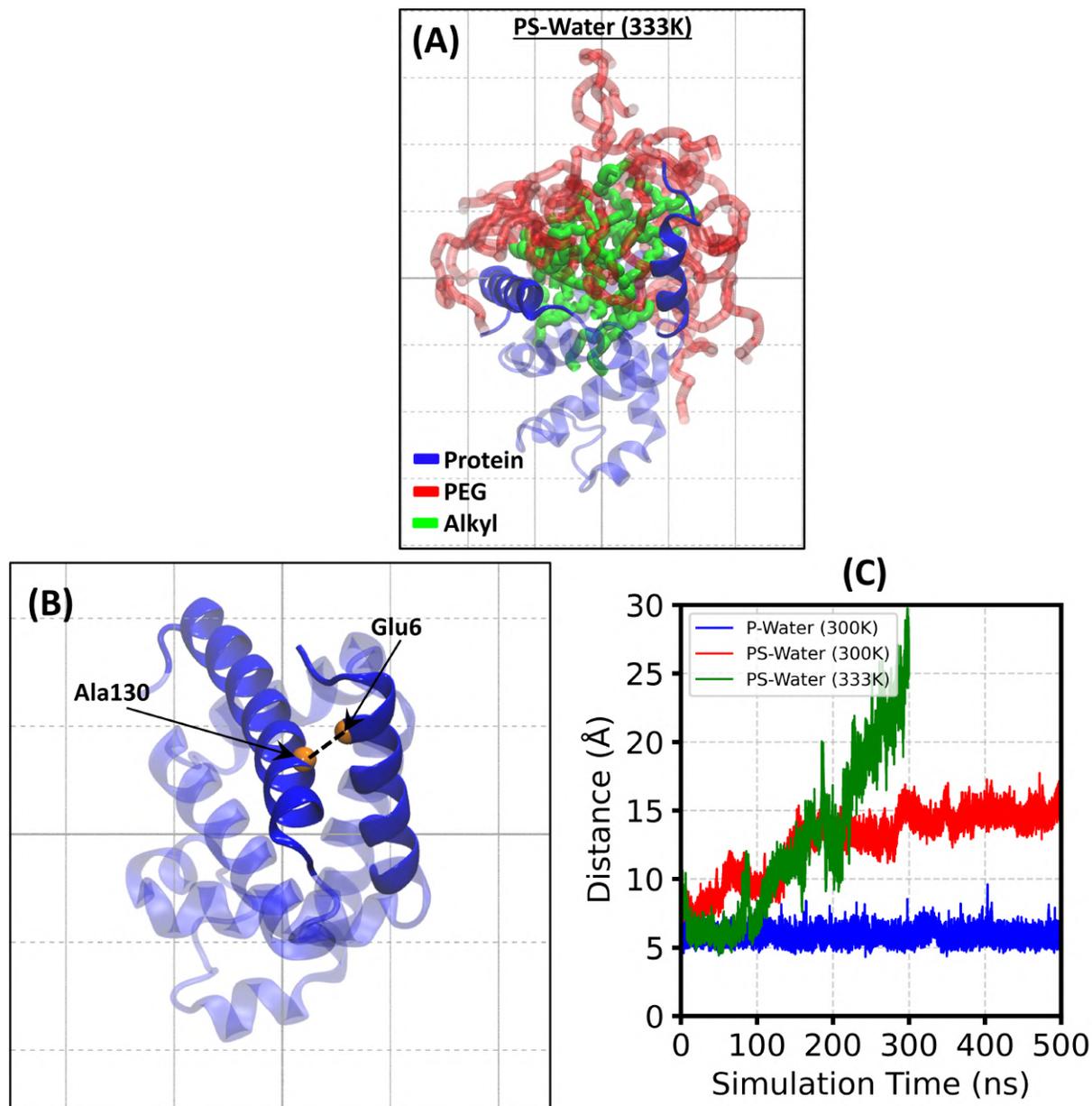


Figure S28: (A) A snapshot of the PS-Water system at 333 K demonstrates the increased entry of the alkyl parts of the surfactants into the hydrophobic core of Myoglobin, further separating Helix-1 and Helix-2. (B) A snapshot of the crystal structure WT Myoglobin showing close contact ( $\sim 6$  Å) between C $\alpha$  of Glu6 and Ala130 (Orange balls). (C) The distance between C $\alpha$  atoms of Glu6 (part of the Helix-1 defined in Figure 5 of manuscript) and Ala130 (part of the Helix-2 defined in Figure 5 of manuscript) in various systems, shows the unfolding of Myoglobin by disruption of the contacts in the region containing Glu6 and Ala130 in the PS-Water systems.

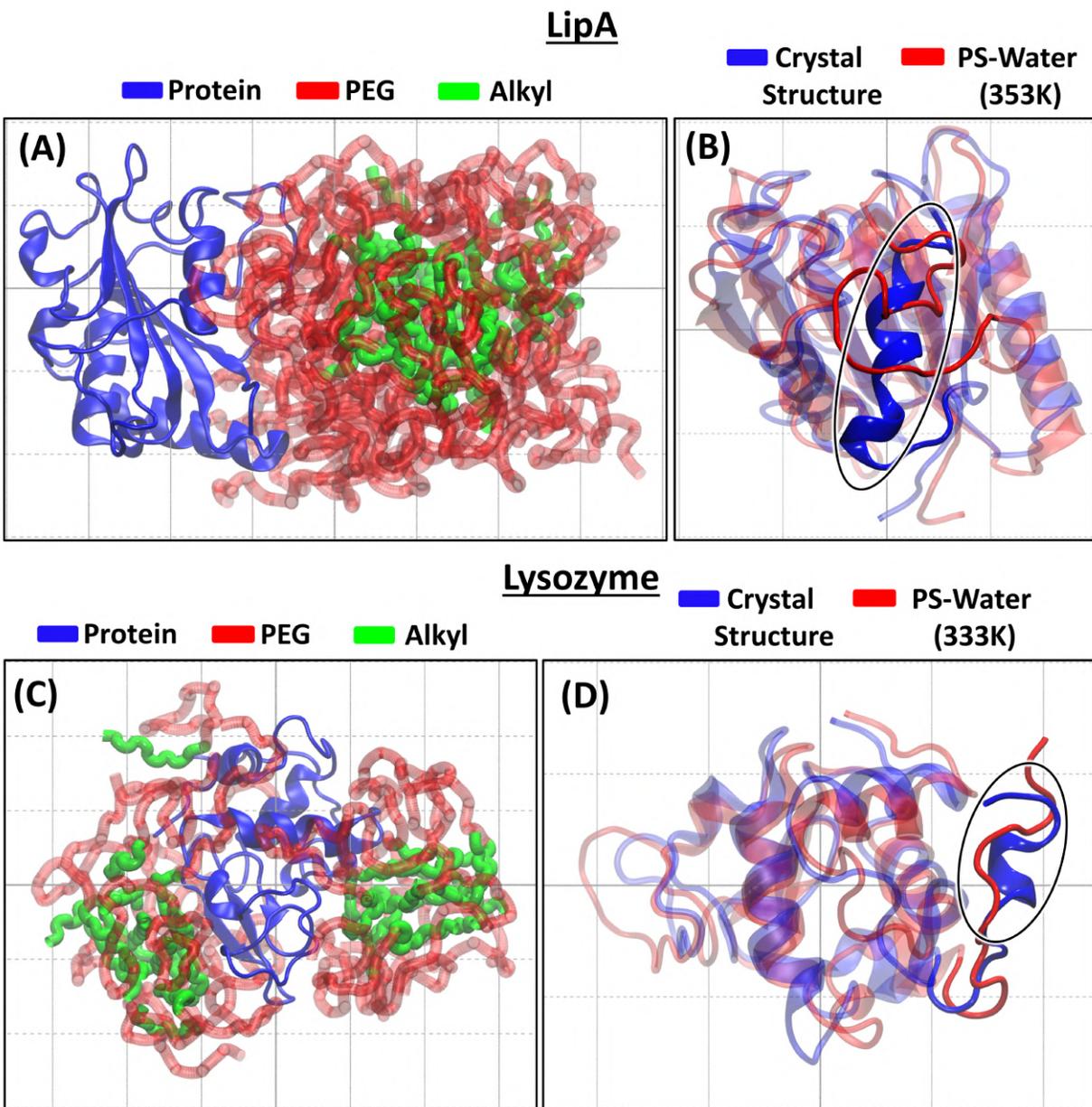


Figure S29: (A) A snapshot of the PS-Water system of LipA at 353 K (config.2 in the SI Figure S7C, where RMSD reaches the highest value), (B) Alignment of LipA from PS-Water system at 353 K with the crystal structure of LipA showing that the increase in RMSD value in config.2 of SI Figure S7C is because of the secondary structure loss highlighted inside the black ellipse (resid Asn18 to Gly30). (C) A snapshot of the PS-Water system of Lysozyme at 333 K (config. 2 in the SI Figure S11B, where RMSD reaches the highest value), (D) Alignment of Lysozyme from PS-Water system at 333 K with the crystal structure of Lysozyme showing that the denaturation of the helix (resid Gln121 to Arg125) highlighted inside the black ellipse is also responsible (other than flexible loops) for the increase in the RMSD.

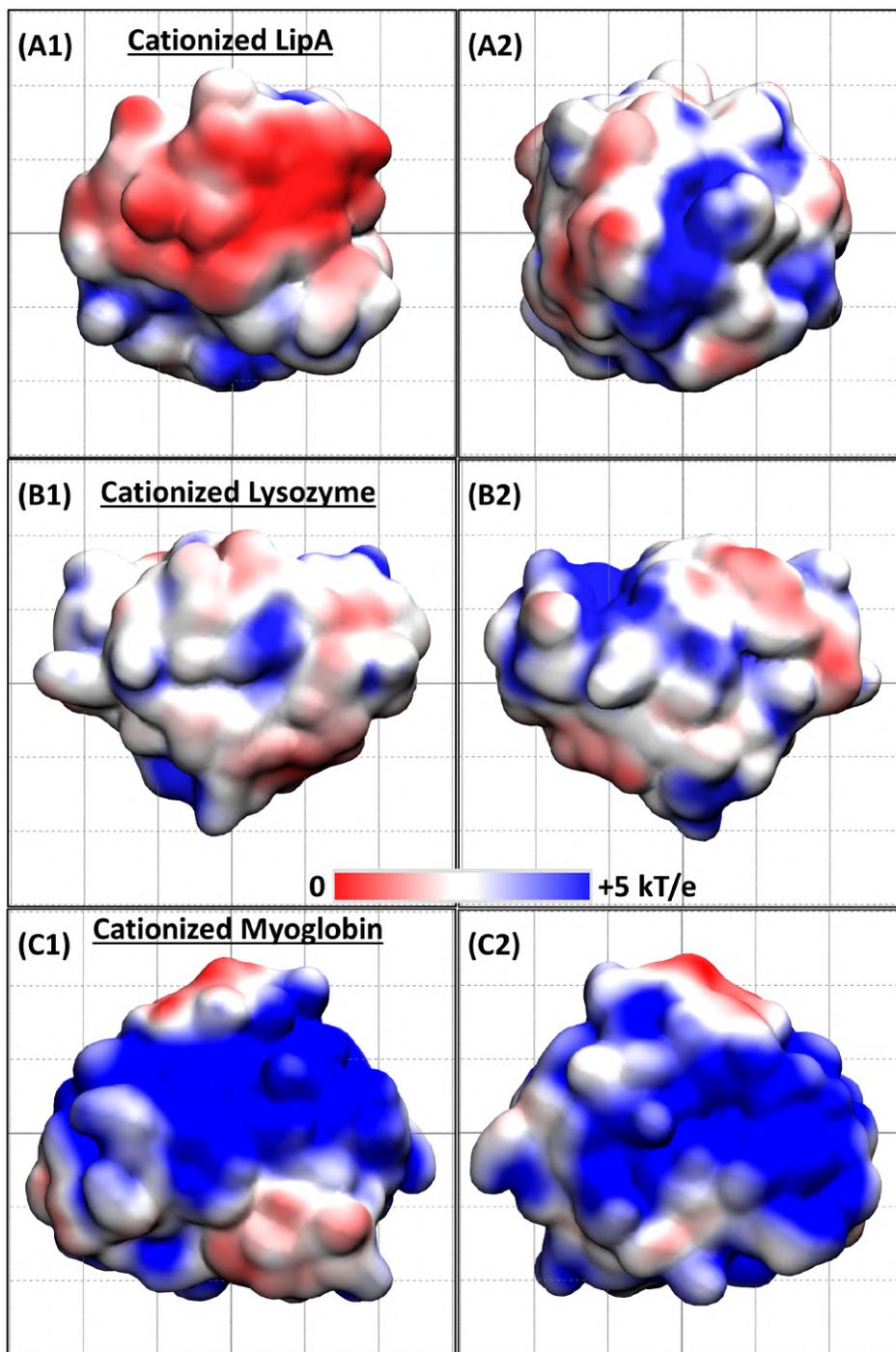


Figure S30: Two different views of the electrostatic potential maps of (A1 & A2) cationized LipA, (B1 & B2) cationized Lysozyme, and (C1 & C2) cationized Myoglobin. These maps clearly demonstrate the decrease in the extent of non-polar regions on the surface of cationized Myoglobin compared to that in cationized LipA and cationized Lysozyme.

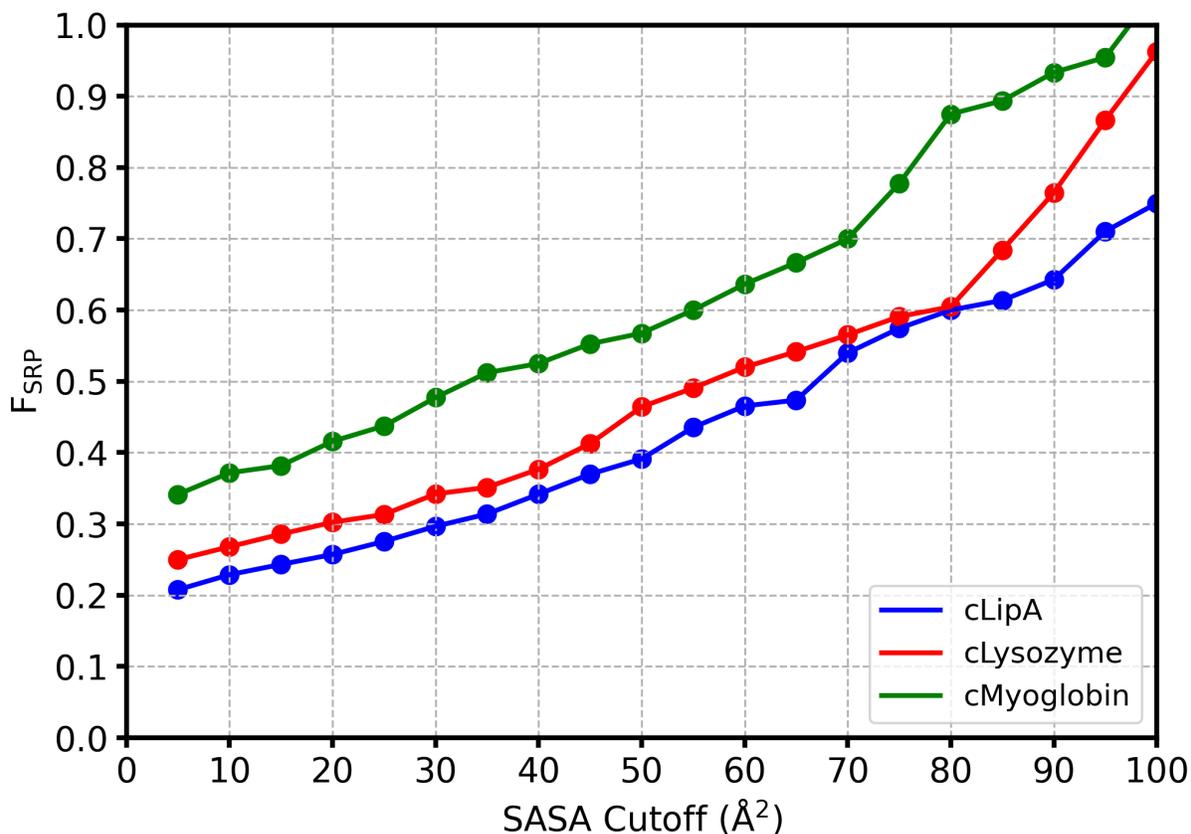


Figure S31: Fraction of surface residues with positively charged side chains ( $F_{SRP}$ ) is defined as the number of positively charged surface residues divided by the total number of surface residues. A residue is defined as surface residue if its SASA is more than a certain cutoff. This figure displays the  $F_{SRP}$  as a function of the SASA cutoff for cationized LipA (cLipA), cationized Lysozyme (cLysozyme), and cationized Myoglobin (cMyoglobin). It shows that in the case of cMyoglobin, the surface has  $\sim 10\%$  higher positively charged residues than that of cLipA and cLysozyme surfaces, which is also independent of the cutoff value.

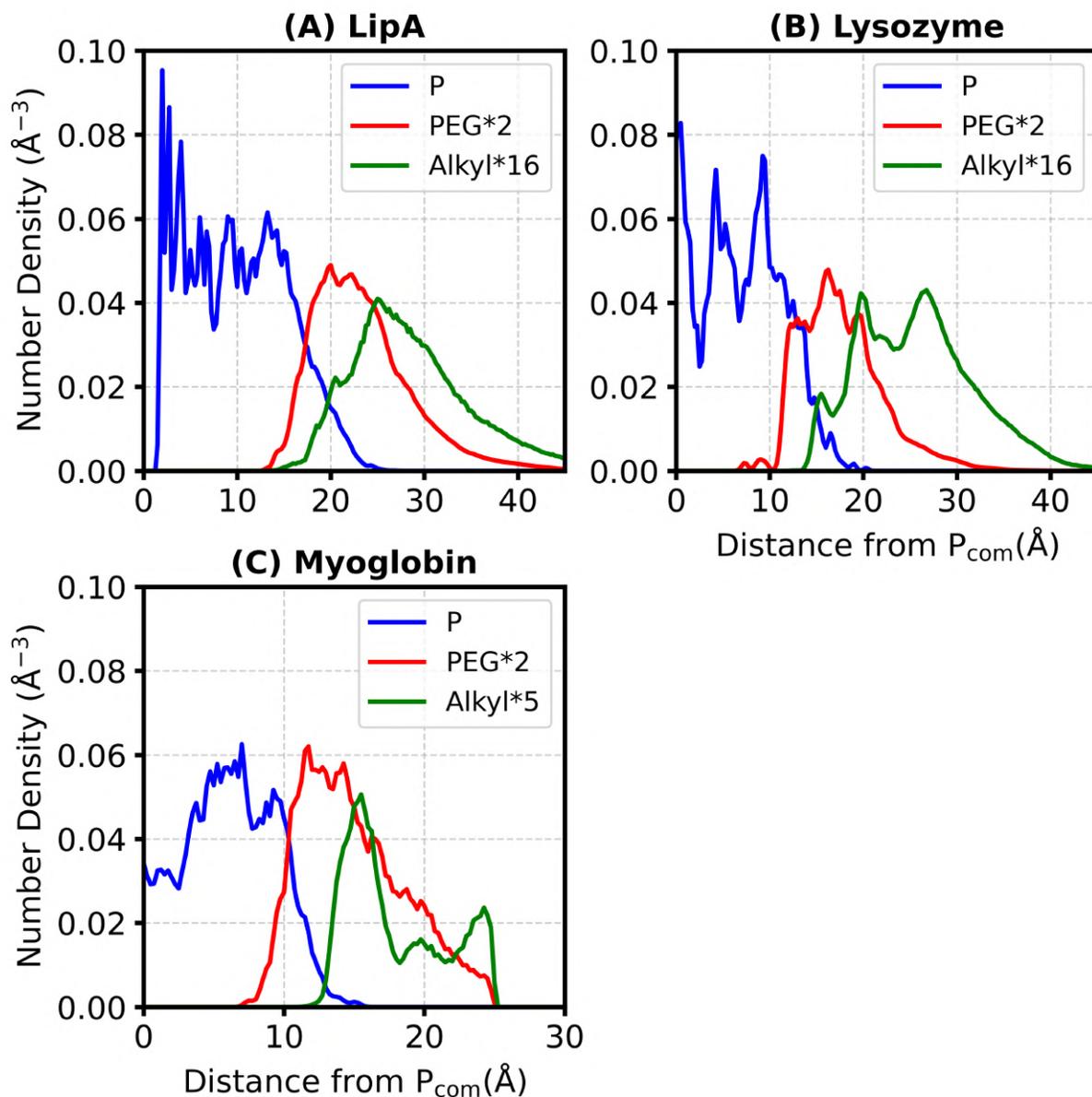


Figure S32: Number density profile of heavy atoms of protein (blue), PEG part (red) and alkyl part of the surfactant (green) in the PS-IL systems for (A) LipA, (B) Lysozyme and (C) Myoglobin. The PEG parts of surfactant are present closer to the protein than the alkyl part, similar to that in the pure SFPL.<sup>21</sup>

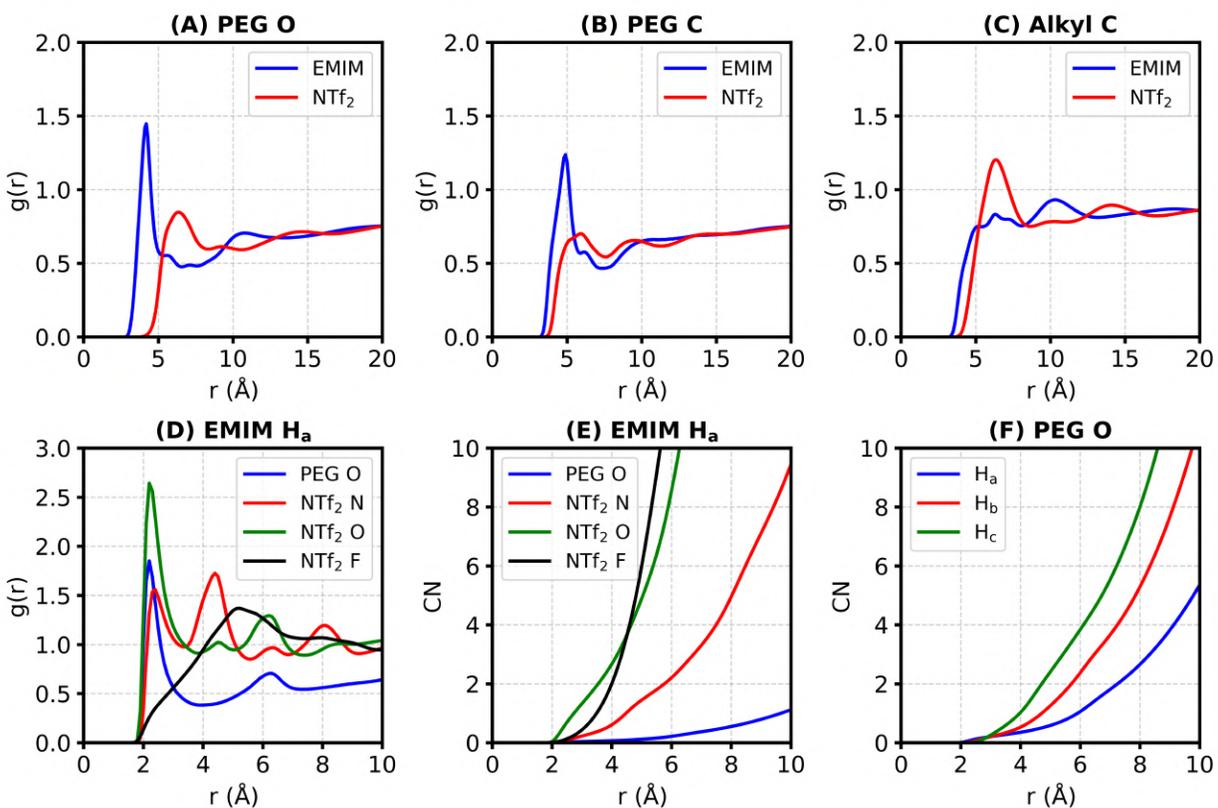


Figure S33: RDFs of (A) PEG Oxygen, (B) PEG Carbon, (C) Alkyl Carbon of surfactant with the EMIM ring center of geometry (COG) and NTf<sub>2</sub> center of mass (COM). (D) RDF of EMIM H<sub>a</sub> with various hydrogen bond acceptors, averaged over the PS-IL of all the independent configurations of the three proteins. Refer to SI Figure S18B for naming convention of EMIM and NTf<sub>2</sub>. (E) The coordination number plot corresponding to the RDF presented in panels D, and (F) the coordination number plot corresponding to the RDF presented in the Figure 7A of manuscript.

## References

- (1) Brogan, A. P.; Hallett, J. P. Solubilizing and stabilizing proteins in anhydrous ionic liquids through formation of protein–polymer surfactant nanoconstructs. *Journal of the American Chemical Society* **2016**, *138*, 4494–4501.
- (2) Brogan, A. P.; Bui-Le, L.; Hallett, J. P. Non-aqueous homogenous biocatalytic conversion of polysaccharides in ionic liquids using chemically modified glucosidase. *Nature Chemistry* **2018**, *10*, 859–865.
- (3) Mondal, A.; Balasubramanian, S. Quantitative Prediction of Physical Properties of Imidazolium Based Room Temperature Ionic Liquids through Determination of Condensed Phase Site Charges: A Refined Force Field. *The Journal of Physical Chemistry B* **2014**, *118*, 3409–3422.
- (4) Ohno, H.; Suzuki, C.; Fukumoto, K.; Yoshizawa, M.; Fujita, K. Electron Transfer Process of Poly(ethylene oxide)-Modified Cytochrome c in Imidazolium Type Ionic Liquid. *Chemistry Letters* **2003**, *32*, 450–451.
- (5) Nakashima, K.; Maruyama, T.; Kamiya, N.; Goto, M. Homogeneous enzymatic reactions in ionic liquids with poly(ethylene glycol)-modified subtilisin. *Org. Biomol. Chem.* **2006**, *4*, 3462–3467.
- (6) Dennington *et al.*, R. GaussView, version 5. **2009**,
- (7) Frisch *et al.*, M. J. Gaussian 09, Revision B.01. 2009.
- (8) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *Journal of Physical Chemistry* **1993**, *97*, 10269–10280.
- (9) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and

- testing of a general Amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157–1174.
- (10) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling* **2006**, *25*, 247–260.
- (11) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE-Antechamber python parser interface. *BMC Research Notes* **2012**, *5*, 1–8.
- (12) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Van Gunsteren, W. F.; Mark, A. E. Peptide folding: when simulation meets experiment. *Angewandte Chemie International Edition* **1999**, *38*, 236–240.
- (13) Martinez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* **2009**, *30*, 2157–2164.
- (14) D’Agostino, C.; Mantle, M. D.; Mullan, C. L.; Hardacre, C.; Gladden, L. F. Diffusion, Ion Pairing and Aggregation in 1-Ethyl-3-Methylimidazolium-Based Ionic Liquids Studied by <sup>1</sup>H and <sup>19</sup>F PFG NMR: Effect of Temperature, Anion and Glucose Dissolution. *ChemPhysChem* **2018**, *19*, 1081–1088.
- (15) Van Pouderoyen, G.; Eggert, T.; Jaeger, K.-E.; Dijkstra, B. W. The crystal structure of Bacillus subtilis lipase: a minimal  $\alpha/\beta$  hydrolase fold enzyme. *Journal of Molecular Biology* **2001**, *309*, 215–226.
- (16) Wang, J.; Dauter, M.; Alkire, R.; Joachimiak, A.; Dauter, Z. Triclinic lysozyme at 0.65 Å resolution. *Acta Crystallographica Section D: Biological Crystallography* **2007**, *63*, 1254–1268.

- (17) Hersleth, H.-P.; Hsiao, Y.-W.; Ryde, U.; Görbitz, C. H.; Andersson, K. K. The crystal structure of peroxymyoglobin generated through cryoradiolytic reduction of myoglobin compound III during data collection. *Biochemical Journal* **2008**, *412*, 257–264.
- (18) Martnez-Rosell, G.; Giorgino, T.; De Fabritiis, G. PlayMolecule ProteinPrepare: A Web Application for Protein Preparation for Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling* **2017**, *57*, 1511–1516.
- (19) Perriman, A. W.; Cölfen, H.; Hughes, R. W.; Barrie, C. L.; Mann, S. Solvent-free protein liquids and liquid crystals. *Angewandte Chemie International Edition* **2009**, *48*, 6242–6246.
- (20) Perriman, A. W.; Mann, S. Liquid proteins-A new frontier for biomolecule-based nanoscience. *ACS Nano* **2011**, *5*, 6085–6091.
- (21) Behera, S.; Das, S.; Balasubramanian, S. An atomistic view of solvent-free protein liquids: the case of Lipase A. *Physical Chemistry Chemical Physics* **2021**, *23*, 7302–7312.
- (22) Brogan, A. P.; Sessions, R. B.; Perriman, A. W.; Mann, S. Molecular dynamics simulations reveal a dielectric-responsive coronal structure in protein–polymer surfactant hybrid nanoconstructs. *Journal of the American Chemical Society* **2014**, *136*, 16824–16831.
- (23) Brogan, A. P.; Sharma, K. P.; Perriman, A. W.; Mann, S. Isolation of a highly reactive  $\beta$ -sheet-rich intermediate of lysozyme in a solvent-free liquid phase. *The Journal of Physical Chemistry B* **2013**, *117*, 8400–8407.
- (24) Perriman, A. W.; Brogan, A. P.; Cölfen, H.; Tsoureas, N.; Owen, G. R.; Mann, S. Reversible dioxygen binding in solvent-free liquid myoglobin. *Nature Chemistry* **2010**, *2*, 622–626.

- (25) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function and Bioinformatics* **2010**, *78*, 1950–1958.
- (26) Fletcher, R. Function minimization by conjugate gradients. *The Computer Journal* **1964**, *7*, 149–154.
- (27) Van Gunsteren, W. F.; Berendsen, H. J. A Leap-Frog Algorithm for Stochastic Dynamics. *Molecular Simulation* **1988**, *1*, 173–185.
- (28) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **1995**, *91*, 43–56.
- (29) Price, D. J.; Brooks, C. L. A modified TIP3P water potential for simulation with Ewald summation. *Journal of Chemical Physics* **2004**, *121*, 10096.
- (30) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.
- (31) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: A Linear Constraint Solver for molecular simulations. *Journal of Computational Chemistry* **1997**, *18*, 1463–1472.
- (32) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *Journal of Chemical Physics* **2007**, *126*, 014101.
- (33) Berendsen, H. J.; Postma, J. P.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, *81*, 3684–3690.
- (34) Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing In Science & Engineering* **2007**, *9*, 90–95.

- (35) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38.
- (36) Wiecek, M.; Hospital, A.; Bayarri, G.; Czub, J.; Orozco, M. Molywood: streamlining the design and rendering of molecular movies. *Bioinformatics* **2020**, *36*, 4660–4661.
- (37) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry* **2011**, *32*, 2319–2327.
- (38) Zhou, Y.; Jones, N. C.; Pedersen, J. N.; Prez, B.; Hoffmann, S. V.; Petersen, S. V.; Pedersen, J. S.; Perriman, A.; Kristensen, P.; Gao, R.; Guo, Z. Insight into the Structure and Activity of Surface-Engineered Lipase Biofluids. *ChemBioChem* **2019**, *20*, 1266–1272.