Temperature dependence of protein dynamics simulated with three different water models


It is advisable to refer to the publisher’s version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1021/ct9006508

Publisher: American Chemical Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
Temperature dependence of protein dynamics simulated with various water models

Dennis C. Glass¹, M. Krishnan¹, David R. Nutt², and Jeremy C. Smith¹

¹University of Tennessee/ORNL Center for Molecular Biophysics
Oak Ridge National Laboratory, Oak Ridge, Tennessee

²The University of Reading, Reading, UK

Abstract

The effect of variation of the water model on temperature dependent protein and hydration water dynamics is examined by performing molecular dynamics simulations of myoglobin with the TIP3P, TIP4P, and TIP5P water models and the CHARMM protein force field at temperatures between 20 and 300 K. The atomic mean-square displacements, solvent reorientational relaxation times, pair angular correlations between surface water molecules and time-averaged structures of the protein are all found to be similar and the protein dynamical transition is described almost indistinguishably for the three water potentials.
The results provide further evidence that variation of the water model without loss of accuracy can be considered, opening up the possibility of choosing the water model employed in protein simulations according to the properties to be investigated.

1. Introduction

Water plays an important role in many chemical and biological processes [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. For example, hydration water strongly influences the three-dimensional structure, dynamics, and function of proteins [2]. Water-protein interactions modify the free energy landscape that determines the folding, structure and stability of proteins [3, 4, 5, 6]. Internal protein dynamics, which are required for biological function, are dependent on the level of hydration [7] and dynamical coupling between the protein and water influences the conformational flexibility of proteins [8, 9, 10, 11].

Protein hydration water can be grouped into two classes: internal water molecules which can play structural and/or catalytic roles [12, 13], and external surface water molecules. Hydration water is experimentally estimated to account for 10-15 % of the total cell water [14, 15] of which a small fraction of \( \approx 0.1 \% \) is internal water molecules. Properties of water in the external hydration shell are modified compared to the bulk, with, for example, changes in av-
verage density [16, 17, 18] and perturbations in translational and rotational dynamics, and these changes have been extensively investigated using techniques such as neutron scattering [14, 19, 20, 21, 22, 23, 24, 25, 26], nuclear magnetic resonance [27, 28, 29, 30, 31, 32], fluorescence spectroscopy [33], mid-infrared pump-probe spectroscopy [34], and molecular dynamics (MD) simulations [22, 35, 36, 37, 38, 39, 40, 41].

Empirical, molecular mechanics force fields, such as CHARMM [42, 43, 44], AMBER [45], GROMOS [46] and OPLS-AA [47] are widely employed in atomistic MD simulations of biological molecules and, in order to represent the protein-solvent energy accurately, most molecular simulation applications employ explicit water models. A large number of water models is available. However, individual biomolecular force fields have normally been parametrized for use with a single water model (see e.g. [48]) such that, during the parameterisation, care can be taken to correctly balance water-water, water-protein, and protein-protein interactions. Nevertheless, the question arises as to whether the water:protein potentials are sufficiently robust that alternative water models may be employed with any given biomolecular force field without serious loss of accuracy. Flexibility in the choice of the water model may be of particular interest when the system is to be simulated at nonphysiological temperatures or pressures, under which circumstances different water models may exhibit significantly different properties (see e.g. [49]).
and/or when water properties are specifically under investigation in which case a
detailed water potential may be appropriate.

In recent work, effects of varying the water model in molecular mechanics and
dynamics calculations on the hydration of N-methylacetamide (NMA) hydration
and other small solute molecules, and a small protein using the CHARMM force
field [42, 43, 44] were examined [50]. The overall description of solvation and
biomolecular properties were found to be similar for the three models tested: TIP3P,
TIP4P, and TIP5P [49, 51]. The CHARMM protein force field was parameterized
for use with the TIP3P potential. However, the results provide an indication that
molecular simulations with the CHARMM force field may in some cases be per-
formed with water models other than TIP3P.

In the present work, we focus on the effect of varying water models on the
temperature dependence of internal protein motion. Experimental and theoreti-
cal studies have found that proteins undergo a transition in internal dynamics at
$T_g \approx 180 - 220 K$ [52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63], characterized by
a rapid increase in the average protein atomic mean-square displacement above
$T_g$. An additional low-temperature transition ($\sim 150 K$), that has been attributed
to the activation of methyl group rotations, is present also in dehydrated proteins
[64, 65, 66, 67, 68, 69], whereas the transition at $T_g$ is strongly coupled to the sol-
vent dynamics [2, 60, 61, 62, 68, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79]. The dynami-
cal transition at $T_g$ can be eliminated when dehydrating the protein or coating it with bioprotectant [80] and the value of $T_g$ can be shifted by changing the solvent composition [81]. MD simulations have demonstrated that the $T_g$ dynamical transition is driven by translational solvent dynamics [61, 82]. For some proteins, the $T_g$ dynamical transition has also been correlated with protein activity [57, 59, 71, 83, 84] although activity has also been observed for $T < T_g$ [85, 86, 87, 88].

The question arises as to whether the dynamical properties of the protein are affected by the choice of the water model used. To investigate this, we have performed simulations of myoglobin at temperatures ranges from 20 K to 300 K using the TIP3P, TIP4P, or TIP5P potentials and the CHARMM protein force field [42, 43, 44]. The protein dynamical transition was found to be unaffected by changes in the solvent model. Moreover, although the bulk properties of the three water models are markedly different, when interacting with the protein surface, the three water models behave similarly.

2. Methods

The set of models considered here is the TIP3P, TIP4P, and TIP5P family [49, 51]. TIP3P is the standard model in the widely-used CHARMM force fields. However, TIP4P and TIP5P are easy to implement for use with CHARMM and exhibit improved water properties, as described below.
The TIP geometries and associated parameters are shown in Fig. 1 and Tab. 1. In the TIP3P model, charges are placed at the hydrogen positions and on the oxygen, resulting in three interaction sites. A van der Waals term provides additional non-bonding interactions involving the oxygen atoms only [51]. Geometrical parameters were determined according to experimental gas phase values.

TIP3Pm is a slightly modified version of TIP3P, commonly used with CHARMM [42, 43, 44], and including additional van der Waals interaction sites at the hydrogen positions [89, 90]. The effect of these additional terms on the properties of TIP3P has been shown to be small [91]. Therefore, in the following, TIP3P is used to refer to the CHARMM-modified version, TIP3Pm.

The combination of TIP3P with the CHARMM force field has proven to be useful for examining the structure and dynamics of biomolecular systems. However, although TIP3P adequately describes the first hydration shell of bulk water, it lacks accuracy for the second hydration shell, for which the corresponding peak in the oxygen-oxygen radial distribution function is almost completely absent [49, 90]. TIP4P, in which the oxygen charge site is moved along the HOH bisector towards the molecular center of mass, better reproduces experimental distribution functions than TIP3P [51]. In TIP5P, there are two lone-pair interaction sites, L, moved from the oxygen along the HOH bisector away from the hydrogens and symmetrically placed out of the HOH plane with $\angle_{LOL} = 109.47^\circ$ [49]. The TIP4P and TIP5P
oxygen carries no charge. TIP5P, explicitly incorporating tetrahedrality in the water model, is especially successful in reproducing water density over a wide range of temperatures [49].

Here, the CHARMM program package version c33b2 was used to perform MD simulations of hydrated myoglobin with the CHARMM22 force field and TIP3P, TIP4P, or TIP5P water [42, 43, 44]. The 1.15 Å-resolution myoglobin structure 1A6G, taken from the RCSB Protein Data Bank (www.pdb.org) [92], was used as the starting protein configuration. The model was constructed as in Ref. [61] to mimic a hydrated powder sample and thus model the experimental neutron scattering setup of Ref. [75], the data from which serve as a reference here. To do this, the protein was solvated by placing it in a box of water, retaining only those 492 molecules closest to the protein. The temperature dependent dynamical properties in the present simulations were found to be very similar to those derived from NPT simulations in solution on the same myoglobin structure [69].

Electrostatic interactions were truncated using a shift function with a 12 Å cutoff, and a switch function was used for truncation of van der Waals interactions between 10 and 12 Å. The SHAKE algorithm was used to constrain all bond lengths involving hydrogens [93]. The structures were energy minimized using 600 steepest descent steps and 2500 conjugate gradient steps with the protein atoms fixed, then with fixed solvent allowing the protein to move, and finally without any constraints.
After heating to the desired temperature in steps of 5 K every 1000 dynamics steps, the system was equilibrated for 400 ps. Subsequently, 1 ns production runs were performed in the NVT ensemble at temperatures from 20 to 300 K in intervals of 20 K and in smaller intervals of 10 K between 120 and 240 K. The system was kept at constant temperature using the weak coupling algorithm of Ref. [94]. A time step of 1 fs was used for the integration of the equations of motion. Coordinates and velocities were saved every 50 steps. The simulation protocol was the same for all temperatures and water models.

In order to examine in more detail the influence of the water model on the time averaged protein structure at 300 K, ten additional simulations of 1 ns length were performed for each water model, starting with different, randomly-chosen, initial velocity assignments.

To avoid potential artifacts no restraining potential was applied to the water molecules to prevent evaporation. Subsequently for $T > 260 K$, a small number of molecules evaporated from the water shell surrounding the protein and these were excluded from all analyses. No evaporation occured for $T < 260 K$, including during extended 7 ns simulations with TIP3P at 180, 220, and 260 K.

In addition to the above calculations, for comparison purposes simulations of bulk water using the TIP water models were performed. For these, GROMACS 4.0 [95] was used to generate a 30 Å cubic box at physiological conditions containing
895 molecules for TIP3P, 886 for TIP4P, and 878 for TIP5P. The electrostatic interactions were treated in the same way as for the above protein simulations, and periodic boundary conditions were applied. The configurations were energy minimized using 600 steepest descent steps, equilibrated for 500 ps at the desired temperatures and, subsequently, 1 ns production runs were performed in the NVT ensemble at temperatures from 20 to 300 K in intervals of 20 K and in smaller intervals of 10 K between 100 and 260 K. The system was kept at constant temperature using the weak coupling algorithm of Ref. [94]. A time step of 2 fs was used. Coordinates were saved every 50 steps. The simulation protocol was again the same for all temperatures and water models.

3. Results

3.1. Time-averaged structures

The influence of the water models on the average structure of myoglobin is first investigated. The time-averaged structures were calculated for all 30 independent simulations at 300 K and were compared using a root mean square deviation (RMSD) per residue, defined here as:

\[
RMSD_{i}^{A,B} = \sqrt{\frac{\sum_{j=1}^{N_i} (x_{ij}^A - x_{ij}^B)^2}{N_i}}
\]

(1)
where $i$ is the residue number, $N_i$ is the number of protein heavy atoms in residue $i$, $A$ and $B$ are any two given time-averaged structures, and $x^A_{i}/x^B_{i}$ denotes the heavy atom coordinates. Convergence was checked by comparing RMSD values from the first and second half of each simulation, which were found to be in close agreement.

Fig. 2 shows the backbone heavy-atom RMSD per residue based on the average over all pairs $(A,B)$ of structures solvated in the same or in different water models. “SELF” refers to RMSD values between protein structures solvated by the same water model whereas “CROSS” refers to deviations in configurations between the models, for example, structures of the protein in TIP3P compared with structures in TIP4P or TIP5P.

Neglecting the five C- and N-terminal residues, the $RMSD^A_{i}/RMSD^B_{i}$, averaged over residues is $0.47 \pm 0.25 \ \text{Å}$, averaged over “SELF” is $0.46 \pm 0.26 \ \text{Å}$, averaged over “CROSS” is $0.49 \pm 0.25 \ \text{Å}$, and the average difference between “SELF” and “CROSS” is $0.035 \pm 0.026 \ \text{Å}$. Therefore, the RMSD per residue is similar for both the “SELF” and “CROSS” data sets, indicating that variation of the water model does not significantly influence the time-averaged protein RMSD.
3.2. Mean-square displacement

The mean-square displacement \( \langle r^2(t) \rangle \) is defined as:

\[
\langle r^2(t) \rangle = \left\langle \frac{1}{N} \sum_{i=0}^{N} (r_i(t + t_0) - r_i(t_0))^2 \right\rangle_{i,t_0}
\]  

(2)

where \( N \) is the number of atoms, \( (r_i(t + t_0) - r_i(t_0)) \) is the displacement of atom \( i \) in time \( t \), and \( \langle \cdot \rangle_{i,t_0} \) represents the ensemble average, approximated as a time average over \( t_0 \) by assuming ergodicity.

The mean-square displacement averaged over the atoms of myoglobin was calculated for all temperatures and water models and the results are shown in Fig. 3 A. The inset shows an expanded view of the low-temperature region, up to \( T \approx 210 \, K \).

The MSD rises linearly for low temperatures until a first change in the gradient at \( T \approx 150 \, K \), which has been attributed to the activation of methyl group rotations [66, 68, 69]. At \( T \approx 220 \, K \), \( \langle r^2(t) \rangle \) exhibits a further increase in the gradient corresponding to the solvent-driven dynamical transition. \( \langle r^2(t) \rangle \), both for \( < 220 \, K \) and \( > 220 \, K \), is similar using all three TIP water models.

Fig. 3 B shows \( \langle r^2(t) \rangle \) of the water molecules in the protein simulations as a function of temperature. The insets B.1 and B.2 give expanded views since the data cover multiple orders of magnitude. The profiles qualitatively resemble Fig. 3 A, i.e., the MSD for the water molecules rises linearly for low temperatures until, interestingly, a first change in slope is seen at \( T \approx 150 \, K \) (Fig. 3 B.2) followed by a
second transition at $T \approx 200K$ (Fig. 3 B.1).

The $\sim 220K$ transition in the protein has been previously observed and the strong coupling between the solvent and protein characterized [55, 58, 61, 71, 74, 76, 78, 82, 96, 97, 98, 99, 100, 101, 102]. However, the low-T, 150 K change in slope for water was unexpected, and the question therefore arises as to whether it would occur independent of the protein. To check this, a set of simulations of pure TIP water was performed, as described in Methods. Fig. 4 A shows the TIP water mean-square displacements as a function of temperature. Comparison between Fig. 3 B and 4 A indicates that, while the dynamical transition behaviour of the protein and protein hydration water are similar, they do not resemble that of the bulk water models. The data for the three TIP models exhibit similar properties in that $\langle r^2(t) \rangle$ is linear for $T \leq T_l$. However unlike in the solution simulations, the temperature at which the slope changes, $T_l$, is markedly different for the three models with $T_l \approx 110K$ for TIP3P, 130 K for TIP4P, and 190 K for TIP5P. Further, again unlike the data for the hydrated protein, the bulk data do not show two distinct changes in slope.

The self-diffusion constant, $D$, was calculated from the linear part of each mean-squared displacement, at i.e. $t > 200$ ps (see Fig. 4 (B)), using the Einstein relation:

$$ \lim_{t \to \infty} \left\langle |r(t_0 + t) - r(t_0)|^2 \right\rangle = 6Dt $$

(3)

where $r(t)$ is the position of the water oxygen atom, and $\langle \cdot \rangle$ denotes averaging over
both time origins $t_0$ and the water molecules. $D$, in units of $10^{-5}\text{cm}^2/\text{s}$, at 300 K is found to be $5.46 \pm 0.02$ (TIP3P), $3.60 \pm 0.06$ (TIP4P), and $2.82 \pm 0.02$ (TIP5P). Corresponding values in the literature vary depending on the simulation setup, but reported values using a similar protocol to that used here (i.e. shift electrostatics) in the NPT ensemble are $5.8 \pm 0.2$ (TIP3P), $3.78 \pm 0.02$ (TIP4P), $2.94 \pm 0.06$ (TIP5P) [103] in good agreement with the present work. For completeness, in Table 2, $D$ is also given for the other temperatures.

3.3. Heterogeneous distribution of anharmonic motions among protein residues.

To quantify the fraction of protein residues that exhibit large anharmonic dynamics, the MSD per residue ($\langle r^2(t) \rangle_i$, where $i$ denotes the residue index) was calculated. $\langle r^2(t) \rangle_i$ is decomposed into harmonic and anharmonic components as follows:

$$\langle r^2(t) \rangle_i = \langle r^2(t) \rangle_{i,\text{harmonic}} + \langle r^2(t) \rangle_{i,\text{anharmonic}}$$ (4)

where $\langle r^2(t) \rangle_{i,\text{harmonic}}$ is obtained by linearly fitting $\langle r^2(t) \rangle_i$ for $T \leq 140K$ and extrapolating to higher $T$.

A residue is denoted as exhibiting anharmonic dynamics if $\langle r^2(t) \rangle_{i,\text{anharmonic}} > (\langle r^2(t) \rangle_{i,\text{harmonic}} + 2\sigma)$, where $\sigma$ is the standard deviation at the onset of anharmonicity (140 K). As a control, it was determined whether the onset of anharmonic-
ity depends on the temperature interval chosen for the estimation of harmonic dynamics: using 20-100 K as the fitting interval and $\sigma_{100K}$, the procedure was found to give similar results.

Fig. 5 shows the temperature dependence of the fraction of residues exhibiting anharmonic dynamics. For $T \leq 140K$, almost all residues exhibit only harmonic motion, with $\langle r^2(t) \rangle_i$ similar to each other, while an abrupt change is evident at higher T, as an increasing fraction of residues exhibits anharmonic dynamics. Even well above the dynamical transition temperature at 260 K, $\sim 25 \%$ of the residues still remain harmonic. This result is consistent with previous simulation work on myoglobin, in which the onset of anharmonicity was found to be gradual with T [62]. Again, when comparing the protein simulations using different water models, there are no statistically significant differences.

3.4. Reorientational relaxation time

The rotational dynamics of water can be characterized by the water dipole orientation autocorrelation function, $C(t)$:

$$C(t) = \langle \vec{e}_i(t + t_0) \cdot \vec{e}_i(t_0) \rangle_{i,t_0}$$

(5)

where $\vec{e}_i(t)$ is a unit vector along the water dipole. For liquids, $C(t)$ decays to zero as the dipole loses its memory of its initial orientation.
Since the structural and dynamic properties of the hydration layer water molecules depend on the heterogeneous surface roughness and charge distribution [22, 25, 40, 41, 104], a multi-exponential decay is expected. From the protein simulations, the reorientational relaxation time of water, $\tau$, was calculated by fitting the following triple exponential function to $C(t)$:

$$C(t) = a_0 e^{-t/\tau_0} + a_1 e^{-t/\tau_1} + (1 - a_0 - a_1) e^{-t/\tau_2}$$

Eq. 6 was found to capture the decay in the target data, whereas the fitting procedure failed or gave worse results for simpler fitting functions. The relaxation time, $\tau$, was derived using the following relation:

$$\tau = a_0 \tau_0 + a_1 \tau_1 + (1 - a_0 - a_1) \tau_2$$

Fig. 6 A shows $C(t)$ together with the fit of Eq. 6 for TIP3P hydration water at different temperatures. $C(t)$ for TIP4P and TIP5P hydration water exhibit similar decay behaviours (not shown). In general, the profiles consist of a fast decay on the picosecond timescale followed by slower dynamics. For most temperatures, $C(t)$ does not fully decay to zero on the present timescale of $\sim 500$ ps. For temperatures $T > 220 K$ (above the dynamical transition temperature) $C(t)$ decays rapidly (< 50
ps) to $\sim 0.5$, but the decay is much slower at lower T.

There is good agreement between the fitted curves of Eq. 6 and the data (Fig. 6 A). Fig. 7 and Tbl. 4 present the temperature dependence of the resulting fitted parameters. At low T ($\leq 180K$), the process associated with relaxation time $\tau_2$ dominates, with a weight factor, $1 - a_0 - a_1 \approx 1$. A decrease of $1 - a_0 - a_1$ (i.e., an increase of $a_0$ and/or $a_1$) is observed for $T > 180K$, indicating the activation of additional relaxation processes. For $T \geq 240K$ the three components are approximately equally weighted.

The temperature dependences of the weights and values of $\tau_0$ and $\tau_1$ behave very similarly to each other, with the weight of the $\tau_1$ component being larger and the values being $\sim 10$ times larger than $\tau_0$. $\tau_0$ and $\tau_1$ have broad maxima at $\sim 260$ K and $210$ K, respectively. For temperatures $T \leq 180K$, the $\tau_1$ component jumps to high values, accompanied by a drop in its weight.

The relaxation times $\tau$, derived by Eq. 7, are given in Fig. 6 B. Relaxation times at 300 K from the full sets of simulations at this temperature are $31 \pm 3$, $35 \pm 3$, and $32 \pm 6$ ps (TIP3P, TIP4P, and TIP5P) and rise by two orders of magnitude as the temperature decreases to 160 K.

At temperatures $T > T^*$, glass-forming liquids exhibit an Arrhenius relaxation mechanism $\tau \propto exp(-E_\beta/k_BT)$ due to the behaviour of the $\beta$-relaxation which, at $T^*$, splits into a fast $\beta$-relaxation and a slow $\alpha$-relaxation [105, 106]. The $\alpha$-
relaxation may often be described with a Vogel-Fulcher-Tammann (VFT) equation over the range $T_g < T < T^*$, i.e.,

$$\tau(T) = \tau_c \exp \left(\frac{A}{T - T_0}\right)$$  \hspace{1cm} (8)

with fitting parameters $\tau_c$, $A$, and $T_0$. Cooperativity of $\beta$-relaxation events has been suggested as the origin of the $\alpha$ relaxation. A similar VFT relationship between $\tau$ and $T$ has been used in glass physics [105], where $T_0$ has been hypothesized to be the Kauzmann temperature and $A = \frac{E_\beta}{R} \frac{T^* - T_0}{T^*}$ with $E_\beta$ as the activation energy for the $\beta$-relaxations. The Kauzmann temperature is the temperature, at which the entropies of the supercooled liquid and the corresponding crystal are in principle be equal [106, 107].

Angell proposed a “fragile-to-strong” classification of liquids in which relaxation times of “strong” liquids follow an Arrhenius trend (e.g. SiO$_2$) whereas “fragile” liquids deviate from such a behaviour [108]. The VFT relation describes well the present data for $T > 160 K$, thus classifying the protein hydration shell water as a “fragile” liquid in Angell’s scheme.

**3.5. Local orientational ordering**

The local orientational ordering of water dipoles can be quantified by the distance-dependent Kirkwood $G$-factor, defined as follows [109]:

\[ G_{K,h}(r) = \left\langle \vec{\mu}_i \cdot \vec{M}(r) \right\rangle \]  \hspace{1cm} (9)

with

\[ \vec{M}(r) = \sum_{r_{ij} \leq r} \vec{\mu}_j \]  \hspace{1cm} (10)

where \( \vec{\mu}_i \) denotes a unit vector in the direction of the dipole moment of molecule \( i \). \( G_{K,h}(r) \) is equal to two if a pair of water dipoles is parallel. Elevated \( G_{K,h}(r) \) therefore corresponds to high angular correlations between water molecules.

Fig. 8 shows \( G^{r_{\text{max}}}_{K,h}(T) \), as a function of temperature for different water models, where \( r_{\text{max}} \) is the most-probable near neighbour distance, taken to be equal to the position of the first peak in the water oxygen-oxygen radial distribution function \( g_{OO}(r) \) (\( \sim 2.8 \, \text{Å} \)). Other simulation studies obtained similar values for \( G^{r_{\text{max}}}_{K,h}(T) \) at \( \sim 2.8 \, \text{Å} \) and 300 K [110, 111].

A systematic decrease in \( G^{r_{\text{max}}}_{K,h}(T) \) is evident above 210 K for all water models, associated with the increased diffusion (see Fig. 3 B). TIP4P exhibits lower angular correlation in the hydration layer, indicating that the TIP4P dipoles are less strongly aligned. The results reported in Figs. 3 A/B and Fig. 8 are consistent with dynamical coupling between the protein and the solvent since the transition consistently occurs at \( \sim 220 \, \text{K} \), captured by all TIP water models investigated here.
4. Conclusions

The effect of the variation of the water model on the temperature dependence of protein and hydration water dynamics have been investigated here by performing molecular dynamics simulations of hydrated myoglobin. Both protein and water properties were analysed, including the time-averaged structures of the protein, the average mean-square displacements of the protein and water atoms, the solvent reorientational relaxation times, and pair angular correlations between the water molecules.

Variation of the water model between TIP3P, 4P, and 5P leads to the same time-averaged structures of the protein to within statistical error. Also, for all three water models, the temperature-dependent mean-square displacement exhibits the well-known dynamical transition at \( T \approx 220 \, K \).

Mid-infrared pump-probe spectroscopy on the dynamics of HDO in H\(_2\)O yielded an orientational lifetime of \( \tau_{PP} = 2.5 \, \text{ps} \) for bulk water and \( \tau_{PP} > 10 \, \text{ps} \) for “immobilized” water in the solvation shell of tetramethylurea [34, 112] which corresponds to a retardation factor of 4. The bulk water relaxation time \( \tau_D \) was derived to be 8.8 ps from molecular dynamics studies and 7 ps from experiments on dielectric relaxation, both at 303 K [113].

\(^1\)Pump-probe experiments measure different lifetimes than dielectric relaxation experiments or the present simulation data: whereas the lifetime \( \tau_{PP} \) from pump-probe experiments is related to the second-order correlation function \( \langle P_2(\cos \theta(t)) \rangle \) (where \( P_2 \) is the second-order Legendre polynomial), the lifetimes \( \tau_D \) from dielectric relaxation and the present work are determined by the first-order correlation function \( \langle P_1(\cos \theta(t)) \rangle \). The ratio between \( \tau_{PP} \) and \( \tau_D \) depends on the details of molecular diffusion: \( \tau_{PP} = 3\tau_D \) for simple rotational diffusion, but is different for other types of dynamics [114].
Experiments and simulations have shown that the reorientational dynamics of protein hydration water is slowed down relative to the bulk \( e.g., [6, 27, 28, 35, 37, 41] \), with the slowing down being influenced by heterogeneous surface roughness and charge distribution \([22, 25, 40, 41, 104]\).

A recent measurement of cell water dynamics estimates the rotational correlation time for water directly interacting with biomolecular surfaces to be 27 ps \([15]\). The present protein hydration water relaxation times at 300 K are \(31 \pm 3, 35 \pm 3, \) and \(32 \pm 6 \) ps for TIP3P, TIP4P, and TIP5P, respectively. These times are very similar to each other and to the values in Ref. \([15]\), and are approximately a factor of 4 slower then the bulk relaxation times measured in Ref. \([113]\).

The pair angular correlations, calculated using the Kirkwood \(G\)-factor, indicates that TIP4P exhibits somewhat lower angular correlation, although the difference is not large \((\sim 10 \%)\). A decrease in the \(G\)-factor accompanies the dynamical transition.

Taken together, the present results suggest that the global dynamical properties of the protein and hydration water are not significantly affected by variation of the water models among TIP3P, TIP4P, and TIP5P. Although they have documented different bulk phase properties, the water models behave similarly on the protein surface for the quantities investigated.

Furthermore, it has been previously shown that the SPC/E water model \([115]\)
with the GROMOS protein force field [116] also reproduces the protein dynamical transition at the same temperature [96].

5. References


[40] D Argyris, NR Tummala, A Striolo, and DR Cole. Molecular structure and


[50] DR Nutt and JC Smith. Molecular dynamics simulations of proteins: Can the explicit water model be varied? *Journal of Chemical Theory and Computa-


[57] BF Rasmussen, AM Stock, D Ringe, and GA Petsko. Crystalline ribonu-


[64] AL Lee and AJ Wand. Microscopic origins of entropy, heat capacity
101038/35078119.


[97] C Arcangeli, AR Bizzarri, and S Cannistraro. Role of interfacial water in the


[104] RK Murarka and T Head-Gordon. Single particle and collective hydration


[111] P Hoechtl, S Boresch, W Bitomsky, and O Steinhauser. Rationalization of the dielectric properties of common three-site water models in terms of their


Table 1: TIP3P, TIP4P, TIP5P, and TIP3Pm potential energy parameters

<table>
<thead>
<tr>
<th>T / K</th>
<th>TIP3P</th>
<th>TIP3Pm</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>0.0018 ± 0.0000 0.0018 ± 0.0002 0.0007 ± 0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>0.0043 ± 0.0002 0.0016 ± 0.0002 0.0007 ± 0.0002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>170</td>
<td>0.0171 ± 0.0007 0.0025 ± 0.0002 0.0009 ± 0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>0.0444 ± 0.0033 0.0036 ± 0.0001 0.0009 ± 0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>190</td>
<td>0.116 ± 0.0045 0.0086 ± 0.0011 0.0009 ± 0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.3104 ± 0.0075 0.0193 ± 0.0004 0.0014 ± 0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>0.5058 ± 0.0146 0.04 ± 0.0007 0.0028 ± 0.0004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>0.8463 ± 0.0032 0.1306 ± 0.0026 0.003 ± 0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>230</td>
<td>1.243 ± 0.0177 0.2699 ± 0.0010 0.0073 ± 0.0016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>1.7191 ± 0.0464 0.4948 ± 0.0044 0.0202 ± 0.0002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>2.3816 ± 0.0269 0.8021 ± 0.0017 0.0459 ± 0.0072</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>2.8328 ± 0.0261 1.1895 ± 0.0040 0.2601 ± 0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>280</td>
<td>3.9614 ± 0.1640 2.1882 ± 0.0674 1.1840 ± 0.0100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>5.4596 ± 0.0158 3.595 ± 0.0581 2.8162 ± 0.0190</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Diffusion constants in units of $10^{-5} \text{cm}^2/\text{s}$ calculated using the Einstein relation for TIP3P, TIP4P, and TIP5P from a 30 Å cubic water box with the same electrostatic treatment as in the protein simulation.
Table 3: Fitted parameter values for Eq. 8

<table>
<thead>
<tr>
<th>$T_c$/ps</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>21</td>
<td>0.5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$A$/K</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>3810</td>
<td>403</td>
<td>1904</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T_0$/K</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>-145</td>
<td>79</td>
<td>-55</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Fitted parameter values for Eq. 6. $\tau_1$, $\tau_2$, and $\tau_3$ in ps.

<table>
<thead>
<tr>
<th>$T$ / K</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>160</td>
<td>0.08</td>
<td>0.10</td>
<td>0.09</td>
</tr>
<tr>
<td>170</td>
<td>0.08</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td>180</td>
<td>0.09</td>
<td>0.09</td>
<td>0.08</td>
</tr>
<tr>
<td>190</td>
<td>0.11</td>
<td>0.10</td>
<td>0.08</td>
</tr>
<tr>
<td>200</td>
<td>0.14</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td>210</td>
<td>0.16</td>
<td>0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>220</td>
<td>0.21</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>230</td>
<td>0.23</td>
<td>0.19</td>
<td>0.18</td>
</tr>
<tr>
<td>240</td>
<td>0.29</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>260</td>
<td>0.40</td>
<td>0.33</td>
<td>0.25</td>
</tr>
<tr>
<td>280</td>
<td>0.42</td>
<td>0.38</td>
<td>0.28</td>
</tr>
<tr>
<td>300</td>
<td>0.30</td>
<td>0.32</td>
<td>0.29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\tau_0$</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>160</td>
<td>0.00</td>
<td>1.88</td>
<td>5.11</td>
</tr>
<tr>
<td>170</td>
<td>0.10</td>
<td>0.11</td>
<td>7.15</td>
</tr>
<tr>
<td>180</td>
<td>0.15</td>
<td>0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>190</td>
<td>0.41</td>
<td>0.17</td>
<td>0.31</td>
</tr>
<tr>
<td>200</td>
<td>0.83</td>
<td>0.41</td>
<td>0.29</td>
</tr>
<tr>
<td>210</td>
<td>1.06</td>
<td>0.91</td>
<td>0.78</td>
</tr>
<tr>
<td>220</td>
<td>1.74</td>
<td>1.26</td>
<td>0.74</td>
</tr>
<tr>
<td>230</td>
<td>1.66</td>
<td>1.01</td>
<td>1.38</td>
</tr>
<tr>
<td>240</td>
<td>2.18</td>
<td>1.66</td>
<td>2.06</td>
</tr>
<tr>
<td>260</td>
<td>2.60</td>
<td>1.85</td>
<td>1.11</td>
</tr>
<tr>
<td>280</td>
<td>1.63</td>
<td>1.39</td>
<td>0.69</td>
</tr>
<tr>
<td>300</td>
<td>0.44</td>
<td>0.55</td>
<td>0.43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\tau_1$</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>160</td>
<td>78.22</td>
<td>1049.77</td>
<td>2710.67</td>
</tr>
<tr>
<td>170</td>
<td>25.05</td>
<td>27.77</td>
<td>1984.47</td>
</tr>
<tr>
<td>180</td>
<td>53.70</td>
<td>41.52</td>
<td>42.55</td>
</tr>
<tr>
<td>190</td>
<td>59.20</td>
<td>61.70</td>
<td>37.20</td>
</tr>
<tr>
<td>200</td>
<td>56.04</td>
<td>79.61</td>
<td>63.44</td>
</tr>
<tr>
<td>210</td>
<td>50.97</td>
<td>64.38</td>
<td>51.01</td>
</tr>
<tr>
<td>220</td>
<td>45.34</td>
<td>44.00</td>
<td>49.49</td>
</tr>
<tr>
<td>230</td>
<td>42.57</td>
<td>44.90</td>
<td>49.49</td>
</tr>
<tr>
<td>240</td>
<td>38.17</td>
<td>29.87</td>
<td>24.68</td>
</tr>
<tr>
<td>260</td>
<td>22.27</td>
<td>19.51</td>
<td>12.89</td>
</tr>
<tr>
<td>280</td>
<td>7.32</td>
<td>10.09</td>
<td>7.46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\tau_2$</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>160</td>
<td>2819.49</td>
<td>3498.03</td>
<td>3679.43</td>
</tr>
<tr>
<td>170</td>
<td>1714.96</td>
<td>1939.45</td>
<td>3007.36</td>
</tr>
<tr>
<td>180</td>
<td>1535.37</td>
<td>1245.80</td>
<td>1650.42</td>
</tr>
<tr>
<td>190</td>
<td>963.95</td>
<td>1182.51</td>
<td>1538.97</td>
</tr>
<tr>
<td>200</td>
<td>890.55</td>
<td>1020.11</td>
<td>1156.62</td>
</tr>
<tr>
<td>210</td>
<td>610.75</td>
<td>968.10</td>
<td>1007.51</td>
</tr>
<tr>
<td>220</td>
<td>580.03</td>
<td>736.41</td>
<td>838.47</td>
</tr>
<tr>
<td>230</td>
<td>594.33</td>
<td>690.84</td>
<td>741.91</td>
</tr>
<tr>
<td>240</td>
<td>515.06</td>
<td>443.10</td>
<td>431.47</td>
</tr>
<tr>
<td>260</td>
<td>398.39</td>
<td>349.69</td>
<td>280.67</td>
</tr>
<tr>
<td>280</td>
<td>269.09</td>
<td>235.99</td>
<td>123.27</td>
</tr>
<tr>
<td>300</td>
<td>86.83</td>
<td>119.99</td>
<td>86.76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$1 - a_0 - a_1$</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>TIP5P</th>
</tr>
</thead>
<tbody>
<tr>
<td>160</td>
<td>0.86</td>
<td>0.90</td>
<td>0.91</td>
</tr>
<tr>
<td>170</td>
<td>0.87</td>
<td>0.87</td>
<td>0.89</td>
</tr>
<tr>
<td>180</td>
<td>0.84</td>
<td>0.83</td>
<td>0.84</td>
</tr>
<tr>
<td>190</td>
<td>0.75</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>200</td>
<td>0.66</td>
<td>0.73</td>
<td>0.78</td>
</tr>
<tr>
<td>210</td>
<td>0.60</td>
<td>0.58</td>
<td>0.69</td>
</tr>
<tr>
<td>220</td>
<td>0.48</td>
<td>0.52</td>
<td>0.59</td>
</tr>
<tr>
<td>230</td>
<td>0.40</td>
<td>0.46</td>
<td>0.49</td>
</tr>
<tr>
<td>240</td>
<td>0.30</td>
<td>0.35</td>
<td>0.38</td>
</tr>
<tr>
<td>260</td>
<td>0.20</td>
<td>0.26</td>
<td>0.31</td>
</tr>
<tr>
<td>280</td>
<td>0.19</td>
<td>0.18</td>
<td>0.29</td>
</tr>
<tr>
<td>300</td>
<td>0.27</td>
<td>0.23</td>
<td>0.28</td>
</tr>
</tbody>
</table>
Figure 1: TIPnP geometries. Dashed lines represent $1/2$ of the van der Waals radius $\sigma_0$. 
Figure 2: Backbone heavy-atom RMSD per residue at 300 K. **Average over pairs of structures solvated in the same or in different water models. “SELF” refers to RMSD values between protein structures solvated by the same water model whereas “CROSS” refers to deviations in configurations between the models.**
Figure 3: (A) Mean-square displacement of myoglobin solvated in TIP3P, TIP4P, or TIP5P as a function of temperature. The inset shows an expanded view of the low temperature region up to $T\approx 210 \text{K}$. (B) Mean-square displacement of the water as a function of temperature. The insets (B.1) and (B.2) give expanded views since the data cover multiple orders of magnitude. (C) Mean-square displacement of myoglobin in TIP3P as a function of time at different temperatures.
Figure 4: (A) Mean-square displacement of TIP3P, TIP4P, or TIP5P water molecules as a function of temperature from bulk water simulations. The inset shows an expanded view of the low temperature region. The solid lines corresponds to linear fits for $T = 20 - 100$ K. (B) Mean-square displacement as a function of time for TIP3P and various temperatures. The insets shows data belonging to further temperatures in an appropriate magnification of the abscissa.
Figure 5: Fraction of residues exhibiting anharmonic dynamics.
Figure 6: (A) Protein hydration water dipole orientational autocorrelation functions (Eq. 5) for TIP3P with Eq. 6 fitted. TIP4P and TIP5P data are similar. (B) Reorientational relaxation lifetimes $\tau$ fitted with the Vogel-Fulcher-Tammann equation fitted (Eq. 8) according to Eq. 7.
Figure 7: Fitted parameter values $a_0$, $a_1$, $1 - a_0 - a_1$, $\tau_0$, $\tau_1$, and $\tau_2$ from Eq. 6. Lines connect points as guide to the eye. Numerical values are given in Tbl. 4.
Figure 8: Local orientational ordering of water dipoles measured with the distance-dependant Kirkwood $G$-factor $G_{K,h}^{max}(T)$ at the maximum of the oxygen-oxygen radial distribution function as a function of temperature for different water models. Lines connect points as guide to the eye.