

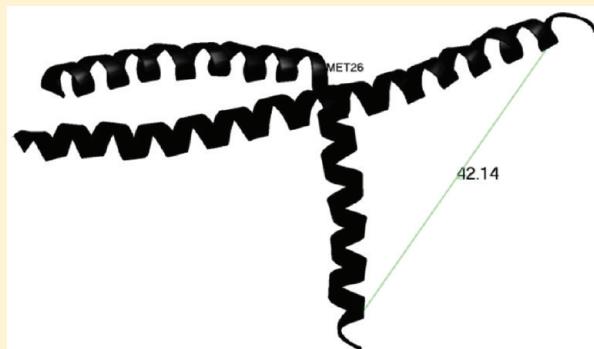
# Simulations of Temperature and Salt Concentration Effects on bZIP, a Basic Region Leucine Zipper

Robert I. Cukier\*

Department of Chemistry Michigan State University, East Lansing, Michigan 48824-1322, United States

 Supporting Information

**ABSTRACT:** Basic region leucine zipper (bZIP) transcription factors are dimeric proteins that recognize DNA. The monomers consist of a leucine zipper subdomain responsible for dimerization and a highly basic DNA recognition subdomain. Twelve explicit solvent molecular dynamics (MD) trajectories were run on the GCN4 bZIP transcriptional factor in the absence of DNA at three temperatures and two ion concentrations (0 mM with  $\text{Cl}^-$  ions to neutralize the bZIP and 200 mM with additional  $\text{Na}^+$  and  $\text{Cl}^-$  ions) to probe the conformational ensemble that the basic region samples. In most trajectories, the basic region exhibits an alligator-jaw-like opening and closing (only one monomer moves), versus scissor-like motion, by a mainly rigid body, hinge motion centered on three “fork” residues that span the basic region to the coiled coil. In this motion, the  $\alpha$ -helical character of the basic region monomers is mostly maintained. A broad range of distances is accessed, consistent with the absence of particular interactions for the basic region monomers. In two of the trajectories, the basic region monomers “collapse” to form a stable state. The coiled coil, leucine zipper subdomain is very stable for all of the trajectories. Ion solvation of the charged residue side chains is transient, on the scale of a few picoseconds. There is no evidence for persistent specific ion salt bridges to charged residues. For 0 mM, only certain basic region positively charged residues are substantially  $\text{Cl}^-$  ion salt bridged. For 200 mM, in addition, some basic region negatively (positively) charged residues are salt bridged to  $\text{Na}^+$  ( $\text{Cl}^-$ ) ions. The different ion solvation patterns at the two ion concentrations are not greatly temperature sensitive, and the conformational sampling found in the MD is remarkably unperturbed by ion concentration and/or temperature.



## 1. INTRODUCTION

Transcription factors are monomeric proteins that homo- or heterodimerize to bind DNA.<sup>1</sup> The basic region leucine zipper (bZIP) transcriptional regulator superfamily, the largest and most highly conserved one in eukaryotes, is involved in the regulation of development, in metabolism, and other cell functions.<sup>2,3</sup> Different subdomains of the monomers are responsible for dimerization and DNA recognition.<sup>4</sup> The dimerization unit is a leucine zipper sequence that produces a stable  $\alpha$ -helical coiled coil upon dimerization. The DNA recognition subdomain (basic region) has a great excess of basic residues. Before DNA binding, the basic region is thought to be partially structured. Once DNA is bound, the basic region also becomes very structured.

This work presents the results of molecular dynamics (MD) simulations on a bZIP transcriptional factor in the absence of DNA. Of interest is the nature of the fluctuations of the monomer-to-monomer separation of the basic region and the stability of the basic region monomers that are important for DNA binding. In particular, do these properties have a significant dependence on temperature and ionic strength, or is bZIP robust and not much influenced by temperature and ion concentration? If the basic region monomers were quite labile, higher temperatures might promote less structure. Some

simulations have shown that protein stability can be influenced by ionic strength.<sup>5–7</sup> Because the basic region is so highly charged, the role of ion concentration could be decisive in the conformational sampling, and in principle, there could be combined effects of temperature and ion concentration.<sup>5</sup> While the basic region has a large net positive charge and the coiled coil zero net charge, the bZIP coiled coil also has numerous charged residues. Thus, one might anticipate that there could be specific ion to charged residue salt-bridge-like interactions throughout bZIP that could have a significant dependence on temperature and ion concentration. We investigate both the fraction of such salt bridges for all of the charged residues as well as their lifetimes.

The MD starting structure is based on the crystal structure of GCN4, the bZIP of the yeast transcriptional factor, which was determined in the presence of DNA.<sup>4</sup> In this structure, the two monomers are bound as a supercoil in the coiled coil, and the ( $\alpha$ -helical) monomers diverge in the basic region and grip DNA, with the DNA perpendicular to the dimer long axis. In the absence of DNA, solution NMR and CD studies<sup>8–12</sup> find

**Received:** January 25, 2012

**Revised:** May 2, 2012

**Published:** May 5, 2012

that the coiled coil subdomain is stable and the basic region samples a broad range of conformations with some preference for helical conformations.

In general,<sup>13–16</sup> and in transcription, in particular,<sup>2</sup> an important role for intrinsically disordered proteins in function has become increasingly appreciated. In the context of transcription factors, the presence of a bias toward secondary structures such as  $\alpha$ -helices in the basic region could provide a reduction of an entropic penalty and present energetically favorable sites for DNA binding as a way of reducing the free-energy cost of having to fold unstructured monomers concomitantly with DNA binding. Thus, we investigate whether binding might occur by conformational selection,<sup>17–19</sup> whereby the basic region extensively fluctuates and, for appropriate conformations, binds DNA, or by an induced fit mechanism.

To study the influence of temperature and ion concentration on these GCN4 bZIP properties, we carried out explicit solvent simulations at three temperatures,  $T = 303, 333$ , and  $363\text{ K}$ , and two ion concentrations. The ion concentrations are “0 mM”, corresponding to 15 added  $\text{Cl}^-$  ions used to neutralize the net GCN4 charge, and  $\sim 200\text{ mM}$ , where an additional 84  $\text{Cl}^-$  and 84  $\text{Na}^+$  ions were added. For each state point, two trajectories were run, each trajectory of at least 30 ns, in order to provide enough time for extensive sampling of the conformational space. The basic region is found, in most trajectories, to undergo an alligator-jaw-like opening and closing (only one monomer moves), versus scissor-like motion, of the basic region that occurs by a dominantly rigid body, hinge motion centered on “fork” residues<sup>20,21</sup> that span the transition from the basic region to the coiled coil. The coiled coil, leucine zipper part is very stable for all of the state points considered.

The ion solvation of the charged residue side chains is found to be very transient, with lifetimes on the order of picoseconds and no evidence for specific ion solvation. While there are differences in the ion solvation found for the different temperatures, they are quite modest. For 0 mM, not surprisingly, only the basic region is found to be solvated by the  $\text{Cl}^-$  ions. At 200 mM, the presence of additional  $\text{Cl}^-$  ions and the  $\text{Na}^+$  ions leads to the coiled coil going from no  $\text{Cl}^-$  ions salt bridged at 0 mM to a finite  $\text{Cl}^-$  ion salt bridge fraction. The presence of the  $\text{Na}^+$  ions not only, of course, populates the coiled coil acid residues but also permits  $\text{Cl}^-$  ions to populate the basic residues of the coiled coil. Despite the different ion solvation patterns at the two ion concentrations, the conformational sampling found in the MD is remarkably unperturbed by ion concentration and/or temperature.

## 2. METHODS

**2.1. System Preparation.** The crystal structure of GCN4<sup>4</sup> in the presence of DNA (PDB accession code 1YSA) was used to initiate the simulation. The DNA was removed from the structure. There were 114 resolved residues (the beginning MET in the first monomer and ending ARG in second monomer were not resolved). In the dimer, the linear long axis dimension was about 75 Å, the maximum separation of the basic ends was about 45 Å, and the width was about 25 Å. The protein was placed in a cubic box of linear dimension 88.78 Å. For the standard density, 23328 SPC waters were added and 1077 removed to accommodate the protein. With standard protonation states for all of the residues, 15  $\text{Cl}^-$  ions were added for neutralization by replacing waters at various locations to give the different initial conditions for what will be referred to as the 0 mM ion concentration. For 200 mM, in addition to

the 15  $\text{Cl}^-$  ions, 108  $\text{Na}^+$  and 108  $\text{Cl}^-$  ions were first added on a lattice and 48 removed (in pairs) to leave 84  $\text{Na}^+$  and 84  $\text{Cl}^-$  at various locations to produce different initial conditions.

**2.2. Molecular Dynamics.** The CUKMODY protein molecular dynamics code<sup>22</sup> with the GROMOS96<sup>23</sup> force field was used to generate the trajectories. SHAKE<sup>24</sup> was used to constrain bond distances, enabling a 2 fs time step, and the temperature was controlled with a Berendsen thermostat<sup>25</sup> with a relaxation time of 0.2 ps. The PME method<sup>26</sup> was applied with a direct-space cutoff of 9.86 Å, an Ewald coefficient of 0.45, and a  $90 \times 90 \times 90$  reciprocal space grid for the evaluation of the electrostatic and the attractive parts of the Lennard-Jones energies and forces. Each system simulated was first equilibrated at constant number, temperature, and volume using one-body restraints on the protein that were gradually reduced to zero over 100 ps. The first 1 ns of simulation data were discarded in the production runs, summarized in Table 1.

Table 1. MD Simulations Carried out on bZIP

state <sup>a</sup>	length (ns) <sup>b</sup>	temperature (K)	ion concentration (mM) <sup>c</sup>
T303_0 mM_1, _2	50	303	0
T303_200 mM_1, _2	58	303	200
T333_0 mM_1, _2	28	333	0
T333_200 mM_1, _2	36	333	200
T363_0 mM_1, _2	36	363	0
T363_200 mM_1, _2	32	363	200

<sup>a</sup>\_1 and \_2 denote the two independent trajectories. <sup>b</sup>Length of each trajectory. <sup>c</sup>0 mM corresponds to 15  $\text{Cl}^-$  ions, and 200 mM corresponds to 15 + 84  $\text{Cl}^-$  and 84  $\text{Na}^+$  ions.

**2.3. Data Analysis.** The atom trajectories that are written out every 1 ps were first processed with ANALYZER<sup>27</sup> to provide Cartesian coordinates after best fitting the trajectory on all of the CA atoms. ANALYZER was then used to obtain the following information:

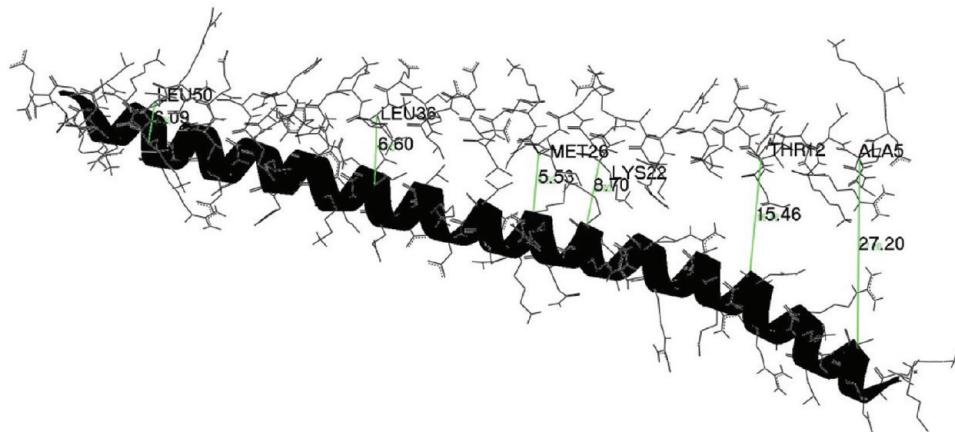
rmst: The root-mean-square fluctuation as a function of time for all atoms and for all backbone atoms.

rmsd (rmsf): the root-mean-squared displacement from the starting (average) structure by a residue of all atoms and backbone atoms.

Distances: All desired distance coordinates of CA atoms between corresponding monomer residues (independent of the best fitting procedure).

PCA: Principal component analysis<sup>28–30</sup> provides a decomposition of a trajectory into modes (linear combinations of the atom coordinates) that successively capture decreasing amounts of the mean-square fluctuation. The first mode captures the largest fraction of the mean-square fluctuation. PCA modes were generated by first best fitting on the monomers' backbone atoms.

$\alpha$ -Helical Content: The  $\alpha$ -helical content was monitored by evaluating the 19 1–4 hydrogen bonds (residues 30–34 to 48–52) in the coiled coil and 17 1–4 hydrogen bonds (5–9 to 21–25) in the basic region. The hydrogen bond criteria are a 1–4 carbonyl oxygen to backbone nitrogen distance less than 3 Å and the corresponding OHN angle between 0 and 30°. The distance was picked from the crystal structure distances<sup>31</sup> of the GCN4



**Figure 1.** The bZIP dimer with the coiled coil to the left and the basic region to the right. The CA–CA monomer-to-monomer distances between three basic region (Ala5, Thr12, Lys22, monomer 1) and three coiled coil (Met 26, Leu36, Leu50, monomer 1) residues are indicated.<sup>4</sup>

leucine zipper. The “tight” angle requirement was used to not introduce many 1–3 and 1–5 hydrogen bonds.

Fraction SB: The “salt bridges” between ions and the charged residue side chains were monitored over a trajectory by finding the closest distance between a numbered ion and a particular side chain atom: Arg (CZ,NZ,NH1,NH2), Lys (NZ), Glu (OE1,OE2), Asp (OD1,OD2). A 5.5 Å cutoff was used to count ions within 4.5 Å of H atoms because, for example, arginine has NH<sub>3</sub> groups. The fraction of the trajectory time that met this cutoff produces each (charged) residue’s fraction SB.

Dwell Time Histogram: The time record of which ion is salt bridged to a residue side chain is available from the fraction SB trajectory data. Thus, this record can be used to obtain the dwell times, the runs of how long a particular ion is salt bridged to a residue before replacement with another ion. A histogram of these dwell times can be constructed to obtain a mean ion residence time.

Number SB: Another measure of ion side chain salt bridging more similar to a radial distribution function than the fraction SB was also constructed. It counts all ions within a given radius around residue side chain charge centers. For Arg(Glu/Asp)[Lys], there are 4(2) [1] specified atoms NZ, CZ, NH1, NH2(OE1,OE2/OD1,OD2)[NZ] that are the main charge centers for ion salt bridging. The trajectories are used to count all such interactions over the trajectory for the indicated charged side chains and then time averaged.

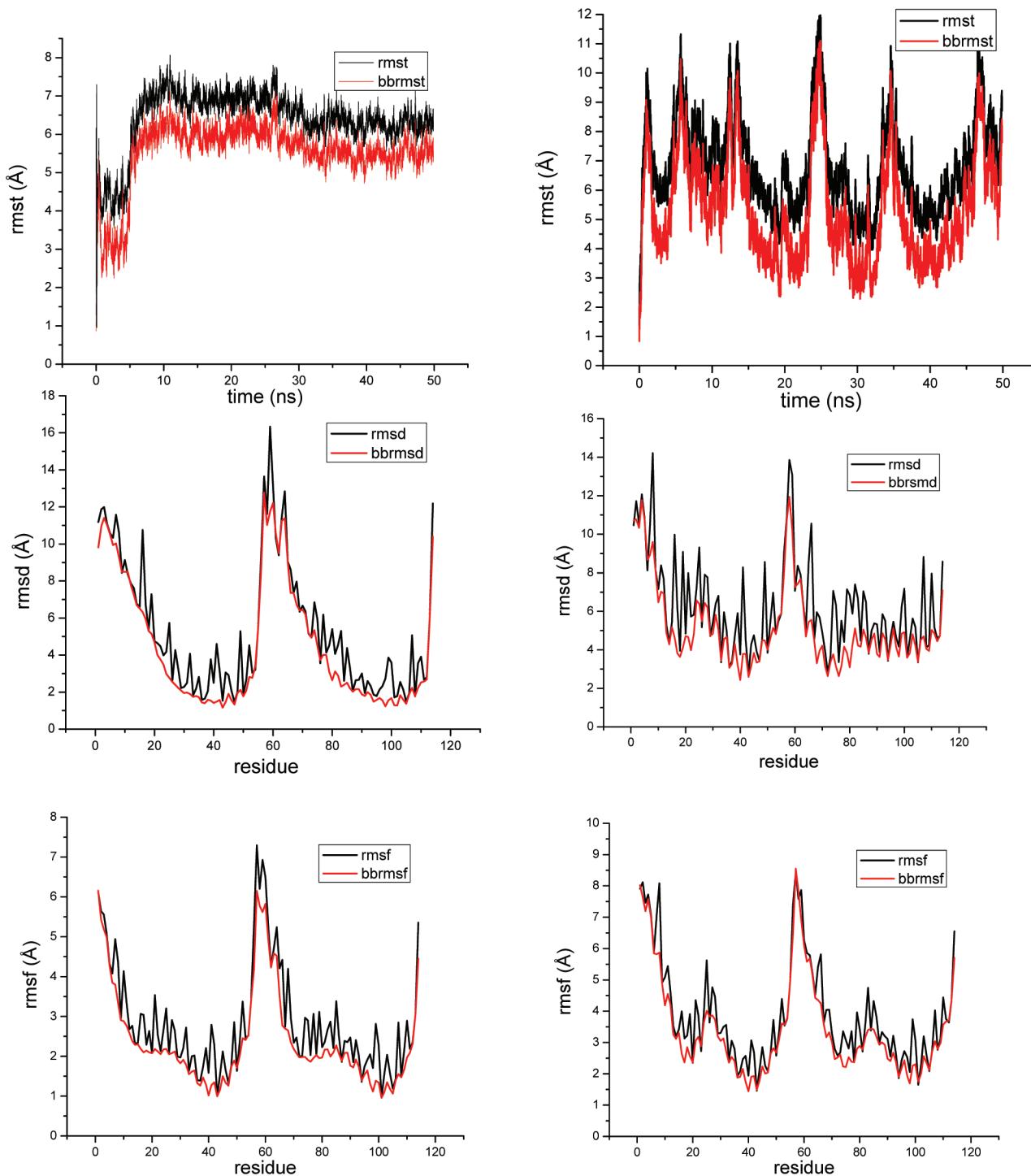
### 3. RESULTS

The simulated dimer from the crystal structure configuration<sup>4</sup> is depicted in Figure 1, with some representative monomer–monomer CA distances noted. The numbering scheme starts from the basic region of monomer 1 (basic region 1–25, coiled coil 26–57) and ends with the coiled coil of monomer 2 (basic region 58–83, coiled coil 84–114). The six state points used and trajectory lengths are given in Table 1, using the notation TX\_YmM\_Z to denote the temperature X, ion concentration Y, and trajectory Z. Each state point was simulated twice starting from different configurations (see section 2.1). The designation 0 mM corresponds to the addition of 15 Cl<sup>−</sup> ions that were required to neutralize bZIP at the standard

protonation states of the residues, and 200 mM corresponds to the further addition of 84 Cl<sup>−</sup> and 84 Na<sup>+</sup> ions.

**3.1. Structure Effects.** Figures 2 (T303\_0mM\_1 and T303\_0mM\_2), 4 (T363\_0mM\_1), and 6 (T303\_200mM\_1) present the root-mean-square (rms) deviation from the crystal structure versus time for all (rmst) and backbone (bbrmst) atoms, along with the rmsd and rmsf measures that are, respectively, the rms deviation from the crystal structure and that from the average structure. These are, for each residue, for all residue atoms (rmsd), and for all backbone (bbRMSD) atoms. From the numbering scheme used, if monomers 1 and 2 were statistically equivalent, the rmsd and rmsf values of residues 1–57 and 58–114 would be identical. Figures 3 (T303\_0mM\_1 and T303\_0mM\_2), 5 (T363\_0mM\_1), and 7 (T303\_200mM\_1) are the corresponding plots of the monomer-to-monomer CA distances between three residues in the basic region and three in the coiled coil, as indexed in Figure 1.

The T303\_0mM\_1,2 trajectories in Figures 2 and 3 show that two very different paths are taken, as is most evident from the rmst trace in Figure 2 and the monomer–monomer distances in Figure 3. In trajectory 1, there is a “collapse” of the basic region, while in trajectory 2, the basic region monomer–monomer distances sample many distances. Note that the only difference in these trajectories was the initial placement of the Cl<sup>−</sup> ions, and the collapse occurs after 10 ns. Thus, it is safe to conclude that the thermal fluctuations can support these two very different conformations. Figures 4 and 5 for T363\_0mM\_1 show behavior similar to the T303\_0mM\_2 trajectory. There is no evidence for a collapse of the basic region. All of the rmstd rmsd and rmsf data are similar to those shown in Figure 1 for trajectory 2. We do not display the T363\_0mM\_2 data (Figures S1 and S2, Supporting Information) since it is similar to the T363\_0mM\_1 data. The data for the T333\_0mM\_1 and T333\_0mM\_2 trajectories (Figures S3 and S4 in the Supporting Information) are also quite similar to those of the T303\_0mM\_2 trajectory; again, there is no evidence for a collapse of the basic region. All of these rmst and rmsd and rmsf data are similar to those shown in Figure 1 for trajectory 2. Thus, with the exception of the basic region in the T303\_0mM\_1 trajectory, all of the 0 mM trajectories show that the coiled coil is stable and the basic region is fluctuating over a large range of monomer-to-monomer distances.

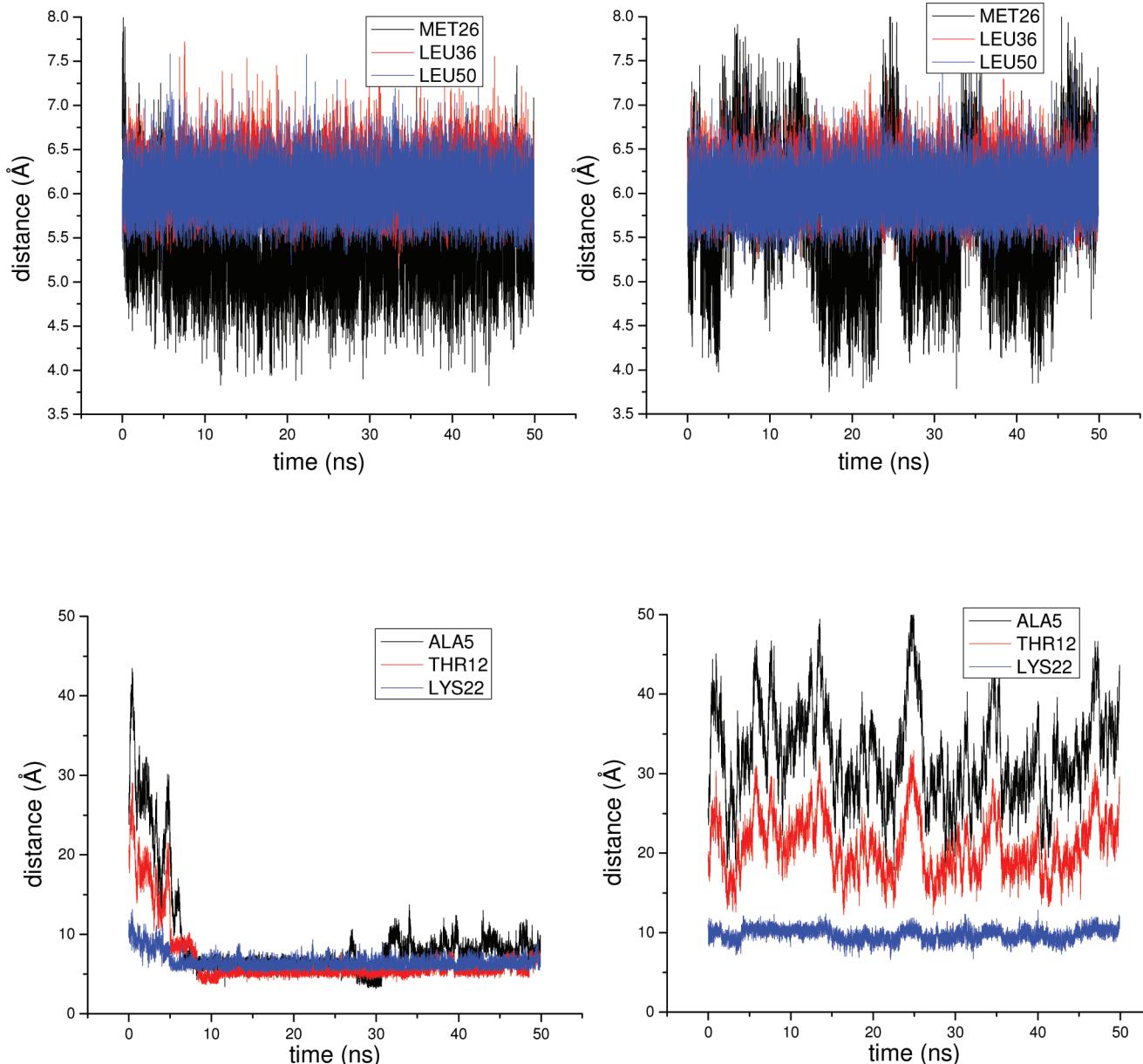


**Figure 2.** (Left) T303\_0mM\_1; (Right) T303\_0mM\_2. rmst: root-mean-square (rms) deviation from the crystal structure versus time for all (rmst) and backbone (bbrmst) atoms; rmsd (rmsf): rms deviation from the crystal structure (average structure) for all and backbone (bb) atoms by the residue. Residue numbering of monomer 1 [2] is (1–25) [58–83] for the basic region and (26–57) [84–114] for the coiled coil.

With one exception discussed below, the 200 mM trajectories are similar in character to the T303\_0mM\_2 trajectory, with its fluctuating basic region behavior. The data for T303\_200mM\_1 (Figures 6 and 7), T303\_200mM\_2 (Figures S5 and S6, Supporting Information), T333\_200mM\_1 and T333\_200mM\_2 (Figures S7 and S8, Supporting Information), and T363\_200mM\_1 (Figure 8) are quite similar to the T303\_0mM\_2 data. The exception is the T363\_200mM\_2 trajectory, where the monomer–monomer

distance plots in Figure 8 show again the basic region collapse behavior. (The rmst, rmsd, and rmsf data are shown in Figure S9, Supporting Information). Thus, at this high temperature and high salt concentration, one of the trajectories, T363\_200mM\_2, has the same collapse behavior as that found at the low-temperature, low-salt-concentration (T303\_0mM\_1) trajectory.

Excluding the two basic region collapse trajectories for the moment, there is no evident dramatic effect from temperature

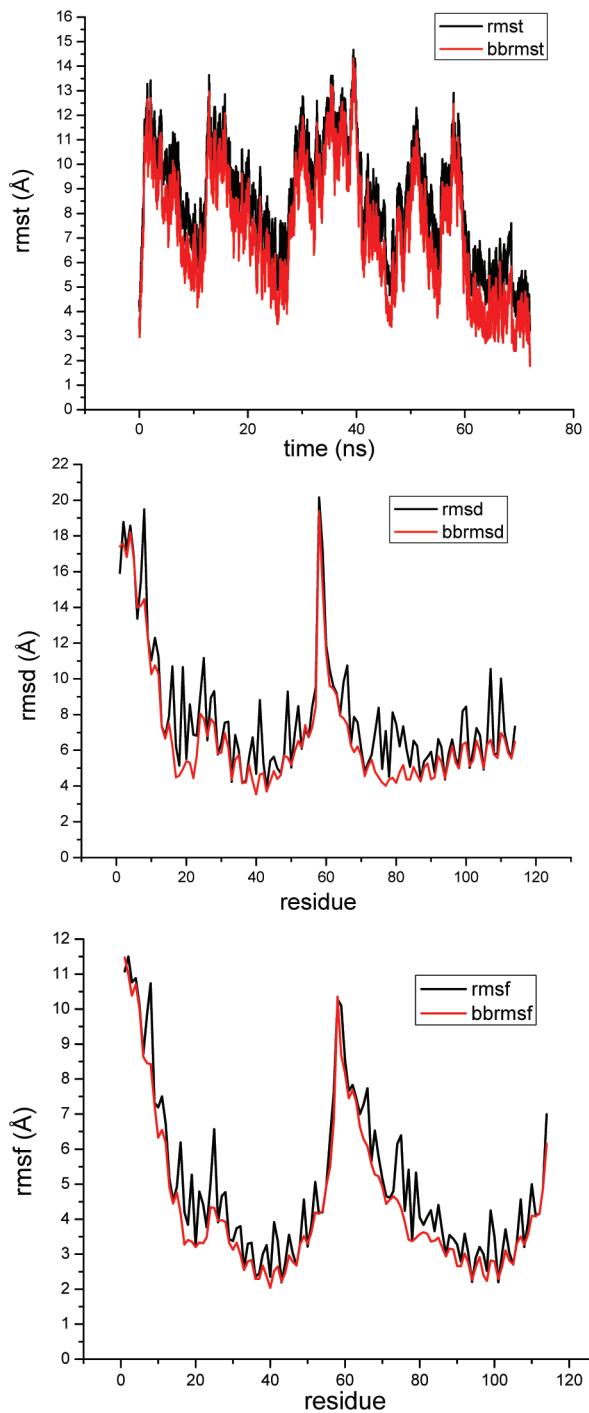


**Figure 3.** (Left) T303\_0mM\_1; (Right) T303\_0mM\_2. Distances between CA atoms of the indicated residues of the monomers (using monomer 1 numbering). The monomer 1 basic region residues span residues 1–25, and the coiled coil residues span residues 26–57. The “collapse” of the basic region is evident in trajectory 1.

or ion concentration. The coiled coil region maintains its “knobs-into-holes” structure,<sup>32</sup> as is characteristic of leucine zippers, while the basic region undergoes opening and closing fluctuations of substantial magnitude. In principle, the dimer could (and typically will in low-protein-concentration experiments) dissociate, but on this MD time scale, it does not. Also, from CD evidence, the coiled coil is  $\alpha$ -helical, as is essential for dimerization, while the basic region is thought to be transiently  $\alpha$ -helical. We will examine this issue more in the Discussion. Here, we present evidence that both regions are mainly  $\alpha$ -helical. Figure 9 displays the fraction (over the trajectory) of  $\alpha$ -helical content for 19 1–4 hydrogen bonds (residues 30–34 to 48–52) in the coiled coil and 17 (5–9 to 21–25) in the basic region for T303\_200mM\_1 and T363\_200mM\_1. (The data for T303\_200mM\_2 are shown in Figure S10, Supporting Information). The coiled coil  $\alpha$ -helical content is very modestly greater than that for the basic region. The trajectory 1–4

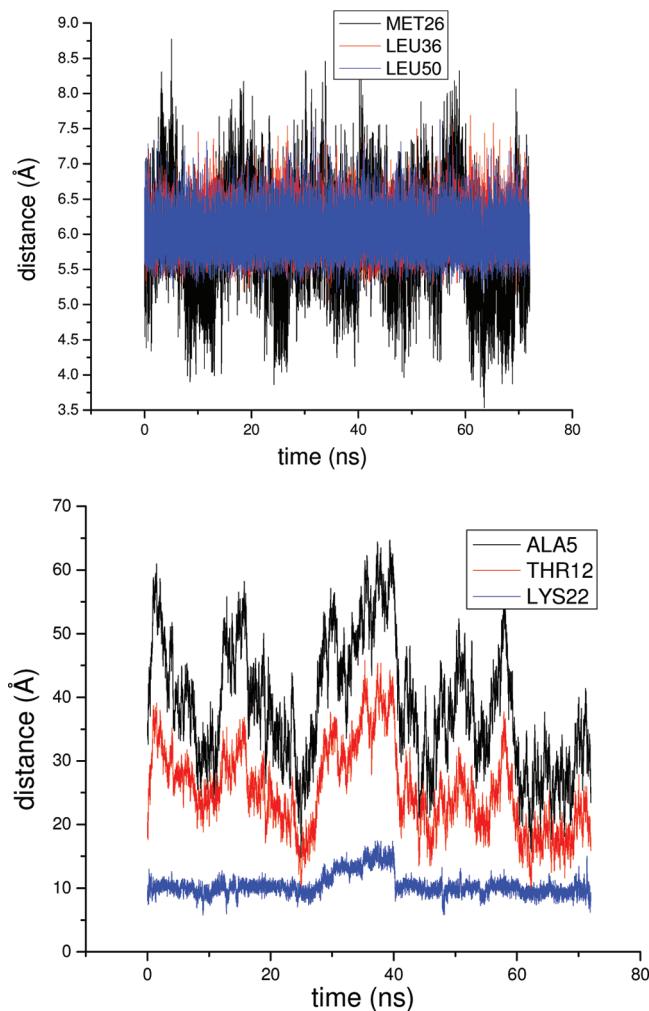
distances have rms  $\approx 0.5$  Å; thus, using this cutoff and a rather stringent angle leads to the conservative definition (see section 2.3) of hydrogen bonding. Changing these parameters mainly scales the fractions; the coiled coil and basic region hydrogen bonds are similar in proportion. Thus, the trajectory data do not lead to an unstructured basic region.

The nature of the basic region fluctuations should be critical to the function of a bZIP. Figure 10 displays a snapshot taken from the T303\_200mM\_1 trajectory. The coiled coil is completely intact, and the basic region has opened to a large degree. From the figure, it appears that the opening is by a rotation about a hinge located close to Met26. Note that the Met26 monomer-to-monomer distance undergoes relatively large fluctuations, as exemplified in Figure 3. To verify that the opening of the “jaw” is around a hinge, we used two available programs that report on hinges, DSSP<sup>33</sup> and DynDom.<sup>34</sup> DSSP lists the helical (H) content of proteins and for the above



**Figure 4.** T363\_0mM\_1. rmst: root-mean-square (rms) deviation from the crystal structure versus time for all (rmst) and backbone (bbrmst) atoms. rmsd (rmsf): rms deviation from the crystal structure (average structure) for all and backbone (bb) atoms by residue. Residue numbering of monomer 1 [2] is (1–25) [58–83] for the basic region and (26–57) [84–114] for the coiled coil.

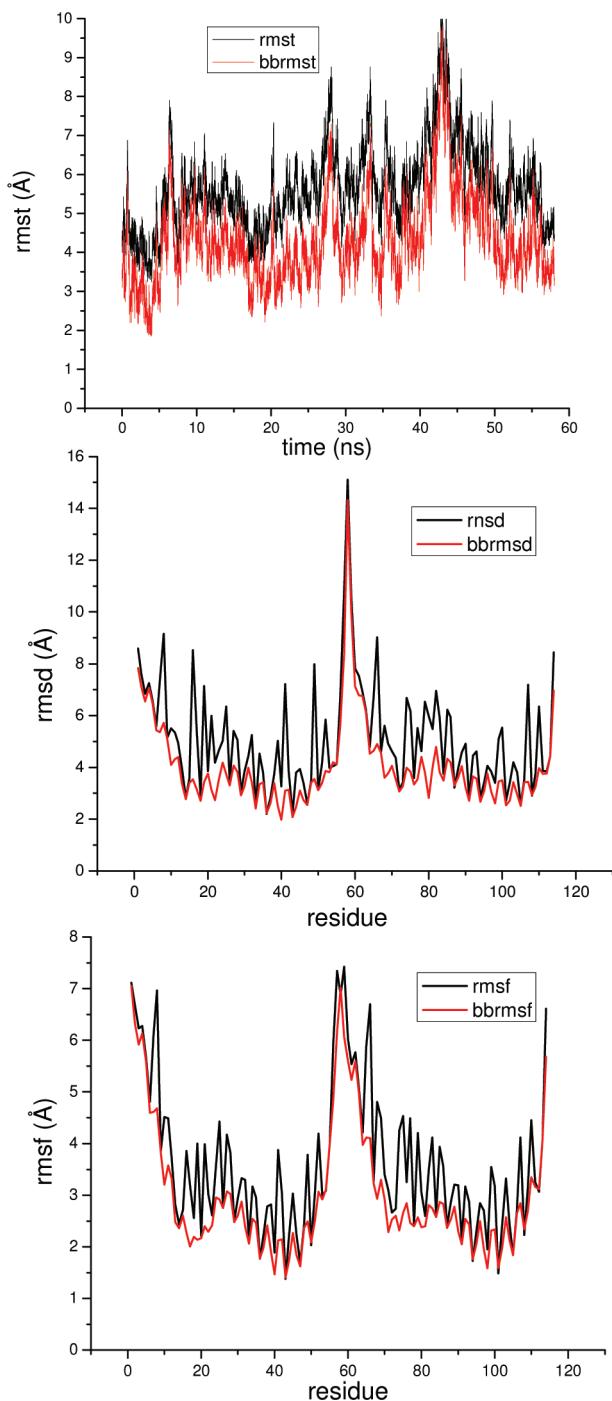
snapshot identifies Gln24 as H, Arg25 as T (turn), and Met26 as H in monomer 1, while the corresponding residues in monomer 2 are all H. (All of the other residues are H.) DynDom relies on two structures to identify hinges in a protein. Using the above snapshot and our initial configuration shows that there is a hinge with the bending region identified as residues 24–26 with large changes in the Arg25 phi and psi angles. The DynDom displaying hinge motion connecting the



**Figure 5.** T363\_0mM\_1. Distances between CA atoms of the indicated residues of the monomers (using monomer 1 numbering). The basic region residues span 1–25, and the coiled coil residues span 26–57.

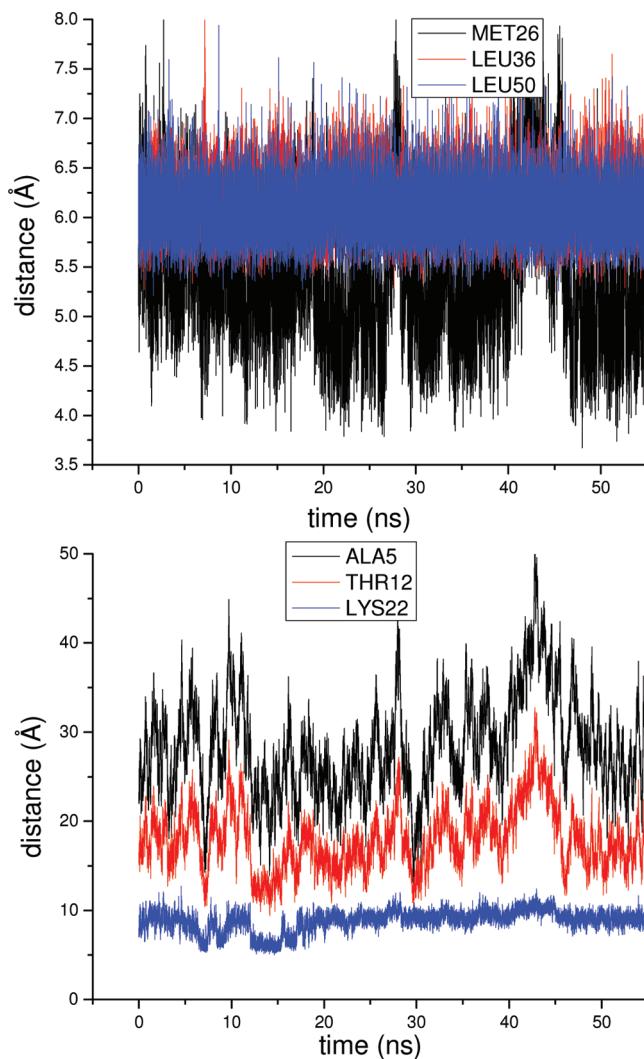
two conformations shows essentially a rigid body motion of the basic region of monomer 1.

The notion of motion around a hinge is supported directly in the trajectory data by computing the correlation coefficient between, for example, the Ala5 and Thr12 monomer–monomer distances. For all of the (noncollapse) data, the correlation coefficient is  $R \approx 0.92$ . This makes it clear that for these trajectories, the basic region is undergoing a correlated, jaw-like open and close motion. The dynamics could occur by some smooth motion or by a two-state, predominantly open or closed, mechanism. By histogramming the Ala5 monomer-to-monomer distance fluctuations, the potential of mean force (PMF) can be obtained for the basic region motion because the distance sampling in, for example, Figure 7 is reasonably good. The T303\_200mM\_1 PMF in Figure 11 (T303\_200mM\_2, Figure S11, Supporting Information) shows that the basic region jaw motion is a fluctuation centered around the crystal structure Ala5 monomer-to-monomer distance of  $\sim 27$  Å. The (harmonic) force constant,  $\sim 0.01$  kcal/mol/Å $^2$ , characterizing the PMF indicates a slow, weak spring motion in what amounts to a collective coordinate in free energy with rms excursions on the 7–8 Å scale.



**Figure 6.** T303\_200mM\_1. rmst: root-mean-square (rms) deviation from the crystal structure versus time for all (rmst) and backbone (bbrmst) atoms. rmsd (rmsf): rms deviation from the crystal structure (average structure) for all and backbone (bb) atoms by residue. Residue numbering of monomer 1 [2] is (1–25) [58–83] for the basic region and (26–57) [84–114] for the coiled coil.

The jaw dynamics could be scissor-like (both basic region monomers move) or alligator-like (one basic region monomer moves with the other fixed). Figure 12 plots parametrically the end-to-end distance of the two monomers for T303\_200mM\_2 and shows that the jaw motion is alligator-like, as illustrated for an extreme configuration in Figure 10. Similar results are obtained for other trajectories (data not shown).

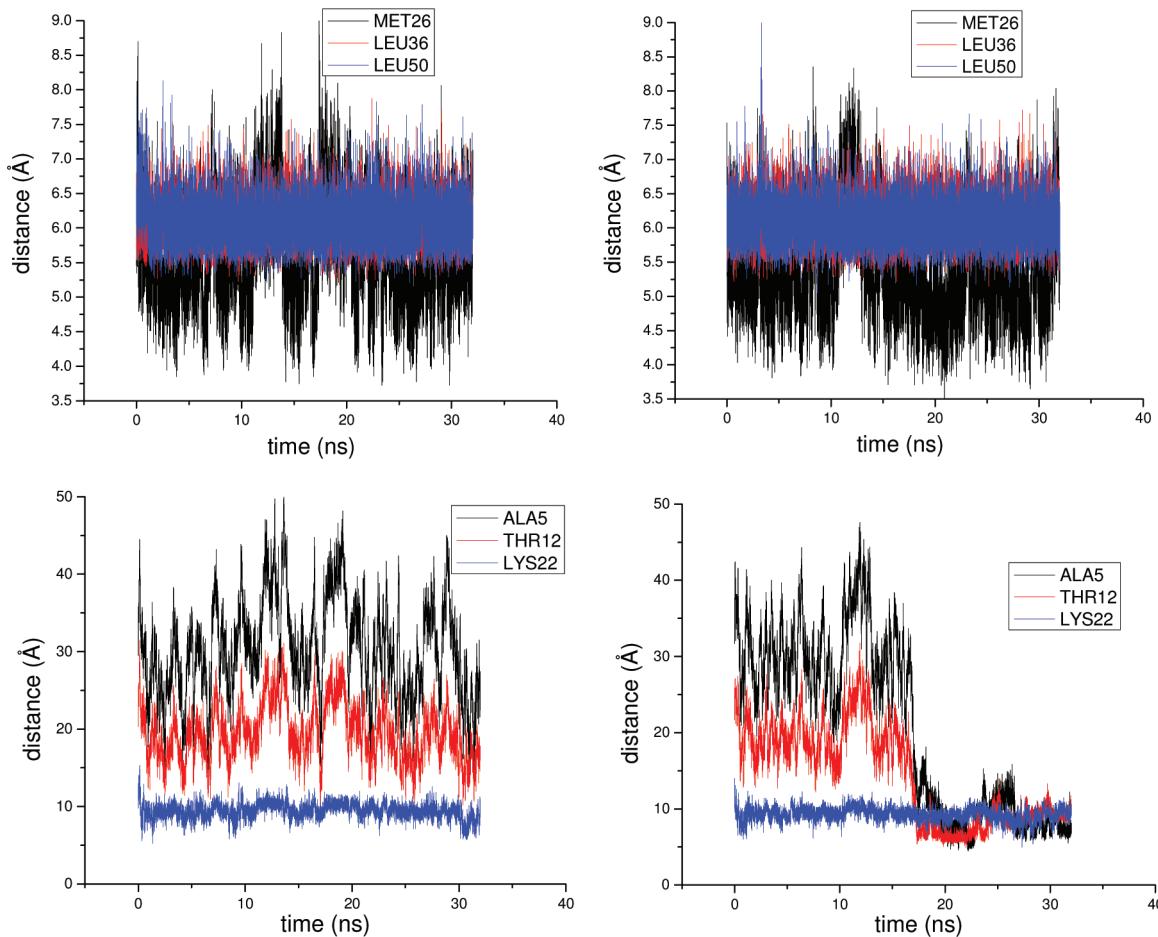


**Figure 7.** T303\_200mM\_1. Distances between CA atoms of the indicated residues of the monomers (using monomer 1 numbering). The basic region residues span 1–25, and the coiled coil residues span 26–57.

Finally, PCA was carried out on each monomer for the T303\_200mM\_1, \_2 and T363\_200mM\_1, \_2 data. The first PCA mode describes about two-thirds of the total fluctuation of each monomer, showing that the PCA picks out a majority of the motion in mode one. (The first 10 modes capture ~95% of the overall trajectory fluctuations.) A parametric plot of mode 1 and the end-to-end distance, for each monomer, show that this mode is strongly correlated with the end-to-end distance. Thus, a great deal of the overall fluctuation of each monomer arises from the conformational sampling of the basic region.

To summarize, the coiled coil part of the dimer stays well dimerized, and the basic region undergoes an alligator-like (one monomer moves) dynamics that is centered on its equilibrium position with large excursions to either side on the scale of 7–8 Å, characterized by a harmonic PMF. All of these features are not very dependent on the temperature and/or ion concentration.

However, a more fine-grained examination of the data does show that there is a modest distinction among the trajectories for the basic region fluctuations. Table 2 compares the average distance between the Ala5 CA atoms of the two monomers. There is a tendency to increasing average distance with



**Figure 8.** (Left) T363\_200mM\_1; (Right) T363\_200mM\_2. Distances between CA atoms of the indicated residues of the monomers (using monomer 1 numbering). The basic region residues span 1–25, and the coiled coil residues span 26–57. The “collapse” of the basic region is evident in trajectory 2.

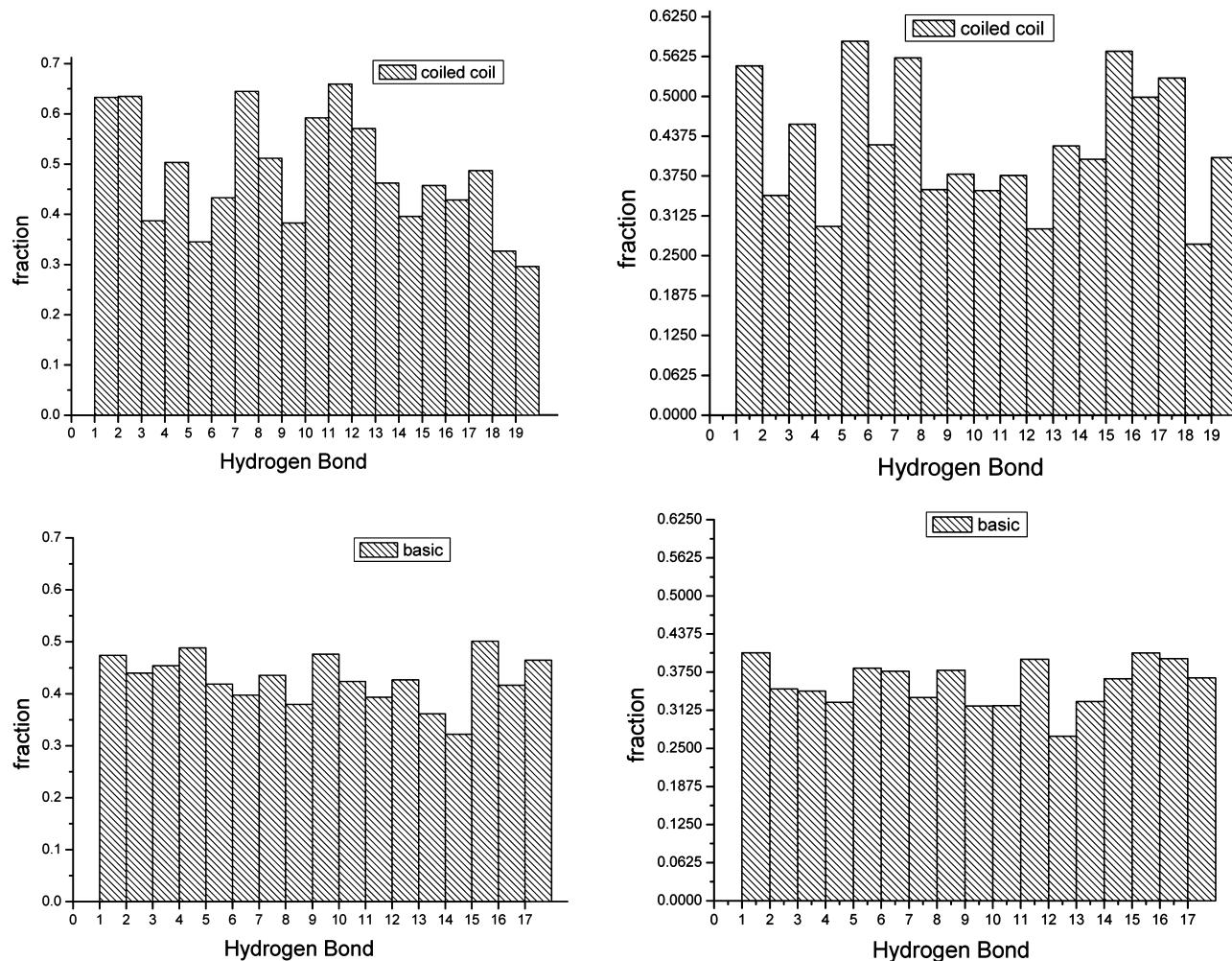
temperature for the 0 mM data. More interestingly, it appears that at each temperature, the 200 mM average distance is smaller than that of the 0 mM distance, indicating some suppression of the basic region opening arising at the higher ion concentration.

**3.2. Ion–Residue Salt Bridge 0 and 200 mM Comparison for T303.** The fraction (over the trajectory) of salt bridged  $\text{Cl}^-$  ions to the side chains of the positively (Arg, Lys) and negatively (Glu, Asp) charged residues is shown in Figure 13. Only the basic region residues are displayed as the corresponding plot for the coiled coil has all fractions less than 0.05. Figure 14 shows the  $\text{Na}^+$  and  $\text{Cl}^-$  fraction SBs for the T303\_200mM\_1 and T303\_200mM\_2 trajectories. Recall that 0 mM has 15  $\text{Cl}^-$  ions, required to neutralize bZIP, and 200 mM has an additional 84  $\text{Na}^+$  and 84  $\text{Cl}^-$  ions. The fraction SB is obtained from the record of the closest ion to an atom in each charged residue side chain in a trajectory snapshot (see section 2.3). Thus, at any instant, there is only one ion considered to be salt bridged to a residue. As shown below, the ion solvation is transient. Thus, it is not useful to look for a specific (numbered) ion being salt bridged; what matters is how often a generic ion is salt bridged to a charged residue. Figure 13 shows that the maximum fraction SB is around 0.55, with fractions concentrated on Arg16, Arg19, Arg21, Lys22, and Arg8. For 200 mM, Figure 14 shows that  $\text{Cl}^-$  is totally dominant in the basic region, with a somewhat enhanced maximum occupancy at  $\sim 0.7$ . The coiled coil has maximum

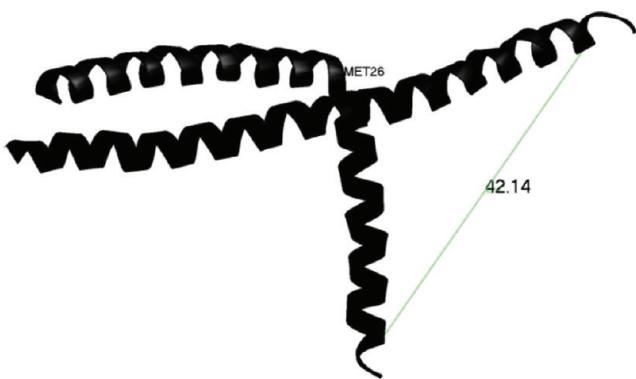
occupancy at  $\sim 0.4$ , and for  $\text{Na}^+$  ions, mainly Asp31, Glu34, Glu35, and Glu56 participate, while Lys27 (close to the basic region) has mostly  $\text{Cl}^-$  ions along with some other Lys and Arg49.

There is a dramatic difference in the 0 versus 200 mM fraction SB data. The coiled coil region goes from no  $\text{Cl}^-$  ions around it at 0 mM to a finite  $\text{Cl}^-$  fraction SB at 200 mM. Note that the coiled coil region is neutral (seven positive and seven negative residues per monomer). Surprisingly, at 0 mM, there are no  $\text{Cl}^-$  ions around the coiled coil; all are around the basic region. While the basic region is much more positive than the coiled coil, there are seven positive residues in the coiled coil that could attract  $\text{Cl}^-$  ions. At 200 mM, with the introduction of more  $\text{Cl}^-$  ions and now  $\text{Na}^+$  ions, a reasonable fraction SB forms around the coiled coil. While it is mainly  $\text{Na}^+$  ions around the negative residues, there also are  $\text{Cl}^-$  ions around some of the positive residues. Thus, the presence of the  $\text{Na}^+$  ions not only of course populates the coiled coil acid residues but also permits  $\text{Cl}^-$  ions to populate the basic residues of the coiled coil.

Another, complementary, view of the ion population around the residues can be obtained by calculating the number of ions that are within some distance of specified atoms in the charged residues. We have done so without distinguishing  $\pm$  charged ions in Figure 15 for the T303\_200mM\_1 and T303\_200mM\_2 data. This measure is similar to a radial distribution function that would count all ions within a given

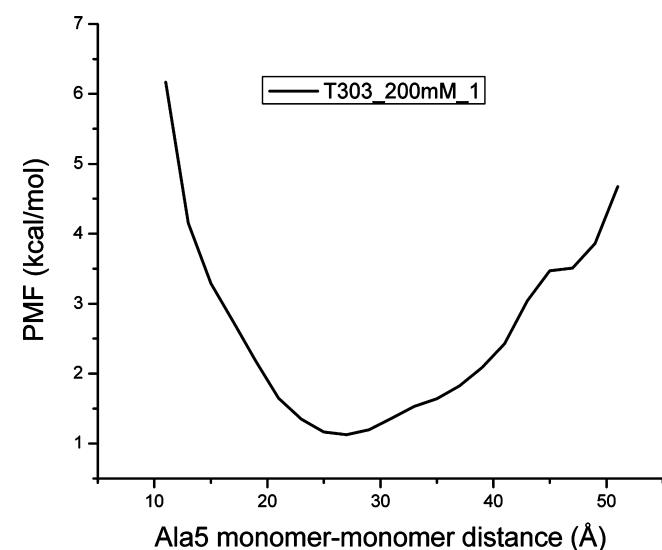


**Figure 9.** (Left) T303\_200mM\_1; (Right) T363\_200mM\_1. Fraction (over the trajectory) of 1–4 hydrogen bonds for the 19 coiled coil (30–34 to 48–52) and 17 basic region (5–9 to 21–25) residues of monomer 1. The criteria for a hydrogen bond are a 1–4 carbonyl oxygen to backbone nitrogen distance less than 3 Å and the corresponding OHN angle between 0 and 30°.

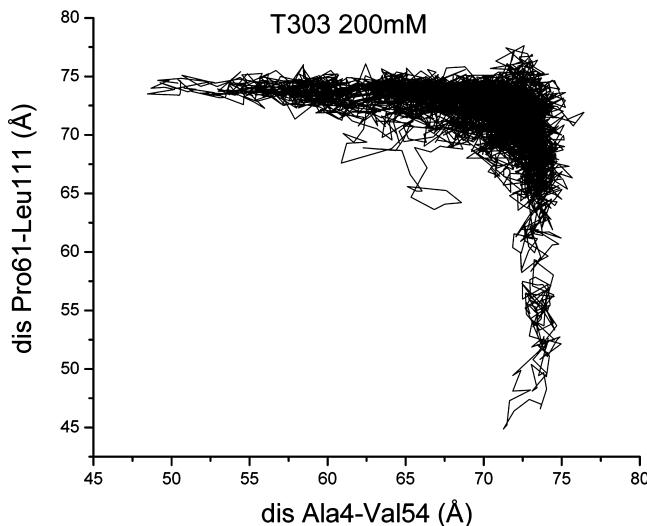


**Figure 10.** T303\_200mM\_1. A snapshot from this trajectory showing that there is a hinge spanning residues 25–27 around Met26 that leads to an open configuration of the basic region while the coiled coil is unperturbed.

radius around some central point. Here, it is implemented specific to the different residues with their differing charge centers. In particular, for Arg(Glu/Asp)[Lys], there are 4(2)[1] specified candidate salt bridging atoms NZ, CZ, NH1, NH2(OE1,OE2/OD1,OD2)[NZ]. Thus, an Arg is 2(4) times more “powerful” than a Glu(Lys). If the Arg values are divided



**Figure 11.** The potential of mean force (PMF) for T303\_200mM\_1 obtained by histogramming the Ala5 monomer-to-monomer distance fluctuations displayed in Figure 7.



**Figure 12.** Parametric plot of the end-to-end distance of the two monomers for T303\_200mM\_2. When one monomer end-to-end distance shortens, the other stays the same, showing that the basic region is undergoing an alligator-jaw-like fluctuation whereby only one monomer bends.

**Table 2. Comparison of the Average Distance between the Basic Region Monomers**

T	ave ALA5 dis (Å) 0 mM <sup>a</sup>	ave ALA5 dis (Å) 200 mM <sup>a</sup>
303	32.36629 <sup>b</sup>	28.54486; 29.44246
333	36.38756; 35.84167	30.97229; 27.94399
363	38.73118; 39.31629	30.40565 <sup>c</sup>

<sup>a</sup>The first number is for trajectory 1, and the second is for trajectory 2.

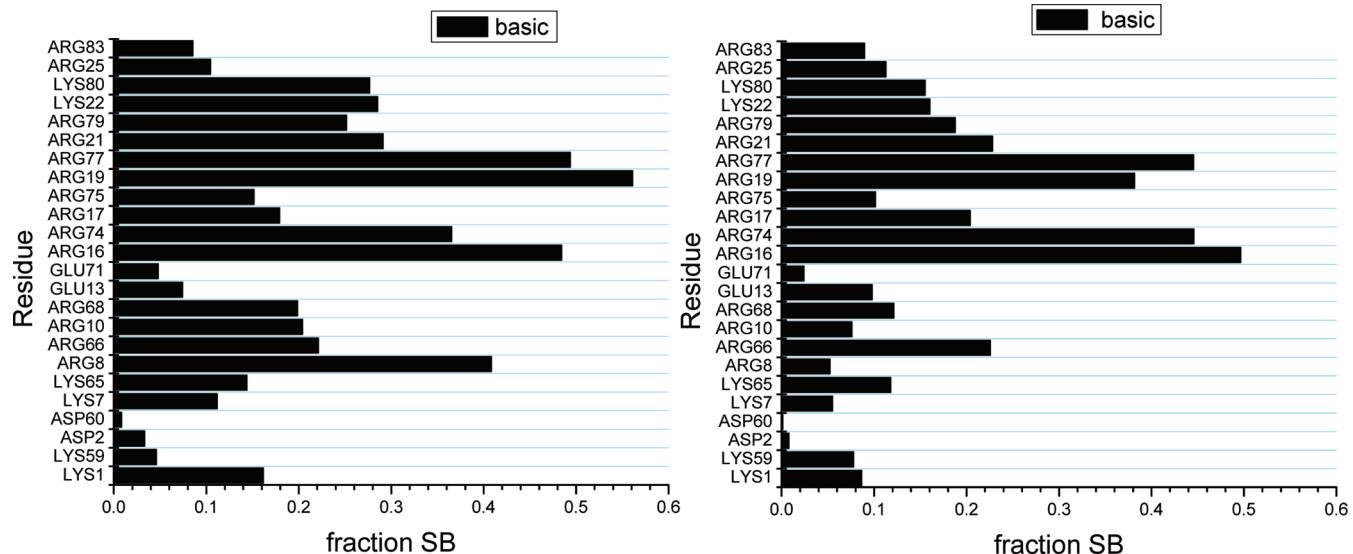
<sup>b</sup>Trajectory 1 closed the basic region. <sup>c</sup>Trajectory 2 closed the basic region.

by 4, the result should be in qualitative accord with the corresponding Arg residue in the fraction SB plot of Figure 14. Comparison of Figures 14 and 15 does show a close correspondence between the patterns across both the basic

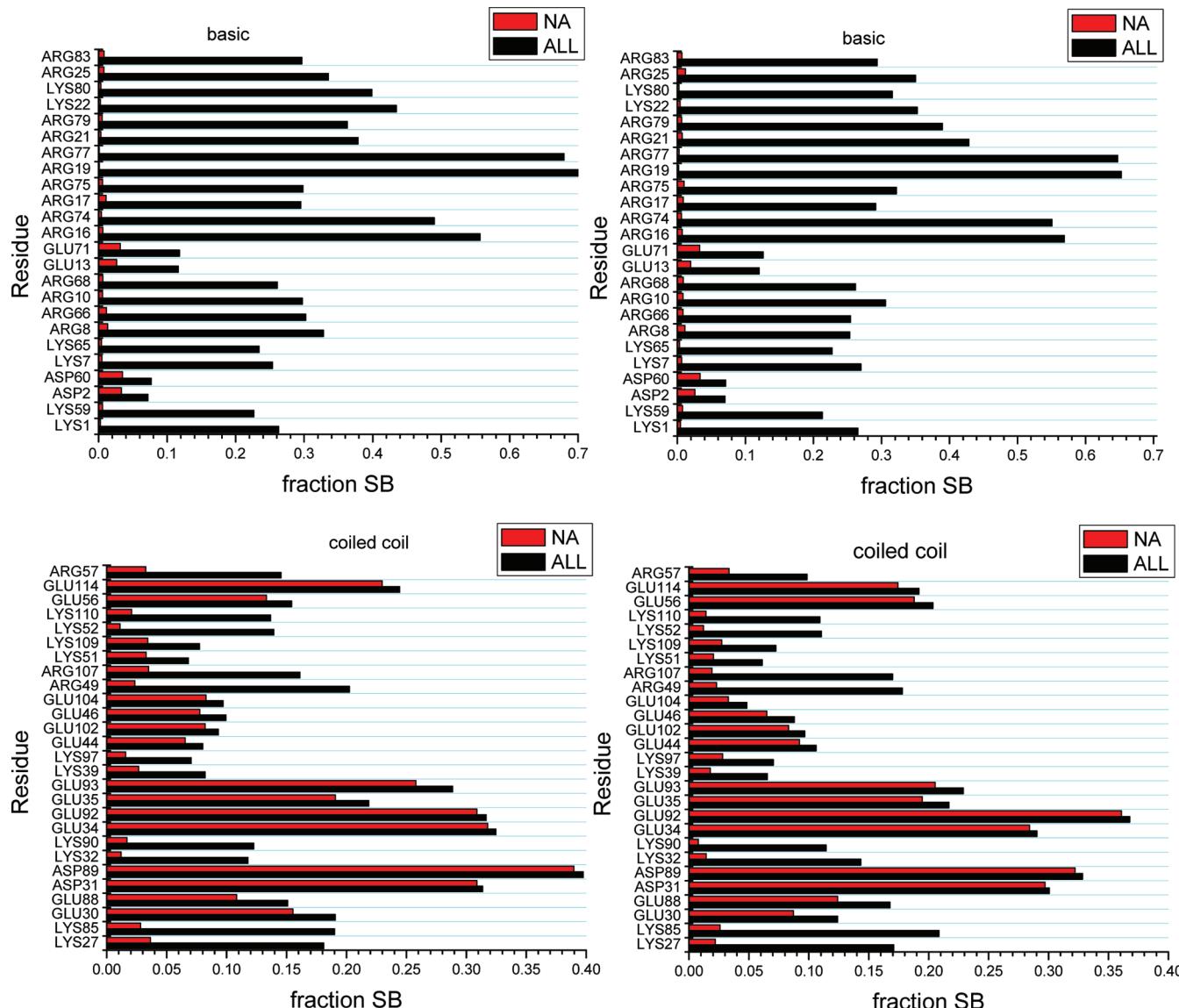
region and coiled coil. Thus, both methods support the found pattern of residue ion solvation.

The ion solvation of the residues is quite transient, as we now explore. The data that underlie the fraction SB plots in, for example, Figure 14 are records of which ion is closest to a charged residue atom at each trajectory snapshot. By tracking the lengths of time an ion is salt bridged to a given residue before being replaced by some other ion and then forming a histogram of these dwell times, a view of the dynamics of the ion solvation can be produced. Assuming a Poisson arrival process, with  $N_t$  as the number of arrivals in  $[0,t]$ , the probability of no new arrival (still the same ion) would follow  $P(N_t = 0) = e^{-t/\tau}$ , with  $\tau$  being the mean ion residence time.<sup>35</sup> Thus,  $P(N_t = 0)$  is what should be histogrammed. Figure 16 shows the (log) dwell time histogram for the average of Arg19 (monomer 1) and Arg77 (monomer 2) for T303\_0mM. Excluding the first (shortest) and the long tail dwell times, the slope provides  $\tau$ . Table 3 lists the mean residence times for a few residues in the basic region for 0 mM and both the basic region and coiled coil for 200 mM, selected because these residues have large fraction SBs. Over the length of the trajectory, about one-half of the ions do salt bridge to, for example, Arg19, though the participation is very nonuniform. For Arg19/77, the mean ion residence time shortens in comparing 0 with 200 mM. The ion participation for Arg19/77 at 200 mM involves about 1/10th of the Cl<sup>-</sup> ions. In the coiled coil, the well-populated Asp and Glu residues have 8–10 ps residence times, and they also involve about 1/10th of the ions, with some evidence for more long (1 ns) residence times. Thus, long dwell times are very rare, one or two ~1 ns dwells in a 50 ns trajectory. In the main, the ion solvation is very transient, and the solvation of a particular residue involves many different ions.

**3.3. Ion–Residue Salt Bridge Temperature Comparison for 0 and 200 mM.** The 0 mM ion solvation pattern for the higher-temperature data (shown in Figure S12, Supporting Information) is remarkably similar to that displayed for T303\_0mM in Figure 13. The difference is the scale of solvation, with the maximum solvation for Arg19 falling from ~0.55 (T303) to 0.4 (T333) to 0.35 (T363). Similar to the



**Figure 13.** (Left) T303\_0mM\_1; (Right) T303\_0mM\_2. The fraction of Cl<sup>-</sup> ions that are salt-bridged to the basic region charged residues. The “neighbor” residues, for example, Lys1 and Lys59, are equivalent residues of the two monomers.



**Figure 14.** (Left) T303\_200mM\_1; (Right) T303\_200mM\_2. The fraction of  $\text{Na}^+$  and  $\text{Cl}^-$  ions that are salt-bridged to the basic region and coiled coil charged residues. The “neighbor” residues, for example, Lys1 and Lys59, are equivalent residues of the two monomers.

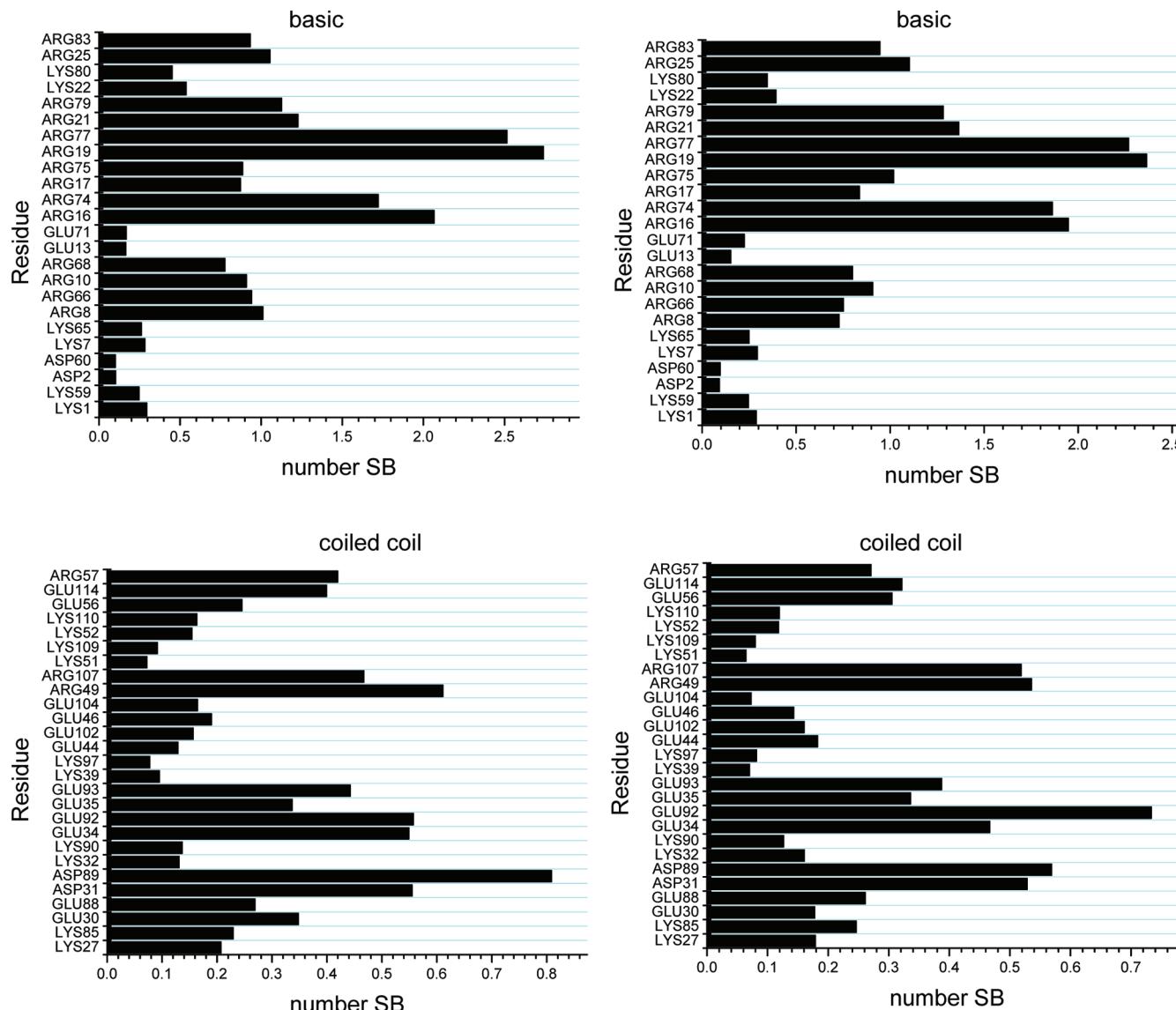
T303 fraction SB data, there is essentially no coiled coil salt bridge population. Thus, there is a temperature effect, but the same basic region residues that have large population at T303 still do so at the higher temperatures. The 0 mM mean ion residence time data in Table 3 shows that at T363, the dwell time shortens by a factor of about 3 relative to T303, indicating again the role of temperature.

For 200 mM, the higher-temperature fraction SB data are essentially identical to those for T303. For example, comparison of Figure 14 for T303 with the T363 data (Figure S13, Supporting Information) shows that the basic region maximum fraction SB is still  $\sim 0.7$  for Arg19/77, and the maximum coiled coil fraction SB is  $\sim 0.4$  for Asp31/89. Comparing the fraction SBs of the residues for the two trajectories TX\_200mM\_1 and TX\_200mM\_2 at T303 and at T363 shows that the extent of the small fluctuations in these fraction SBs is on the same scale as the differences in fraction SBs between the two temperatures. Similar conclusions apply to the T333\_200mM data (not shown). Thus, the 200 mM ion concentration leads to essentially indistinguishable temperature

behavior for both the basic region and coiled coil. The other measure of ion solvation, the number SB, displayed for T303 in Figure 15 also produces a similar number SB plot for T333 (Figure S14, Supporting Information).

The dwell time comparison for T303 and T363 for 200 mM shows that for the basic region, the mean ion residence time may shorten, and for the coiled coil, it does so modestly. It should be noted that a 1 ps residence time is at the limit of detection because the trajectory data is only written out every 1 ps. It may in fact be shorter for this higher-temperature trajectory.

Comparison of the T303 basic region ion solvation pattern between 0 (Figure 13) and 200 mM (Figure 14) shows that they are quite similar. The 200 mM data do have a somewhat higher maximum fraction SB of  $\sim 0.7$  versus  $\sim 0.55$  for 0 mM. As the temperature increases, the maximum basic region ion solvation decreases for 0 mM and remains the same for 200 mM. The mean residence times in the basic region are longer at 0 mM versus those at 200 mM, with the ratio  $\tau(0 \text{ mM})/\tau(200 \text{ mM})$  decreasing as the temperature increases.

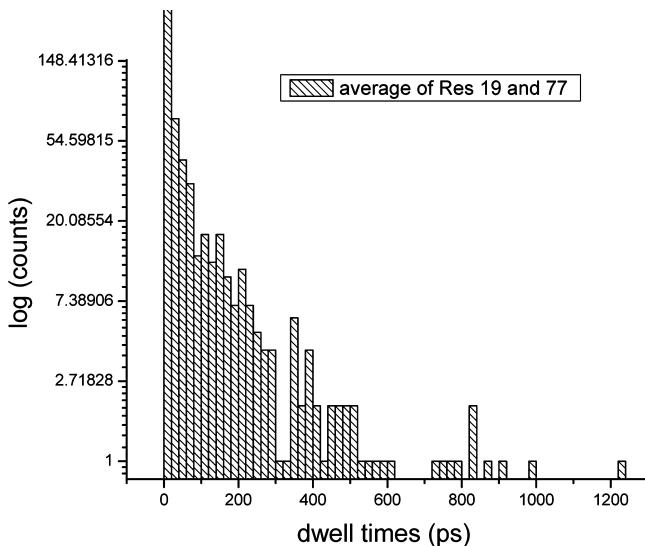


**Figure 15.** (Left) T303\_200mM\_1; (Right) T303\_200mM\_2. The relative number of ions close to the indicated charged residues for the basic region and coiled coil. When account is taken of the numbers of charge sites in the residues (see text for details), the results confirm the ion solvation pattern in Figure 14.

**3.4. Residue–Residue Salt Bridges.** Salt bridges between residues have been suggested to be important to leucine zipper stability.<sup>36,37</sup> However, in the bZIP crystal structure,<sup>4</sup> there are very few salt bridges, and they are concentrated at the carboxy terminus of the coiled coil, leading to the conclusion<sup>4</sup> that bZIP stability is primarily from van der Waals interactions. In leucine zippers, the most likely candidates for salt bridges are between oppositely charged side chains of residues  $i$  and  $i + 3$ ,  $i + 4$  on the same monomer and between residues  $i'(g)$  and  $i + 5(e)$  (for parallel coiled coils) on different monomers. Examination of the potential four intramonomer and four intermonomer salt bridges in the bZIP coiled coil for the T303 0 and 200 mM trajectories shows that the intramonomer Glu–Lys and Glu–Arg (in each monomer) salt bridges are present for about 50% of the trajectory. The intermonomer ones are at most present 25% of the time. Thus, in agreement with the crystal structure data, salt bridges do not seem critical to the bZIP coiled coil stability.

#### 4. DISCUSSION

**4.1. Collapse Trajectories.** Of the six trajectories, two (T303\_0mM\_1 and T363\_200mM\_2) show “anomalous” behavior whereby the basic region collapses to small monomer–monomer distances. These events occur on a relatively long (greater than 5 ns) MD time scale. For T303\_0mM\_1, the collapse may be due to the low ion concentration. For T363\_200mM\_2, the collapse may be due to the elevated temperature. Clearly, one would have to run many more trajectories to know how often this occurs. Once collapsed, the mean time for separation would be very long on a MD time scale, because there is no evidence for reseparation in these trajectories. For the other trajectories that sample a broad, noncollapsed set of configurations, Figure 11 indicates that the basic region monomer separation oscillates around the distance appropriate to the crystal structure.<sup>4</sup> From the perspective of the ion solvation, the T303\_0mM fraction SB data are essentially independent of the collapsed or not collapsed trajectory. The same is true of the calculated residence times.



**Figure 16.** Dwell time distribution for ions salt bridged to Arg 19 (monomer 1) and Arg 77 (monomer 2) averaged over the monomers. See the text for details.

**Table 3. Average Ion Dwell Times for Selected Residues**

T (K)	0 mM $\tau$ (ps); residue	200 mM $\tau$ (ps); residue
303	15–20; Arg19/77	2; Arg19/77 8–10; Asp31/89 8–10; Glu34
363	5–6; Arg19/77	1–2 <sup>a</sup> ; Arg19/77 5–6; Asp31/89 6–8; Glu34

<sup>a</sup>Three out of four are 1 ps.

The T363\_200mM fraction SBs and residence times in both the basic region and coiled coil are also refractory to collapse or not collapse. Thus, from the perspective of ion solvation patterns or residence times, the anomalous trajectories do not discriminate between the structurally different trajectories.

At first glance, it is counterintuitive that the basic region could collapse in view of the 17 positively charged residues (net charge of +13 between the two monomers). Examination of snapshots from the collapsed part of the T303\_0mM\_1 trajectory shows that the basic region takes on the appearance of a porcupine, with the quills, the Arg and Lys side chains, sticking far out into the solvent. Clearly, there is a strong driving force to solvate these charged residues. That permits the packing forces in the basic region to operate and most likely leads to these configurations. While, again, only many more trajectories could address the issue definitively, it seems that there is a competition between the alligator-jaw-like motion based states, driven in part by entropic gain, and the collapsed states driven by charged residue solvation and monomer–monomer interactions.

**4.2. Fluctuating Trajectories.** For the “normal” trajectories where the basic region monomer–monomer distances fluctuate around the crystal structure separation, broad ensembles of conformations are found that, presumably, could be used to trap a DNA ligand. The mechanism would fall into the category of conformational selection,<sup>17–19</sup> whereby a broad set of distances is sampled for ligand binding, versus induced fit, with a more-or-less preformed basic region conformation. The motion is alligator-jaw-like (only one

monomer bends), as is clear from the correlation plot in Figure 12. There is an essentially rigid body, hinge motion pivoting around residues Gln24–Arg25–Met26 centered on Arg 2S, as supported by the DSSP and the DynDom analyses of bent snapshots, illustrated in Figure 10. Furthermore, the first PCA mode corresponds to about 2/3 of the overall mean-square fluctuation of each monomer, and this mode’s time dependence correlates well with the end-to-end monomer distance trajectory. Thus, a picture emerges of a stable coiled coil with a fixed extension of one monomer and with the other monomer acting as a rigid body that moves to provide varying distances for capturing DNA. The three residues that span the transition between the basic region and the coiled coil, referred to as a “fork” region, have also been identified in other, heterodimer bZIPs as important to entraining DNA.<sup>20,21</sup> The key role of Met26 has been noted before. In particular, EPR studies of systematically spin-labeled GCN4 have shown that when Met26 is labeled, this construct leads to unstable dimers,<sup>9</sup> while Alber<sup>38</sup> noted that it may be important to the leucine zipper opening. Our simulations suggest that an important role of the fork region is to provide the pivot point for the hinge motion that produces a set of conformations for DNA binding.

The PMF (Figure 11) for this monomer–monomer distance sampling does not show evidence of stable states other than the minimum around the crystal structure distance. Rather, a broad range of distances is accessible that is consonant with the absence of particular interactions for the sampled basic region separations. While not dramatic at T303 but clear at the higher temperatures, the AlaS average intermonomer distances listed in Table 2 show that the 200 mM data are more constrained than the 0 mM data, indicating a possible role of the ion concentration on the basic region sampling.

**4.3. Ion–Residue Interactions.** The residence time analysis summarized in Table 3 shows that the ion solvation is very transient. While there are a few long (order of nanoseconds) dwell times, the mean residence time  $\tau$  is a few picoseconds, with the time shortening as temperature increases. The shortening with the increase in ion concentration is in part a function of the approximately 5-fold-higher Cl<sup>−</sup> ion concentration at 200 mM. However, the concentration effect should scale with proximity, which scales with the cube root of the concentration ratio; thus, the residence time shortening factor should be ~1.6. The shortening from, for example, 15 to 2 ps may in part be due to the presence of the Na<sup>+</sup> ions as competitors with positive residues such as arginine for Cl<sup>−</sup> ions.

Because there is no evidence for persistent ion association with a specific residue, whereby there would be a specific structural implication as in an enzyme that requires a particular ion–residue association, it does not make sense to look for a specific (numbered) ion being salt bridged to charged residues. What is a useful metric is the fraction of time a generic ion is salt bridged to a charged residue in the trajectory, as recorded by our fraction ion plots in Figures 13 and 14 and Figures S12–S14 in the Supporting Information. (Our other metric, related to a radial distribution function, the average over the trajectory of the number of ions associated with charged residues, confirms the fraction ion results, as shown in Figures 15 and S14 in the Supporting Information.) The pattern of which residues have the largest fraction SB is quite similar when comparing the different temperatures. Remarkably, the pattern of specific (mainly arginine) residues for the basic region persists even when comparing the fraction SBs for the different ion concentrations. In this sense, there is specificity in which

residues have the largest fraction SB, though the solvation is very transient.

That proteins bind ions with specificity in residue type and, in particular, residue has been explored by extensive MD simulations and by a survey of the pdb.<sup>39</sup> The pdb survey shows that  $\text{Na}^+$  binds preferentially in the order Asp > Glu, while  $\text{Cl}^-$  ions bind in the order Arg > Lys, as found consistently over our simulations. The simulations carried out<sup>39</sup> also find specificity as to which residues of a certain type will bind ions, consistent with our findings.

Comparing the T303  $\text{Cl}^-$  ion solvation at 0 and 200 mM for the basic region shows only a slight increase in the fraction SB. This means that the basic region has some basic residues that become ion saturated even at 0 mM; there are enough  $\text{Cl}^-$  ions to satisfy the basic residues. For the higher temperatures, the 0 mM fraction SBs decrease, while those of 200 mM more or less remain the same, thus not saturating at these temperatures. The higher  $\text{Cl}^-$  ion concentration for 200 mM can maintain the basic region ion atmosphere similar to that found for T303.

At 200 mM, the introduction of more  $\text{Cl}^-$  ions and the  $\text{Na}^+$  ions produces a substantial fraction SB that is mainly  $\text{Na}^+$  around the coiled coil acidic residues but with  $\text{Cl}^-$  ions around some of the coiled coil basic residues. Thus, there is a dramatic difference in 0 versus 200 mM. The coiled coil goes from no  $\text{Cl}^-$  ions around it at 0 mM to a finite  $\text{Cl}^-$  fraction SB at 200 mM. The presence of the  $\text{Na}^+$  ions not only, of course, populates the coiled coil acid residues but also permits  $\text{Cl}^-$  ions to populate the basic residues of the coiled coil.

Our focus has been on ion solvation of the side chains of the charged residues. Not surprisingly, an examination of the fraction SB metric for the backbone atoms shows that they are not involved in ion solvation, with the exception of fraction SBs at  $\sim 0.1$  for some of the residues at the ends of monomers for T303\_200mM. Thus, it is safe to conclude that the ion solvation effects are confined to the charged residue side chains.

**4.4. Dimer and  $\alpha$ -Helix Stability.** The bZIP simulations were initiated from the crystal structure of GCN4 with bound DNA.<sup>4</sup> In solution, it is known that the bZIP dimer is unstable at low concentration but has a dimer population at higher concentrations,<sup>12</sup> while leucine zippers (the coiled coil part of bZIP) by themselves have dimers strongly favored.<sup>40,41</sup> The suggested mechanism of leucine zipper dimerization has either unstructured or partially structured monomers coming together and forming their super helix and monomer  $\alpha$ -helices in the course of dimerization.<sup>36,37,42,43</sup> The sequencing of the events of  $\alpha$ -helical formation of the monomers and dimerization is not simple to unravel as  $\alpha$ -helix formation before and after monomer encounter can lead to similar folding kinetics.<sup>44</sup> The same issues arise in the formation of bZIP dimers in the absence of DNA. Presumably, in a MD trajectory, the dimer could separate in a sufficiently long trajectory. However, estimates of the dimer lifetime<sup>12</sup> are greater than 1 ms and up to 1 s and would preclude such events from occurring on a realistic MD time scale.

Simulations of leucine zipper formation started from different monomer distances/angles with  $\alpha$ -helical monomers and led to dimer formation only when the monomers were initially close.<sup>45</sup> Once formed, they were stable on the 10 ns simulation time scale. Longer (50–100 ns) MD simulations<sup>46</sup> on the C- and N-terminal halves of the leucine zipper monomer from GCN4 for a variety of  $\text{pK}_a$  and ion concentrations show that these fragments are mainly  $\alpha$ -helical. On these MD time scales, there is scant evidence for helix unfolding. Our enhanced

sampling MD simulations of melting curves<sup>47</sup> and dimer formation<sup>48</sup> of a GCN4 leucine zipper show that it is very stable. Thus, it is not a surprise that on a 50 ns MD time scale for all six trajectories, the bZIP dimer always remains dimerized.

Regarding the  $\alpha$ -helical character of the basic region, as noted by Oakley<sup>49</sup> and co-workers, the CD spectrum of GCN4 implies that it is  $\sim 75\%$   $\alpha$ -helical. They point out that, allowing for the unstructured residues of the basic region and the coiled coil, if the basic region were unstructured, the percentage would be around 50% instead of 75%, providing evidence for  $\alpha$ -helical character of the basic region. Other heterodimeric bZIPS can have the basic region of one monomer  $\alpha$ -helical.<sup>21</sup> The rmsf plots in, for example, Figure 2 and all of the others certainly show that the C-terminal ends of the coiled coil monomers are not  $\alpha$ -helical, nor are the N-terminal ends of the basic region. Solution studies<sup>8–12</sup> concluded that the basic region fluctuates around  $\alpha$ -helical structures. Indeed, at low temperature, GCN4 is mostly  $\alpha$ -helical.<sup>12</sup> Thus, the picture emerges of a stable  $\alpha$ -helical coiled coil and a basic region that fluctuates around  $\alpha$ -helical monomers. A comparison of our rmsf plots (e.g., Figure 2) with plots of NMR relaxation data<sup>8</sup> and EPR spin labeled data<sup>9</sup> that also reflect the fluctuations of the residues are very similar. Namely, starting at the N-terminus of the basic region, the residue rmsfs are large and decrease until they reach fairly constant values over most of the coiled coil region, and then, they increase again as the C terminus is approached. In the coiled coil, the MD and NMR/EPR data are unequivocally pointing to  $\alpha$ -helical behavior, except for the terminal residues. For the basic region, the NMR and EPR data are ambiguous in the sense that even if the (interior) residues are  $\alpha$ -helical, the residue rmsfs can be large from basic region monomer fluctuations. Of course, once DNA binds, the conformational space of the basic region is greatly limited, and a very specific set of interactions are formed, with the basic region  $\alpha$ -helical.<sup>4</sup>

The 1–4 hydrogen bond data analyses displayed in Figures 9 and S10 (Supporting Information) show only modest evidence for the basic region being significantly less  $\alpha$ -helical than the coiled coil, with the 200 mM\_T363 data showing the greatest distinction between the two subdomains. From the lengths of the trajectories, it would be rather surprising to see a dramatic transition from the mainly  $\alpha$ -helical character of the basic region for longer MD times, but of course, that cannot be ruled out. Supposing that such a transition occurred, regaining  $\alpha$ -helical character would be similar to folding a protein; thus, it could take on the order of microseconds or greater of simulation time to probe such events.

## 5. CONCLUSIONS

The multiple-trajectory bZIP simulations for combinations of three temperatures and two ion concentrations carried out lead to qualitatively similar results, with the exception of the collapse trajectories. Despite the highly charged basic region, there is a modest influence from ion concentration on the conformational sampling of this subdomain with regard to the monomer–monomer distance and monomer structure fluctuations. The same can be said of the effect of temperature. Thus, the simulations suggest that the structure and fluctuations of bZIP are quite robust, as may be required for function in a variety of conditions.

In order for DNA to be entrained by bZIP, the basic region has to be in a relatively open conformation. The simulations show that the basic region undergoes an essentially rigid body, hinge motion centered on Arg 2S in the fork region. The PMF

for the hinge motion suggests that the monomer-to-monomer basic region fluctuations may provide the broad variety of distances required for DNA capture. These results favor a conformational selection<sup>17–19</sup> mechanism whereby ligand binding occurs by sampling from a broad conformational distribution. There is inferential experimental evidence for conformational selection from measurements of residue-by-residue fluctuations. NMR relaxation measurements<sup>8</sup> show that spectral densities of the basic region residues fall sharply from the N-terminus to the fork region, which may reflect the monomer-to-monomer distance fluctuations. A similar effect was found in EPR<sup>9</sup> by spin-labeling selected sites along the bZIP sequence.

No evidence for stable salt bridges between a residue and an ion was found, as monitored by the evaluation of residence times. While some residues are substantially ion-solvated, the lifetime of a particular ion is very short. In a dimer formed from mainly  $\alpha$ -helical monomers, the charged residues are mostly solvent-exposed, versus buried in the protein's interior, and are solvated by generic ions in a transient fashion.

In bZIP, it should be noted that the coiled coil subdomain, while overall neutral, has numerous charged residues. An intriguing effect of ion concentration is that, even so, at 0 mM, the  $\text{Cl}^-$  ions (required to neutralize the basic region) only populate the basic region. However, at 200 mM, the presence of  $\text{Na}^+$  ions not only populates the coiled coil acid residues but also permits  $\text{Cl}^-$  ions to populate the basic residues of the coiled coil.

One can speculate as to the various energetic and entropic considerations that contribute to DNA binding. As has been pointed out in many contexts,<sup>13,50</sup> including transcription factors,<sup>51</sup> it may be advantageous for DNA binding to select from preformed  $\alpha$ -helical-like basic region monomers that broadly sample a variety of intermonomer distances. On the other hand, it may be more favorable energetically for the basic region monomers to require the presence of DNA to induce  $\alpha$ -helical formation in the process of binding to DNA. The results of these simulations do favor the former mechanism.

## ASSOCIATED CONTENT

### Supporting Information

Plots of rmst, rmsd, rmsf, monomer–monomer CA distances, hydrogen bond fractions, potential of mean force, and fraction and number of residue-to-ion salt bridges for some temperatures and ion concentrations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: cukier@chemistry.msu.edu. Phone: 517-355-9715 x 263. Fax: 517-353-1793.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

Computational support from the Michigan State University High Performance Computing Center and the Department of Chemistry is gratefully acknowledged.

## REFERENCES

- (1) Latchman, D. S. *Int. J. Biochem. Cell Biol.* **1997**, *29*, 1305–1312.
- (2) Miller, M. *Curr. Protein Pept. Sci.* **2009**, *10*, 244–269.
- (3) Vinson, C.; Myakishev, M.; Acharya, A.; Mir, A. A.; Moll, J. R.; Bonovich, M. *Mol. Cell. Biol.* **2002**, *22*, 6321–6335.
- (4) Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223–1237.
- (5) Formaneck, M. S.; Ma, L.; Cui, Q. *J. Am. Chem. Soc.* **2006**, *128*, 9506–9517.
- (6) Ibragimova, G. T.; Wade, R. C. *Biophys. J.* **1999**, *77*, 2191–2198.
- (7) Pfeiffer, S.; Fushman, D.; Cowburn, D. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 206–217.
- (8) Bracken, C.; Carr, P. A.; Cavanagh, J.; Palmer, A. G. *J. Mol. Biol.* **1999**, *285*, 2133–2146.
- (9) Columbus, L.; Hubbell, W. L. *Biochemistry* **2004**, *43*, 7273–7287.
- (10) Saudek, V.; Pasley, H. S.; Gibson, T.; Gausepohl, H.; Frank, R.; Pastore, A. *Biochemistry* **1991**, *30*, 1310–1317.
- (11) Saudek, V.; Pastore, A.; Morelli, M. A. C.; Frank, R.; Gausepohl, H.; Gibson, T.; Weil, F.; Roesch, P. *Protein Eng.* **1990**, *4*, 3–10.
- (12) Weiss, M. A.; Ellenberger, T.; Wobbe, C. R.; Lee, J. P.; Harrison, S. C.; Struhl, K. *Nature* **1990**, *347*, 575–578.
- (13) Mohan, A.; Oldfield, C. J.; Radivojac, P.; Vacic, V.; Cortese, M. S.; Dunker, A. K.; Uversky, V. N. *J. Mol. Biol.* **2006**, *362*, 1043–1059.
- (14) Schlessinger, A.; Schaefer, C.; Vicedo, E.; Schmidberger, M.; Punta, M.; Rost, B. *Curr. Opin. Struct. Biol.* **2011**, *21*, 412–418.
- (15) Xie, H. B.; Vučetić, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Uversky, V. N.; Obradović, Z. *J. Proteome Res.* **2007**, *6*, 1882–1898.
- (16) Prakash, M. K. *J. Am. Chem. Soc.* **2011**, *133*, 9976–9979.
- (17) Bahar, I.; Chennubhotla, C.; Tobi, D. *Curr. Opin. Struct. Biol.* **2007**, *17*, 633–640.
- (18) Goh, C. S.; Milburn, D.; Gerstein, M. *Curr. Opin. Struct. Biol.* **2004**, *14*, 104–109.
- (19) Kumar, S.; Ma, B. Y.; Tsai, C. J.; Sinha, N.; Nussinov, R. *Protein Sci.* **2000**, *9*, 10–19.
- (20) Kim, J.; Tzamarias, D.; Ellenberger, T.; Harrison, S. C.; Struhl, K. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4513–4517.
- (21) Podust, L. M.; Krezel, A. M.; Kim, Y. *J. Biol. Chem.* **2001**, *276*, 505–513.
- (22) Lou, H. F.; Cukier, R. I. *J. Phys. Chem. B* **2006**, *110*, 24121–24137.
- (23) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P. *Biomolecular simulation: the GROMOS96 manual and user guide*; Vdf hochschulverlag AG an der ETH: Zürich, 1996.
- (24) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (25) Berendsen, H. H. C.; Postma, J. P. M.; Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (26) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (27) Lou, H.; Cukier, R. I. *Analyzer*, 2.0 ed.; Michigan State University: East Lansing, MI, 2008.
- (28) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412–425.
- (29) García, A. E. *Phys. Rev. Lett.* **1992**, *68*, 2696–2699.
- (30) Jolliffe, I. T. *Principal Component Analysis*, 2nd ed.; Springer Science: New York, 2004.
- (31) O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. *Science* **1991**, *254*, 539–544.
- (32) Crick, F. H. C. *Acta Crystallogr.* **1953**, *6*, 689–697.
- (33) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.
- (34) Hayward, S.; Berendsen, H. J. C. *Proteins: Struct., Funct., Genet.* **1998**, *30*, 144–154.
- (35) Parzen, E. *Modern probability theory and its applications*; Wiley: New York, 1960.
- (36) Matousek, W. M.; Ciani, B.; Fitch, C. A.; Garcia-Moreno, B.; Kammerer, R. A.; Alexandrescu, A. T. *J. Mol. Biol.* **2007**, *374*, 206–219.
- (37) Steinmetz, M. O.; Jelesarov, I.; Matousek, W. M.; Honnappa, S.; Jahnke, W.; Missimer, J. H.; Frank, S.; Alexandrescu, A. T.; Kammerer, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 7062–7067.
- (38) Alber, T. *Curr. Biol.* **1993**, *3*, 182–184.

- (39) Friedman, R. *J. Phys. Chem. B* **2011**, *115*, 9213–9223.
- (40) Grigoryan, G.; Keating, A. E. *J. Mol. Biol.* **2006**, *355*, 1125–1142.
- (41) Woolfson, D. N. The Design of Coiled-Coil Structures and Assemblies. *Adv. Protein Chem.* **2005**, *70*, 80–106.
- (42) Dragan, A. I.; Privalov, P. L. *J. Mol. Biol.* **2002**, *321*, 891–908.
- (43) Mason, J. M.; Arndt, K. M. *ChemBioChem* **2004**, *5*, 170–176.
- (44) Meisner, W. K.; Sosnick, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13478–13482.
- (45) Pinero, A.; Villa, A.; Vagt, T.; Koksich, B.; Mark, A. E. *Biophys. J.* **2005**, *89*, 3701–3713.
- (46) Missimer, J. H.; Steinmetz, M. O.; Jahnke, W.; Winkler, F. K.; van Gunsteren, W. F.; Daura, X. *Chem. Biodiversity* **2005**, *2*, 1086–1104.
- (47) Su, L.; Cukier, R. I. *J. Phys. Chem. B* **2009**, *113*, 9595–9605.
- (48) Cukier, R. I. *J. Chem. Phys.* **2011**, *134*, 045104.
- (49) Hollenbeck, J. J.; McClain, D. L.; Oakley, M. G. *Protein Sci.* **2002**, *11*, 2740–2747.
- (50) Fuxreiter, M.; Simon, I.; Friedrich, P.; Tompa, P. *J. Mol. Biol.* **2004**, *338*, 1015–1026.
- (51) Dunker, A. K.; Uversky, V. N. *Curr. Opin. Pharmacol.* **2010**, *10*, 782–788.