JCTC Journal of Chemical Theory and Computation

Article

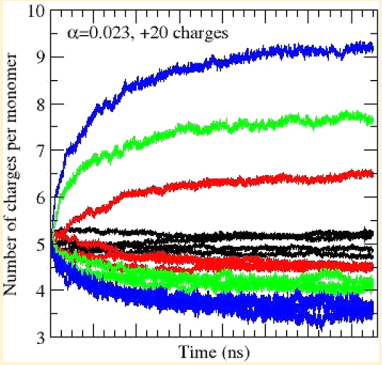
pubs.acs.org/JCTC

A Charge Moving Algorithm for Molecular Dynamics Simulations of Gas-Phase Proteins

Sarah K. Fegan and Mark Thachuk*

Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, BC, V6T 1Z1, Canada

ABSTRACT: A method for moving charges in a coarse-grained simulation of gas-phase proteins is presented which uses a Monte Carlo approach to move charges between charge sites. The method is used to study the role of charge movement in the dissociation mechanism of protein complexes in order to better understand experimentally observed mass spectra from CID studies. The charge hopping process is analyzed using energy distributions and a pair correlation plot. Hopping rates, charge distributions, and structural parameters (radius of gyration and RMSD) are also calculated. The importance of charge movement for the unfolding of protein complexes is demonstrated. The algorithm is implemented in the GROMACS molecular dynamics software package. In this study, transthyretin (TTR) tetramer is used with the MARTINI force field as a model system, and comparisons to experiments are made. The hopping and unfolding are found to be controlled by the Coulomb repulsion among the charges in the complex.



INTRODUCTION

Using mass spectrometry, protein complex dissociation producing asymmetric charge distributions has been observed in protein dimers[1] and in larger multimeric complexes.[2−5] Asymmetric charge dissociation occurs when the mass to charge ratios of fragments differ, that is one of the fragments has a larger fraction of the total charge than it has of the total mass. With slow dissociation methods, such as collisionally induced dissociation (CID), unfolding and ejection of one monomer from the complex is universally observed. Typically, the ejected monomer carries away up to half the total charge on the complex. In some cases, the remaining complexes again eject another charge-enriched monomer. With fast dissociation methods like surface induced dissociation (SID), more symmetric dissociation products are seen.[6,7] That is, single monomer ejection does not dominate, and the charge distribution of the fragments is close to what is expected for a uniformed charged protein complex ion.

In practice, usually the weakest bound monomer in a complex unfolds and dissociates upon CID activation. The predominance of this dissociation pathway excludes other pathways which may shed more light on the binding of monomers in the complex and allow one to build a better picture of their assembly. These monomer interactions are often the key to understanding the relationship between structure and biological function in these complexes. Understanding the detailed mechanism for monomer ejection will help to provide methods by which this dissociation can be controlled and possibly other dissociation pathways selected. Hopefully, this will allow for more structural information about proteins to be gathered from experiments.

A variety of modeling studies have examined charged protein complexes with various distributions of charge and structure.[8−13] The essential conclusion of these studies is that the monomer ejection process is controlled by Coulomb repulsion. In particular, charges arrange themselves so as to maintain a constant surface charge density, and the lowest barrier for dissociation occurs when two fragments have the same charge. Almost all experimental results can be rationalized with this Coulomb repulsion model.

The monomer ejection pathway requires charges to adopt low energy configurations as the monomer unfolds; hence charges must be mobile. There is experimental and theoretical confirmation that charges are mobile in small gas-phase peptides,[14−20] and a mobile proton model has been developed. For reviews of the mobile proton model, see refs 15 and 21. Peptide fragmenting D/H exchange experiments performed by Jorgensen et al.[19] were able to selectively label the C- or N-terminal ends of peptides with deuterium and found 100% scrambling on CID. However, theoretical or experimental studies of proton transfer in large proteins or protein complexes are less common.

A rigorous treatment of proton transfer involves quantum mechanical calculation of the potential energy during proton motion. Such an approach has been used to study proton motion in water (for reviews see refs 22 and 23). In the present case, using a quantum mechanical method to describe proton transfer in protein complexes is computationally very intensive. Further, the time scale for the monomer ejection pathway is much larger than the time scale for proton transfer. Thus, the precise description of this transfer is not expected to play a major role. As well, we want to use a coarse-grained potential to speed up the calculations. It is not possible to include the details of proton motion in a coarse-grained potential, because the degrees of freedom associated with protons are not present.

Received: October 19, 2012
Published: April 23, 2013

dx.doi.org/10.1021/ct300906a | J. Chem. Theory Comput. 2013, 9, 2531−2539

Donnini et al.[24] performed simulations where the protonation states were changed in order to maintain constant pH. However, in an isolated (gas-phase) protein complex, the total charge remains constant rather than the pH.

Lill and Helms created a method (Q-HOP MD) for including proton transport in molecular dynamics (MD) simulations by modeling hopping rates using transition state theory.[25] Their general procedure calculates energies and then uses those energies to determine the hopping probability. They used a multicopy method to enhance sampling at the proton transfer step. It would be technically challenging to implement this for charge hopping in protein complexes.

Instead, a simple method to incorporate moving charges into the simulation of a gas-phase protein complex that is compatible with a coarse-grained force field will be introduced. This method can be used for gas-phase proteins (no solvent). More specifically, this method will use the MARTINI force field,[26,27] which we have already shown should be a suitable starting point for gas-phase studies.[28] The GROMACS program[29−31] is used for molecular dynamics. The method involves adding a charge hopping routine to the loop over MD steps, using simple instantaneous hopping between sites and a Monte Carlo like scheme. We will apply this to study transthyretin, a system for which experimental data are available.[4,32−34]

In this paper, we first present details of the method. Results from some simulations of a homomeric tetramer are then discussed to show the nature of the hops which occur and the effects of charge hopping. In addition to structural parameters, properties specific to hops are studied, including the hopping rate, distributions of the change in energy for a hop, and correlations between donor and acceptor sites. Conclusions follow at the end.

METHOD

Overall, a charge hopping algorithm must have sites at which charges reside and criteria by which charges are moved. Since most mass spectrometry studies of protein complexes use electrospray ionization in positive ion mode, positive charges are sufficient for the current study. We chose basic amino acid residues (arginine, lysine, and histidine) and the N-terminus of chain A as sites for the charges. Thus, a positive charge can hop from one basic site to an uncharged basic site. All other amino acid residues remain neutral.

We employ two criteria for moving charges. The first is a simple distance cutoff. If the distance between the charge donor and charge acceptor sites is less than r_{cutoff}, then a hop is possible. The second criterion depends upon energy and uses a Monte Carlo like scheme to determine hopping probability.

The hopping scheme has been implemented in GROMACS[29−31] with the MARTINI force field.[26,27] The MARTINI force field is a coarse grained force field that uses one to four beads for each amino acid. It does not use partial charges, so it is easy to change the charges on the side chain beads of the amino acids when a hop occurs, since this involves simply increasing or decreasing the total charge on a single bead. In particular, no structural changes are required to change the charge state, so moving charges does not create any geometrical problems like bad contacts. More technical details of the algorithm are given below.

To organize sites, two lists are created. One is a list of all the sites containing charges. These sites act as charge donors for the hops. The other is a list of the uncharged basic sites which are available as charge acceptors.

The hopping algorithm examines the structure of the complex for possible hops before the MD algorithm propagates it to the next time step. First, for each element in the list of charged sites, the empty basic site nearest to it is found. If the distance to that site is less than r_{cutoff} the pair is added to a list of possible hops. The r_{cutoff} was set to 0.57 nm. This value was chosen because it is the diameter of a MARTINI coarse-grained bead plus 1 Å. In practice, the value of r_{cutoff} tunes the hopping rate. The larger the r_{cutoff}, the greater the number of donor/acceptor encounters and hence the greater the probability a charge hop will occur. However, having too large a value of r_{cutoff} simulates proton transfer over larger distances, which at some point is hard to physically justify. A value was chosen small enough to consider the transfer to be localized but not so small as to decrease too severely the hopping rate. Results were also calculated for r_{cutoff} = 2.0 nm, and while the hopping rates did vary (as alluded to above) the behavior of the protein structure did not. In other words, the results are not particularly sensitive to the value of r_{cutoff}. In an atomistic model, the angles between the proton donor and acceptor would affect the probability of proton transfer. However, in the coarse-grained model, this structural detail is lost, so the angles between charge donor and acceptor sites are not considered.

Second, for each pair of possible hops, the change in energy that would result from the hop is calculated. The Coulomb and proton binding energies are considered for this calculation, but not the Lennard-Jones energy because during the hop the coarse-grained beads do not move. The proton binding energy is the difference in energy between the neutral and protonated forms of the amino acid. Because this energy is not included in the MARTINI force field, it must be explicitly included in the energy calculation. Table 1 lists the values of the proton binding energies for the basic sites used in this study. It should be noted that these proton binding energies are for bare amino acids, and the self-solvation properties of the protein environment are not included. *Ab initio* calculations of the energies of bare protonated and nonprotonated amino acids were used to calculate the binding energies.[11] These calculations do not consider the environment of the amino acid, but in the energy calculation the electrostatic environment is accounted for by the Coulomb energy contribution. The change in energy is calculated using the equation

Table 1. Proton Binding Energy Values for Basic Amino Acid Sites[11]

site	$\Delta E_{binding}$ (kJ/mol)
arginine	−1028.8
histidine	−957.7
lysine	−937.2
N-terminus	−959.8

$$\Delta E = E_{final} - E_{initial} \quad (1)$$

The energy of a particular configuration is

$$E = \sum_{i=1}^{N} E_{binding,i} + \sum_{j>i} \frac{e^2}{4\pi\epsilon r_{ij}} \quad (2)$$

where $E_{binding,i}$ is the proton binding energy of the charged site i, ϵ is the dielectric constant in a vacuum, and r_{ij} is the distance

between the centers of charged sites i and j. A relative dielectric constant of 1 is used in these calculations. The value of $E_{initial}$ is the energy E of the current configuration while that of E_{final} is the energy of the configuration that would result if the charge were moved from the donor to the acceptor site. All possible hopping pairs are sorted by their ΔE values from lowest to highest.

Starting with the one of lowest ΔE, each possible hop is considered in turn. If $\Delta E \leq 0$, the charge is moved from the donor to the acceptor. If $\Delta E > 0$, then the probability of a hop occurring, p, is calculated using the Monte Carlo equation[35]

$$p = \exp\left[\frac{-\alpha \Delta E}{kT}\right] \tag{3}$$

where α is an arbitrary scaling factor, k is the Boltzmann constant, and T is the temperature. The factor α is an adjustable parameter that allows the hopping rate to be increased or decreased without changing the temperature for the dynamics of the protein. As will be shown below, it will become necessary to increase the hopping rates in this way in order to make the time scales for charge hopping events smaller. At the same time, one does not wish to increase the hopping rates by raising the temperature greatly, since this will directly affect protein dynamics. By adjusting α, the hopping rates can be manipulated independently of the temperature. The value of p is compared to a random number between 0 and 1. Only if the random number is less than p is the charge moved from donor to acceptor.

When a hop occurs, the charge in the GROMACS mdatoms record for the donor site is decreased by one and the charge for the acceptor site is increased by one. The bead indexes in the lists of charged and uncharged basic sites are then swapped. The bead which had a charge of +1 before the hop is given a charge of 0 and is moved to the list of uncharged basic sites, and what had been the corresponding uncharged basic site is given a charge of +1 and is put in the list of charged sites. In the MARTINI force field, charged beads have different Lennard-Jones parameters than uncharged beads. However, in the implementation of the algorithm used in this study, this difference was not taken into account, so the bead types were left unchanged. The total charge state of the system does not change. After each hop, the values of ΔE are recalculated for all the remaining possible hopping pairs, and the hopping pair list is resorted.

Five options have been added to the GROMACS mdp parameter file. To select moving charges, the parameter mobile_charge is set to yes (the default is no). The largest distance between beads over which a charge can move is r_{cutoff} (called mobile_cutoff in the mdp file and given a default value of 0.57 nm). The mobile_prob parameter is the scaling factor α (the default value is 1.0). The mobile charge algorithm is executed every mobile_freq MD steps (the default value is 10). The seed for the random number generator is set using mobile_seed (the default value of −1 uses time-of-day and process id to generate the seed).

The algorithm generates two output files (in addition to the standard GROMACS output). One lists the charged sites and the other gives information about all the hops considered each time the algorithm is executed. The time, temperature, number of hops considered, and the number of successful hops are given. As well, the charged donor site index, uncharged acceptor site index, ΔE, p, and whether the hop was successful are listed for each move considered.

To study the hopping, a transthyretin (TTR) tetramer (PDB ID: 3GRG)[36] was used as a model system with both high (+20) and low (+8) total charge states. TTR is a small homotetramer which has been studied experimentally[4,32−34] using mass spectrometry employing collisionally activated dissociation with total charge states ranging from +8 to +15. In our previous work[28] the MARTINI force field was found to be a bit too attractive; therefore the maximum charge state in the simulations was chosen to be +20 to increase, as an offset, the Coulomb repulsion among the charges. Simulations were performed for temperatures of 300, 400, 500, and 600 K. Increasing the temperature from 300 to 600 K has the effect of adding energy to the system giving it the energy needed to overcome the dissociation and unfolding barriers, mimicking the CID experiments. The temperature is being used to speed up the dynamics, so it does not give the true temperature dependence of the unfolding or an estimate of the experimental protein temperature. At each temperature, 10 trajectories were run with each of 10 different charge configurations for a total of 100 trajectories. In the present work, simulations were run with α = 1 and 0.023. The α = 0.023 value was chosen because at 300 K this gives a hopping probability of 0.5 or greater for ΔE values less than +75 kJ/mol (approximate midpoint of ΔE distribution for the α = 1 simulations). Each charge configuration had five positive charges placed on each monomer, and the monomers were labeled as chains A, B, C, and D. The temperature was controlled by a Nosé-Hoover thermostat.[37] A 30 fs time step was used for 75 ns simulations. This time step is the same as the one used in ref 28, and because a coarse-grained force field is used, this time step can be much larger than the 1 fs time step commonly used for all-atom simulations. For the electrostatics, no cutoff or periodic boundary conditions were used, and the relative dielectric constant was 1. Local elastic network bonds were used to stabilize flexible parts of the structure.

RESULTS AND DISCUSSION

Hopping Rate and Charge Distributions. In order to study the behavior of the hopping algorithm, a number of quantities were examined. Figure 1 shows the average hopping rate as a function of time. For each trajectory, the simulation time was divided into bins of 150 ps length, and the number of hops was counted in each bin and averaged over all trajectories. This number was then divided by 150 to calculate a hopping rate in units of hops per picosecond and was normalized for the number of charges in the system. Figure 1 shows that the hopping rate is similar for all the temperatures considered, but slightly lower at 300 K than at higher temperatures. It is higher at short times and then after about 10 ns reaches a steady state value of about 0.002 hops/ps for the high charge state and about 0.1 hops/ps for the low charge state when α = 1. This is much lower than transfer rates of 0.25 hops/ps, 0.11 hops/ps, and 0.46 hops/ps reported for an excess proton in water in refs 25, 38, and 39, respectively. The hopping rate is higher for the low charge state because that state has more available acceptor sites so each donor site has a higher probability of coming close to an acceptor site. When α is decreased to 0.023, the hopping rate increases to just above 0.1 hops/ps for the high charge state and 0.4 to 0.6 for the low charge state. The decrease in hopping rate at short times is expected because the initial charge configurations were not optimized and the charges quickly rearrange to find lower energy configurations. However, at long times the hopping rate is close to but not exactly zero.

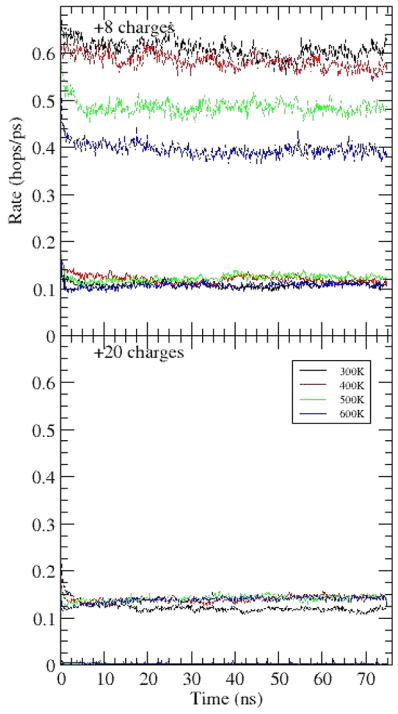


Figure 1. A plot of the average hopping rate (hops/ps) as a function of time for the +20 (bottom) and +8 (top) charge states. The temperatures 300, 400, 500, and 600 K are shown as black, red, green, and blue lines, respectively. The solid and dashed lines represent $\alpha = 1$ and $\alpha = 0.023$, respectively.

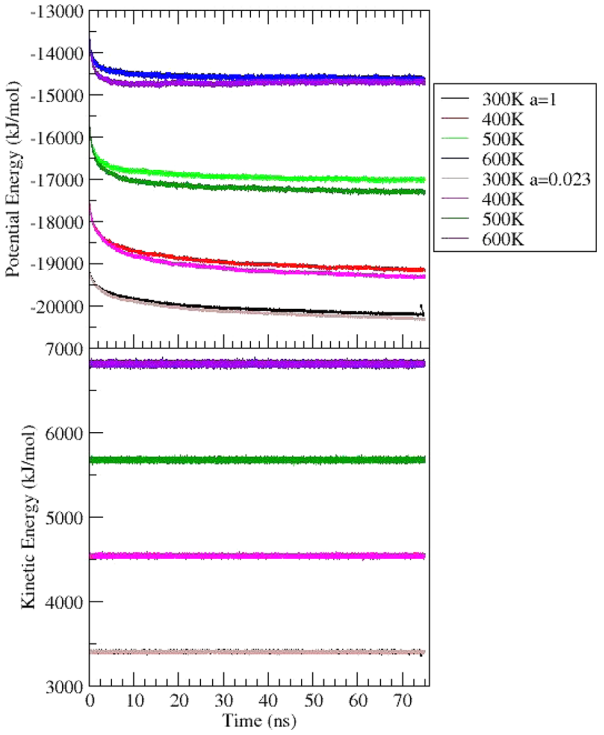


Figure 2. The potential and kinetic energies averaged over all 100 trajectories as a function of time for the +20 charge state at different temperatures. The temperatures 300, 400, 500, and 600 K are shown as black (brown), red (magenta), green (dark green), and blue (purple) lines for $\alpha = 1$ ($\alpha = 0.023$), respectively.

This implies the system is moving among multiple charge distributions which are quite close to the lowest energy. This is consistent with predictions from simulations with static charge which imply that protein relaxation can stabilize many different charge configurations within a small band of lowest energy.[11] More hopping would be expected at higher temperatures, and the hopping rate does show a slight temperature dependence.

The hopping algorithm used here is a simple approximation to the real proton motion. Protein self-solvation and any other local environment effects are not included. The use of $r_{cutoff} = 0.57$ means that it is modeling direct proton transfer, and any other possible mechanism is not considered. In effect, this α parameter is allowing us to tune the hopping rate without using a large increase in temperature which would change the dynamics of the whole system. Decreasing α from 1 to 0.023 increases the hopping rate for the high charge state from near zero to the same order of magnitude as the hopping rate reported for an excess proton in water.[25,38,39]

The potential and kinetic energies averaged over all 100 trajectories are shown in Figure 2 for the high charge state. For each temperature, the kinetic energy is constant, as expected since a thermostat was used. The potential energy decreases at the beginning as the charges rearrange and then becomes constant. This change in potential energy is expected as the charges move to more favorable arrangements. The plateau occurs when the charges are sampling microstates with the lowest energy. The decrease in potential energy matches the higher hopping rate at short times, while the flattening out occurs at the same time as the hopping rate becomes constant. With the higher hopping rate, more states are accessed. This is because the increased hopping probability makes hops from arginine to lysine more frequently accepted. Thus each charge can undergo more movement. This is seen in Figure 2 where the lower α simulations reach a lower potential energy.

The average number of charges on each monomer chain is shown as a function of time in Figure 3. For the high charge state at 300 K, all the chains have an average of +5 charges (a symmetric charge distribution). At 600 K, chain A increases in average charge to +6, and the other chains decrease to +4.5 for the high charge state when $\alpha = 1$. Once the charge of chain A begins to increase, it does not return to lower charge values. This asymmetry in charge starts quickly (within the first 2 ns) and monotonically increases with the greatest changes occurring at small times. Intermediate temperatures show an increasing asymmetry in the charge configuration. The charge accumulation on chain A is larger (+9.5 compared to +6) when α is decreased to 0.023 from 1 (see Figure 3) for the high charge state. This is accompanied by a complementary decrease in the number of charges on the other three chains. For the low charge state, the distributions are centered at $+2 \pm 1$ at both 300 and 600 K, and this does not change when α is changed. This change is on the same time scale as changes in secondary protein structure. These changes easily fall within experimental time scales in typical time-of-flight mass spectrometry studies (on the order of milliseconds). The fluctuations in charge for each chain are small compared to the average charge.

While Figure 3 shows the average (over 100 trajectories) number of charges on each monomer, Figure 4 shows the distribution of charges at 300 and 600 K for $\alpha = 0.023$. The last 100 configurations of each trajectory were counted for the charge distribution histograms. At 300 K, all the chains are equivalent and have distributions centered at +5 and +2 (for the high and low charge states respectively) with the majority of populations within ± 1 charge of the average. At 600 K, the high

dx.doi.org/10.1021/ct300906a | J. Chem. Theory Comput. 2013, 9, 2531−2539

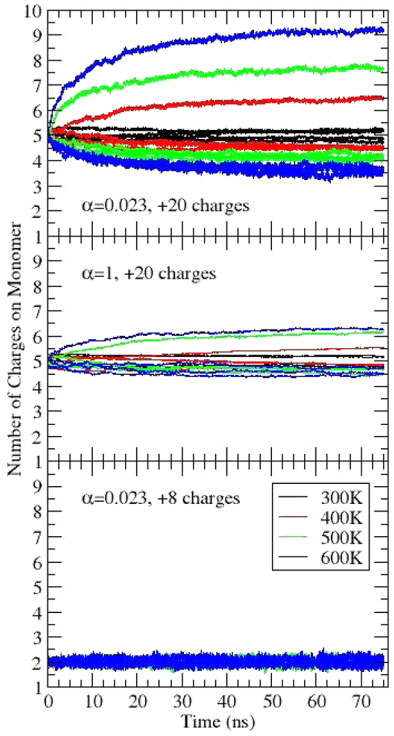


Figure 3. The average number of charges on each monomer is shown averaged over 100 trajectories as a function of time for different temperatures for the +20 charge state using $\alpha = 1$ (middle) and $\alpha = 0.023$ (top), and the +8 charge state with $\alpha = 0.023$ (bottom). The colors have the same meanings as in Figure 1 (there are four lines at each temperature, one for each chain).

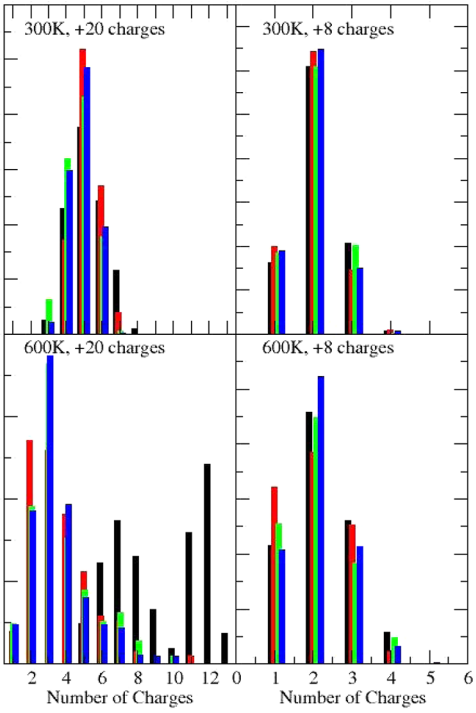


Figure 4. A plot of the charge distributions in the last 100 steps of the trajectories for the +20 charge state (left) and the +8 charge state (right) both using $\alpha = 0.023$ (the top panel is at 300 K and the bottom panel is at 600 K). Each chain is in a different color.

charge state distributions are broader overall having widths of about ±2 charges, and there is asymmetry. One chain has a distribution of charges centered near +10, while the other three chains have lower charges (centered near +3).

While these simulations used a +20 total charge state, the highest charge state as reported in experiments was +15. For a +15 TTR tetramer, CID experiments show fragments with an average charge state +8 ± 1 and trimers with the remaining charge.[4,33,34] The experiments do not have a well-defined temperature, but increasing collision energies are expected to correspond to higher internal temperatures. The 600 K simulation results and the experiments both show one monomer (out of the four) having half the total charge, but the 600 K simulations have a broader distribution of charge states than the experiments. This broader distribution could be caused by the simulation time being shorter than the experimental time. If some of the trajectories have unfolded and have the asymmetric charge while others are caught in the middle of this process, it would make the distribution wider than it would be after a longer time.

Structural Properties. Plots of the radius of gyration (averaged over 100 trajectories) of each chain are shown in Figure 5 as a function of time for different temperatures for the high charge state. Generally, the values stay relatively constant in time except for chain A, which grows with time. This growth becomes more pronounced as the temperature increases. Decreasing α increases the extent to which the radius of gyration increases. This large increase in radius of gyration

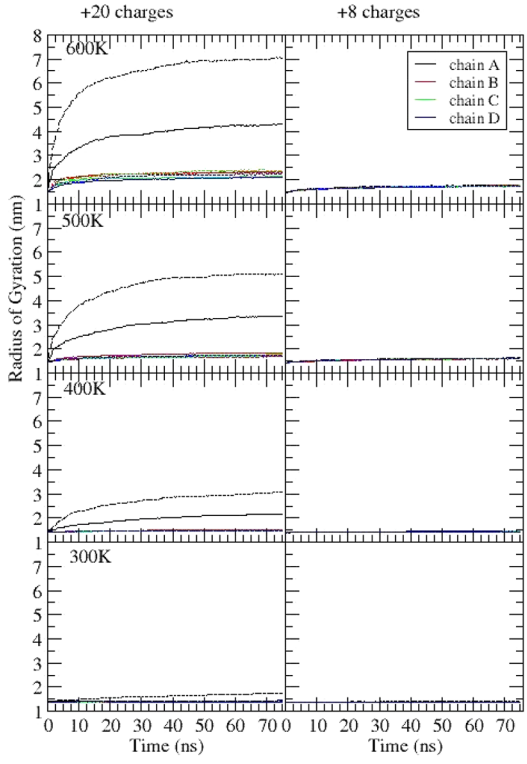


Figure 5. The radius of gyration averaged over 100 trajectories as a function of time for trajectories run at different temperatures for the +20 charge state (left) and the +8 charge state (right). The first (bottom), second, third, and fourth (top) panels show results at 300, 400, 500, and 600 K, respectively. Chains A, B, C, and D are represented by black, red, green, and blue lines, respectively. The solid and dashed lines correspond to $\alpha = 1$ and $\alpha = 0.023$, respectively.

dx.doi.org/10.1021/ct300906a | J. Chem. Theory Comput. 2013, 9, 2531−2539

indicates structural changes. Figure 6 shows snapshots of the protein complex at 300 and 600 K, demonstrating the

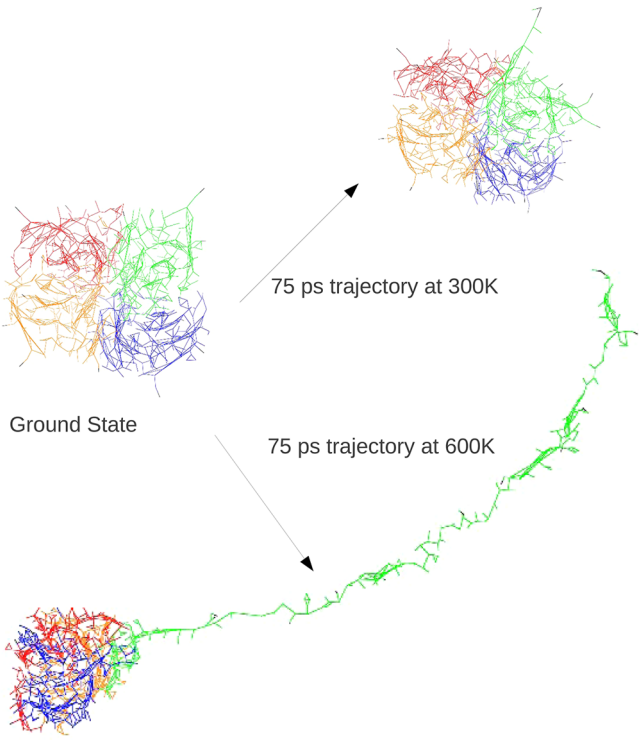


Figure 6. Snapshots of the ground state tetramer and the final structures at the end of two trajectories for the +20 charge state with α = 0.023 (one at 300 and the other at 600 K). Each chain is in a different color, and charged beads are represented in black.

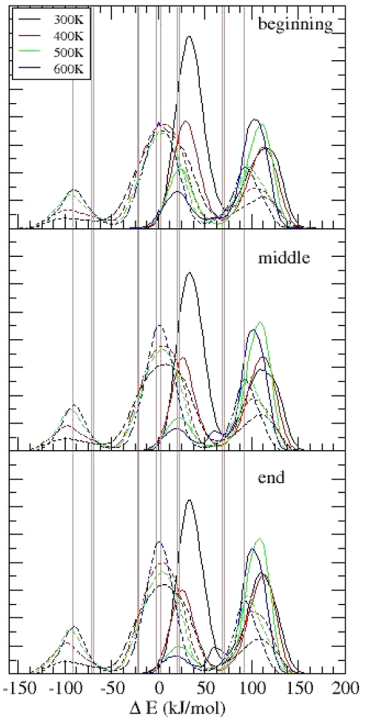


Figure 7. Distributions of ΔE for all attempted hops taken at times near the beginning, middle, and end of the trajectory propagation. The vertical lines show the differences in proton binding energies among all possible pairs of residue types, as calculated from the values in Table 1. The colors have the same meaning as in Figure 1. The top panel shows data collected from the first 25 ns, the middle panel the next 25 ns, and the last panel the remaining 25 ns of the trajectories. The solid and dashed lines represent results for α = 1 and α = 0.023, respectively.

unraveling of chain A. For the low charge state, the radius of gyration does not change for any of the chains at any temperature, indicating that there is no unfolding.

The root-mean-square deviation (RMSD) was also calculated and shows the same qualitative pattern as the radius of gyration with chain A increasing in size with increasing temperature. Quantitatively, the increase is larger with the higher hopping rate (lower α).

Energy Distributions and Pair Correlations. In order to examine the hopping algorithm in more detail, statistics related to the hopping process itself were examined. The first of these is shown in Figure 7 in which the values of ΔE of eq 1 are plotted as distributions for all attempted hops. Each trajectory was divided into three time domains. The beginning, middle, and end correspond to $t \leq 25$ ns, 25 ns $< t \leq 50$ ns, and $t > 50$ ns, respectively. All the instances with $\Delta E < 0$ represent successful hops independent of temperature. For $\Delta E > 0$, the temperature dependent hopping probability is determined from ΔE using eq 3, so only a fraction of hops will be successful in this case. The energy distributions have two peaks, one at approximately 30 kJ/mol and the other at about 100 kJ/mol for α = 1. Most of the ΔE values are positive, which is consistent with the low hopping rate. The lower α simulations have an additional peak in the ΔE distributions (see Figure 7) at approximately −100 kJ/mol, which was not seen when α was 1.

The vertical lines in Figure 7 represent the differences in $E_{binding}$ among all possible pairs of basic sites, as calculated from Table 1. Recall that ΔE has two contributions: (i) from the differences in $E_{binding}$ and (ii) from the differences in Coulomb energy resulting from moving a charge between two basic sites. If the difference in $E_{binding}$ dominated ΔE, one would expect to see values of ΔE peaked near the vertical lines in Figure 7. The peaks are not located at the same values as the differences in binding energy (but they are the same order of magnitude), and they have a peak width of about 40 kJ/mol, indicating the importance of the Coulomb energy.

Plotting the energy distributions separated by the donor site type produces the graph seen in Figure 8. We further broke down the energy distributions by acceptor site type (data not shown). The arginine proton binding energy is approximately 90 kJ/mol lower than that of histidine or lysine. Thus, if the electrostatic environment was the same for all hops, one would expect that it would take 90 kJ/mol more energy to move a proton from an arginine than to move a proton from a lysine or histidine. This difference in binding energy is the source of the peaks seen in Figure 7. The hops with ΔE closer to 0 are from hopping pairs where the donor and acceptor have the same type (or lysine/histidine pairs), and the other peaks are when the donor and acceptor are different.

In the first few steps, the charges tend to hop to arginine. In the low hopping rate situation, those charges remain on the arginine so most of the attempted hops seen in Figure 7 are unsuccessful hops with an arginine donor. The high charge state has more charges than there are arginines, so there are still some charges on lysines and histidines which can hop to other lysines or histidines. In the high hopping rate situation, the charges do hop from arginine to other basic sites; thus there are

dx.doi.org/10.1021/ct300906a | J. Chem. Theory Comput. 2013, 9, 2531−2539

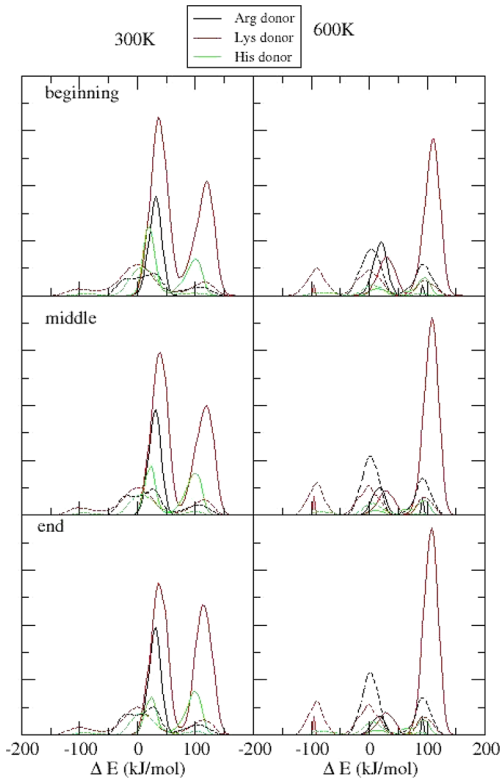


Figure 8. Energy distributions for all the hop attempts of the +20 charge state separated by the type of the donor site (arginine, lysine, and histidine donors are shown in black, red, and green, respectively). The left column is at 300 K, and the right column is at 600 K. Each trajectory was split into three time regions as in Figure 7. The solid and dashed lines represent results for $\alpha = 1$ and $\alpha = 0.023$, respectively.

more attempted hops with nonarginine donors (and more with arginine acceptors).

The discussion above relates to the high charge state ΔE distributions. The low charge state ΔE distributions have peaks at the same ΔE values but with different peak heights.

The hopping pairs graph shown in Figure 9 plots the bead number of the donor site on the x axis and of the corresponding acceptor site on the y axis for all successful hops. Data are shown for 300 and 600 K. Overall, Figure 9 shows the correlations between donor and acceptor sites counted over all successful hops in the 100 trajectories. The number of hops in all 100 trajectories between any pair of sites is indicated by the color: black is 1−999, red 1000−9999, and green 10 000−49 999. Chains A, B, C, and D are represented by bead numbers 1−254, 255−508, 509−762, and 763−1014, respectively. Most of the hops are between residues in the same monomer. Hops between different monomers do occur, but they are less frequent. Many of the hops are between pairs that are close together in the sequence. These appear as points near the diagonal. The plot is approximately symmetric about the line $y = x$. This means that if a hop from residue A to residue B occurs, then the hop from residue B to residue A will often occur with a similar frequency (not necessarily exactly the same). This seems counterintuitive since according to the values of $E_{binding}$ hops should be preferred when moving charge to more basic sites and not vice versa. However, hops between sites of the same type would be equally likely no matter which one was the donor.

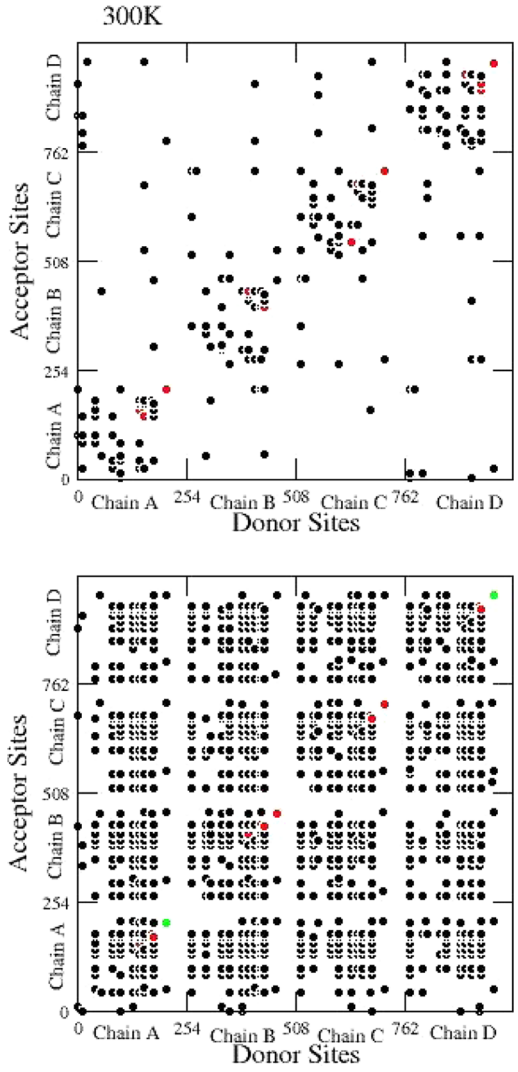


Figure 9. Hopping pairs graph. For each successful hop in the +20 charge state, the bead number of the donor site is plotted on the x axis with the bead number of the corresponding acceptor site on the y axis. Results are shown for 300 K (top) and 600 K (bottom) both using $\alpha = 1$. The number of hops between any pair of beads is indicated by the color: black is 1−999, red 1000−9999, green 10 000−49 999, and blue more than 50 000.

The number of pairs for which hops occur is much larger at 600 K than at 300 K. In particular, there are many more hops between different chains. At both temperatures, the pairs with the most frequent hops are those between sites that are close together in the sequence (near the diagonal).

All the possible basic sites have a positive charge at some point in the 100 trajectories. Not all sites are visited with the same frequency. The TTR tetramer has 16 arginine residues, 16 histidine residues, and 24 lysine residues. Arginine is the most basic of the four types of basic site considered here, and the charges spend significantly more time on arginine than on any other residue. Lysine is the least basic, but it is the most common so lysine has a larger percentage of the charges than histidine. The low occupation of the N-terminus is also explained by the number of residues of each type. As the temperature increases, the percent occupation for arginine increases significantly (from 57% to 77%), and the occupation of the N-terminus increases slightly while the others decrease.

dx.doi.org/10.1021/ct300906a | J. Chem. Theory Comput. 2013, 9, 2531−2539

CONCLUSIONS

Charge motion was studied in the +20 and +8 charge states of the TTR tetramer. Overall, the comparison of the high and low charge states shows that the Coulomb repulsion among the charges is a major contributor to unraveling a monomer in the complex, thereby producing a state ready for dissociation. As seen in Figure 5, monomer unfolding does not occur for the low charge (+8) states, regardless of the temperature or the charge hopping rate (that is, the value of α). For the high charge (+20) state, charge enrichment of a single monomer and monomer unfolding can occur. The rate of these processes is very slow when the temperature and hopping rate are low, but they increase dramatically both with higher hopping rates and temperatures, as seen in Figures 3 and 5. In this sense, the temperature and value of α act to decrease the time scale for these events but do not change their qualitative behavior.

This also speaks to the coupling between charge migration and monomer unfolding. With charge migration turned off, no monomer unfolding occurs regardless of the temperatures studied. Thus, it is the combination of charge migration, needed for charge enrichment, and energy activation, accomplished by increasing temperature, that leads to monomer unfolding. This is consistent with experimental studies.

Because the Coulomb repulsion among the charges is the dominant factor for monomer unfolding, it is expected that particular microscopic details of the force field or protein structure are not critical for determining the qualitative behavior of the monomer unfolding process. Rather it is a result of larger scale variations in the charge distribution. This is a key assumption in these calculations and in the interpretation of the results. For example, the MARTINI force field provides strong bonds to maintain secondary structures, such as helices, so the particular way a monomer unfolds using the MARTINI force field will differ from that for an all-atom force field. In effect the MARTINI force field has a higher barrier for unfolding than typical all-atom force fields. However, the protein is still able to unravel. The more flexible coils between the strong secondary structures allow movement which contributes to the ability of the protein to reach an extended structure. This is show graphically in Figure 6 where it can be seen that a monomer can extend quite significantly, even though there are pockets of structure still locked in by the strong secondary structure bonds. These pockets do not change the fact that charge is associating with the unfolding monomer, and that this charge is being moved farther from the remaining trimer complex. The drop in Coulomb repulsion still occurs as a result. For this reason, the qualitative sequence of events leading to monomer unfolding should be reasonable using the MARTINI force field. In the simulations, the MARTINI bead types were not changed when the charge on a bead was changed. Again, the small change in potential parameters associated with the fitting for these different charge states is not expected to affect the qualitative behavior of the unfolding process.

Overall, the results of this study show that for the purposes of understanding charge motions in large, charged, gas-phase protein complexes, a detailed description of proton transfer is not necessary since processes are dominated by Coulomb repulsion. This is consistent with experimental observations, and with previous static charge simulations from which the Coulomb repulsion model[11] was developed. For highly charged complexes, as studied here, total charge is the main physical parameter governing the process. The method used here is general enough to be applied to simulations of different protein complexes. Since the total charge and the associated Coulomb repulsion govern the process rather than the detailed protein structure, the general trends will be relevant for many protein complexes (although the quantitative details would differ).

For future work, we wish to learn more about the unfolding, in particular the key factors responsible for its initiation and propagation. One of the questions we are most interested in is the role of the N-termini. For example, does the N-terminus need to be charged in order for the monomer to unravel, and in what way does the charge state of the N-terminus effect the distribution of charge on the rest of the monomer. This knowledge should allow some degree of control over the process, which is the larger goal of this research.

AUTHOR INFORMATION

Corresponding Author

*E-mail: thachuk@chem.ubc.ca.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada. All computations were performed using WestGrid computing resources, which are funded in part by the Canada Foundation for Innovation, Alberta Innovation and Science, BC Advanced Education, and the participating research institutions (www.westgrid.ca).

REFERENCES

(1) Jurchen, J. C.; Williams, E. R. *J. Am. Chem. Soc.* **2003**, *125*, 2817−2826.
(2) Felitsyn, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**, *73*, 4647−4661.
(3) Benesch, J. L. P.; Sobott, F.; Robinson, C. V. *Anal. Chem.* **2003**, *75*, 2208−2214.
(4) Sobott, F.; McCammon, M. G.; Robinson, C. V. *Int. J. Mass Spectrom.* **2003**, *230*, 193−200.
(5) Wysocki, V. H.; Jones, C. M.; Galhena, A. S.; Blackwell, A. E. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 903−913.
(6) Jones, C. M.; Beardsley, R. L.; Galhena, A. S.; Dagan, S.; Cheng, G.; Wysocki, V. H. *J. Am. Chem. Soc.* **2006**, *128*, 15044−15045.
(7) Beardsley, R. L.; Jones, C. M.; Galhena, A. S.; Wysocki, V. H. *Anal. Chem.* **2009**, *81*, 1347−1356.
(8) Csiszar, S.; Thachuk, M. *Can. J. Chem.* **2004**, *82*, 1736−1744.
(9) Sinelnikov, I.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 617−631.
(10) Sciuto, S. V.; Liu, J.; Konermann, L. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1679−1689.
(11) Wanasundara, S. N.; Thachuk, M. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 2242−2253.
(12) Wanasundara, S. N.; Thachuk, M. *J. Phys. Chem. A* **2009**, *113*, 3814−3821.
(13) Wanasundara, S. N.; Thachuk, M. *J. Phys. Chem. B* **2010**, *114*, 11646−11653.
(14) Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 417−431.
(15) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399−1406.
(16) Paizs, B.; Csonka, I. P.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 637−650.

(17) Csonka, I. P.; Paizs, B.; Lendvay, G.; Suhai, S. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1457−1472.
(18) Kulhanek, P.; Schlag, E. W.; Koca, J. *J. Am. Chem. Soc.* **2003**, *125*, 13678−13679.
(19) Jorgensen, T. J. D.; Gardsvoll, H.; Ploug, M.; Roepstorff, P. *J. Am. Chem. Soc.* **2005**, *127*, 2785−2793.
(20) MacDonald, B. I.; Thachuk, M. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 2946−2954.
(21) Boyd, R.; Somogyi, A. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1275−1278.
(22) Voth, G. A. *Acc. Chem. Res.* **2006**, *39*, 143−150.
(23) Marx, D. *Chem. Phys. Chem.* **2006**, *7*, 1848−1870.
(24) Donnini, S.; Tegeler, F.; Groenhof, G.; Grubmuller, H. *J. Chem. Theory Comput.* **2011**, *7*, 1962−1978.
(25) Lill, M. A.; Helms, V. *J. Chem. Phys.* **2001**, *115*, 7993−8005.
(26) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.
(27) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.
(28) Fegan, S. K.; Thachuk, M. *J. Chem. Theory Comput.* **2012**, *8*, 1304−1313.
(29) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(30) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(31) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(32) Ruotolo, B. T.; Hyung, S.-J.; Robinson, P. M.; Giles, K.; Bateman, R. H.; Robinson, C. V. *Angew. Chem.* **2007**, *119*, 8147−8150.
(33) Hyung, S.-J.; Robinson, C. V.; Ruotolo, B. T. *Chem. Biol.* **2009**, *16*, 382−390.
(34) Pagel, K.; Hyung, S.-J.; Ruotolo, B. T.; Robinson, C. V. *Anal. Chem.* **2010**, *82*, 5363−5372.
(35) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: San Diego, CA, 2002.
(36) Palmieri, L. d. C.; Lima, L. M. T. R.; Freire, J. B. B.; Bleicher, L.; Polikarpov, I.; Almeida, F. C. L.; Foguel, D. *J. Biol. Chem.* **2010**, *285*, 31731−31741.
(37) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511−519.
(38) Schmitt, U. W.; Voth, G. A. *J. Phys. Chem. B* **1998**, *102*, 5547−5551.
(39) Vuilleumier, R.; Borgis, D. *Chem. Phys. Lett.* **1998**, *284*, 71−77.