Aqueous Guanidinium-Carbonate Interactions by Molecular Dynamics and Neutron Scattering: Relevance to Ion-Protein Interactions

Mario Vazdar,^{†,‡} Pavel Jungwirth,[†] Philip E. Mason[†]*

[†]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610 Prague 6, Czech Republic

‡ Rudjer Bošković Institute, Division of Organic Chemistry and Biochemistry, P.O.B. 180, HR-10002 Zagreb, Croatia

Keywords: Guanidinium, ion, force-field, aqueous, molecular dynamics, neutron scattering, protein.

ABSTRACT: Guanidinium carbonate was used in this study as a simple proxy for the biologically relevant argininecarbonate interactions in water. Molecular dynamics (MD) simulations of guanidinium carbonate were performed with non-polarizible water using two implementations of the ion force fields. In the first the ions had full charges, while in the second the ions had reduced charges in order to effectively account for electronic polarization effects of water. The results from the simulations were then compared to data from previous neutron scattering experiments. It was found that there were significant discrepancies between the full charge force field MD simulations and the experimental results, due to excessive ion pairing and clustering in the former. In contrast, reducing the ionic charges yields a more regular solution with a simulated structure which fits well the experimental data.

Introduction

Arginine is one of the more fascinating amino acids. With its sidechain essentially equating to a propyl group terminated with a cationic guanidinium ion, it is the third largest (by mass) of all the natural amino acids (second only to tyrosine and tryptophan). The peculiar behavior of arginine, and of the isolated guanidium ion have generated a lot of interest in the literature.¹⁻³ Notably guanidinium ions in aqueous solution are found to associate with each other, despite both ions bearing a full positive charge.4-9 It has been suggested that this homo-ion pairing of the arginine side chains is actually a significant factor both in protein folding and in protein aggregation.^{10, 11} Further, its interactions with anionic species are thought to be similarly important. For example arginine is known to make salt bridges with the side chains of glutamate and aspartate.^{12, 13} Arginine also interacts directly with small oxyanions such as the di-anionic carbonate ion. It is typically found double hydrogen bonded to the carbonate ion as in the example of the highly conserved binding site of many iron transport proteins¹⁴. Of these by far the largest and most important family is that of the lactoferrins,¹⁵ where the direct interaction of the arginine with the carbonate ion (chelated to an iron ion) is thought to play a key role in regulating the function of the protein.¹⁶

A previous MD study¹⁷ suggested that in solutions of guanidinium carbonate the Gdm⁺-CO₃²⁻ double hydrogen bond type interaction of the two ions was relatively

strong. Indeed in this study it was found to be strong enough to lead to nanosecond timescale ion pairs and resulted in aggregated ion structures on the nanometers lengthscale.¹⁷ The comparison of the molecular dynamics predications to the neutron scattering data seemed to provide some support for this conclusion.¹⁷

Here we re-examine this study in the light of new work which suggests that in non-polarizable simulations containing fully charged ions in a solution with simple point charge water the ion-ion interactions are significantly over-emphasized.¹⁸ Calculations have suggested that in order to more accurately represent ions in simple point charge water simulations the charge on the ions needs to be reduced by 25 % (*vide infra*) to account for the electronic part of the polarizability of the aqueous solution. A similar 'charge scaling' is commonly used in nonpolarizabile MD simulations of ionic liquids.¹⁹ This reduced charge model is called the electronic continuum correction or ECC model throughout this study.

It would be of a considerable value to determine if such factors are relevant in accurately modeling ion-protein interactions. One way to do this is via the comparison of predictions from MD simulations with direct structural experimental measurements. In this study guanidinium carbonate was chosen as a test case to examine this point of contest. As mentioned above, guanidinium is a proxy for the side chain of the amino acid arginine, and carbonate is an important biological ligand. This study examines ion-pairing in aqueous solutions of guanidinium carbonate via MD simulations with either full charges on ions or with scaled charges within the ECC model and then compares the results to previous neutron scattering measurements on the same system.¹⁷ We show that the ECC correction, which leads to a significant reduction in ion pairing and clustering, is crucial for obtaining a quantitative agreement with experiment.

Simulations Details

Two independent 50 ns MD simulations (after 1 ns of equilibration) were run from identical starting coordinates. Both simulations contained 48 Gdm⁺ ions, 24 CO₃² ions and 889 water molecules. In the first simulation, the ions bear full formal charge, i. e. charge is +1 for the Gdm⁺ ion (atom charges: C, 0.64; N, -0.80; H, 0.46) and -2 for the CO_3^{2-} ion (atom charges: C,0.676, and O, -0.892), respectively, as previously used.¹⁷ In the second simulation, electronic polarization effects were accounted for in an effective way by introducing ECC which is practically realized by rescaling all ionic charges by $1/\sqrt{\varepsilon_{el}}$, where ε_{el} = 1.78 is the electronic part of the static relative permittivity of water (i. e., the scaling factor equals to 0.75).^{18, 20, 21} In this way, Gdm⁺ ion has the charge of 0.75 (atom charges: C, 0.48; N, -0.60; H, 0.345) and the charge is -1.5 for the CO_3^{2-} ion (atom charges: C, 0.507, and O, -0.669), respectively. Water molecules were described using the SPC/E model.22

Periodic boundary conditions were used with long range electrostatic interactions beyond the nonbonded cutoff of 10 Å accounted for using the particle-mesh Ewald procedure.²³ The Berendsen thermostat with temperature of 300 K was used²⁴ and the box length was rescaled to 31.7840 Å - this yielded the correct physical number density (0.101 atoms Å⁻³). The SHAKE algorithm²⁵ was employed to constrain all bonds containing hydrogen atoms. The time step used in all simulation was set to 1 fs. Molecular dynamics simulations were performed with AMBER 11.²⁶

These simulations were then used to calculate the function measured in the previous first order ^{nat}N/¹⁵N neutron diffraction with isotopic substitution (NDIS) experiment.¹⁷ Specifically the reciprocal space function:

$${}^{n}\Delta_{N}(Q)_{\text{null}} = \frac{4.46 S_{NO}(Q) + 0.91 S_{NN}(Q) + 0.38 S_{NC}(Q)}{4.46 + 0.91 + 0.38} - 1$$
(1)

and the real space version of this function,

$${}^{n}G_{N}(r)_{null} = \frac{4.46g_{NO}(r) + 0.91g_{NN}(g) + 0.38g_{NC}(r)}{4.46 + 0.91 + 0.38} - 1$$
(2)

where the superscript n denotes that these measurements were made on a partially deuterated solutions in which the average scattering length of the hydrogen/deuterium mixture was zero. This effectively renders the exchangeable hydrogens 'neutron invisible' in these solutions. There are two chemical types of carbon in these solutions (carbon in carbonate and carbon in guanidinium) and two types of oxygen (oxygen in water and oxygen in carbonate). While these functions can be fur-

ther resolved, in the current case it is only useful to do so with the oxygens, of which 7.5 % are on the carbonate (O_{carb}) and 92.5 % are on the water (O_w) .

Results from Molecular Dynamics

Representative snapshots from the full charges and ECC simulation are shown in Figure 1. From these trajectories density maps were calculated for the distribution of carbonate around guanidinium (Figure 2). The radial distribution functions of the relevant atoms around the nitrogen atoms in these systems were also calculated. These were then weighted by the neutron scattering prefactors shown in equation 2 (Figure 3) and summed to emulate the experimental neutron scattering result that would be obtained from these systems. To further characterize these systems the coordination number of the oxygens (both those on carbonate and those on water) around each Gdm⁺ ion was calculated (Table 1).

Table 1: The coordination number of oxygen on water (O_w) and oxygen on carbonate (O_{carb}) within 3.5 Å of an atom on guanidinium

	Full Charges	ECC
Coordination number O _{carb}	4.8	2.3
Coordination number O _w	3.7	6.4
Coordination number O _{total}	8.5	8.7

As with previous studies of aqueous solutions of K_2CO_3 ,²⁷ the change in the ion-ion behavior upon the implementation of the ECC model (versus the full charges model) is dramatic (Figure 1).



Figure 1: Snapshot of the two $1.5m \text{ Gdm}_2\text{CO}_3$ solutions simulated in this study. Left contains the full charges and has the ions almost entirely clustered, while right has the electronic continuum correction model and has far less ion pairing. For clarity, water molecules are not shown.

In the full charges simulation the ions are practically completely ion paired into a single cluster with a lifetime in the 10-100s of nanoseconds (which is very similar to previous MD studies of Gdm_2CO_3)¹⁷, while in the ECC simulation the fraction of ion pairing is significantly lower. This observation is further confirmed by the examination of the number of oxygens within 3.5 Å of any N on the guanidinium ion (Table 1). The total coordination numbers of oxygens for the full charge and ECC models are 8.5 and 8.7 respectively with the former having a slightly lower coordination number due to the penetration of other species into this solvation shell. In the full charges model 56 % of this is due to oxygens on the carbonate (with the rest coming from water), even though the oxygens on carbonate only constitute ~7.5 % of the oxygens in the system. In the ECC model only 26 % of the oxygens in the first solvation shell are from the oxygens on the oxyanion. This factor of two in the coordination number of O_{carb} has a significant effect on the bulk behavior of these ions. In the full charge simulation the ions are almost entirely clustered, whereas in the ECC simulation ion pairing and clustering is much weaker.

The density maps of carbonate oxygen around Gdm⁺ also reveal this behavior (Figure 2). Not only is the coordination far more intense with the full charges simulations, but also carbonate comes closer to Gdm⁺ than in the ECC simulation. When both density maps are put on the same plot (Figure 2), the more extended yellow cloud from the full charge simulation thus penetrates closer to the Gdm⁺ ion than the smaller red cloud from the ECC simulation.



Figure 2: The density map of carbonate oxygen around Gdm^+ . In both cases the contour level is the same (0.07 atoms Å⁻³, 31 times the bulk density of O_{carb}). The full charges simulation is shown in yellow, while the electronic continuum correction model is shown in red. Clearly the full charges model has a higher level of ion pairing and also, the approach of the two ions is closer than in the ECC model.

Comparison of Molecular Dynamics with Neutron Scattering

The experimental measurement from the previous study¹⁷ constitutes the summation of three weighted pairwise correlations of the substituted nitrogens to the nitrogen, carbon and oxygen in the system (Figure 3).



Figure 3: Left: the components of the function ${}^{n}G_{N}(r)_{null}$ as calculated from the full charges simulation. Right: the same functions for the electronic continuum correction model. In both cases the color scheme is the same, with upper (blue) being the NN component of the function ${}^{n}G_{N}(r)_{null}$, red being the NO component, and green being the NC component. In the case of the NO component this is further broken down into the fraction of this due to the NO_{carb} (black) and NO_w (grey) component. Lower black is shown the total function as calculated from the MD simulations. Grey is shown the equivalent function from the experimental data.

As can be seen, the function calculated for each respective simulation reflects the different ion-ion behavior of each system. This is visible in all three of the components (NN, NO, and NC). Further it turns out to be a significant feature of these simulations that the molecular NN peak is slightly shifted from 2.25 for the full charges simulation to 2.30 for the ECC simulations. This change in position is mostly due to the stronger ion pairing in the full charges simulation distorting the molecular structure of the Gdm⁺ ion. Similarly the more intense ion pairing and clustering in the full charge simulation compared to the ECC simulation is clearly visible in the carbonate oxygen component of the g_{NO}(r) term. This can be seen primarily in the relative heights of the peaks at ~2.9 Å in Figure 3.

To generate the experimental function from the simulation data, all three of these weighted pairwise distribution functions are summed. This can then be back transformed to give the Q-space version of the data²⁸ (the one that was actually measured at the diffractometer¹⁷). These are shown for both the full charges and ECC simulation in Figure 4, demonstrating the much better agreement with neutron scattering for the latter.



Figure 4: The upper set of graphs is the real space function $({}^{n}G_{N}(r)_{null})$, lower the Q-space functions $({}^{n}\Delta_{N}(Q)_{null})$. In each case the experimental data are shown in grey, while the functions calculated from the MD are shown in color (blue for full charges and red for the electronic continuum correction model). Left panel contains the full data range, middle panel has been set to the low r limit in the range o-1.5Å and right panel in the range o-2.7Å. Lower panels are the backtransforms of the upper panels data.

The general relationship between Q-space and *r*-space data is that sharp intramolecular correlations (which tend to be at shorter r) yield longer wavelength features in Qspace. This can most easily be demonstrated by deleting the first peak in the experimental and MD r -space data (setting the function to the low r limit over this data range, then back transforming both the MD and experimental data sets, middle panel in Figure 4). Direct comparisons between simulation and experiment are better performed in Q-space on several grounds. Firstly, if there is a very sharp molecular correlation type peak in the rspace data, it will require an extended Q-range to fully describe it. If, however, that Q-range is not available in the experimental data, the feature will appear broader in the experimental r-space data than in the MD r-space data. However, if the comparison is done in Q-space, the data should match fully for the entire Q-range. Secondly, many of the longer range structures that we are mostly interested in for these solutions occur at higher *r* and are broader features. This means that if the comparison is performed in r-space, these broad features will be difficult to discern over the 'ringing' noise which tends to hamper the interpretation of such data. If, however, the comparison is done in Q-space, then this longer range structure will manifest itself as a sharper 'high frequency' feature at lower Q (typically below $5 A^{-1}$ for most hydration type structures).

After the deletion of the first CN peak at ~1.4 Å in both MD and experimental data, the discrepancies in the comparison of the full charges with the experimental Q-space data become more evident (lower panels of Figure 4). It is

clear that the experimental lower Q data region (which mostly corresponds to broader higher *r* structures such as the hydration of the ion) is fitted much better with the ECC data than with the full charges data. On top of this, the fit of the data in the range 5-10 Å⁻¹ is also much better with the ECC data than with the full charges data. This can be shown to be due to the second molecular peak in the function ${}^{n}G_{N}(r)_{null}$. When the first intramolecular peak at 1.5 Å is deleted (the function is set to the low r limit in the range o-1.5 Å), and the data is backtransformed into Q-space, the discrepancy between the full charge simulation and the experimental data remains (middle panels of Figure 4). But when this process is repeated with the intramolecular guanidinium peak at 2.3 Å, the discrepancy in the Q-space data is no longer found. The distortion of the Gdm⁺ ion due to the strong ion pairing in the full charges simulation is less consistent with the experimental data than the practically undistorted Gdm⁺ from the ECC simulation.

Arguably the simplest way to see these small differences in the intramolecular structure of these two solutions (top left panel of Figure 4) is by comparing the relative heights of the two molecular correlations (the CN and NN correlations) of both the full charges and ECC simulations. Given that the integration of these two peaks must be the same for both simulations (due to the same molecular topology), the mere fact that the peaks are of different heights for the two simulations demonstrates that there must be different intramolecular configurations. Further, after all of the molecular correlations have been deleted (furthest right panel of Figure 4), when this data is backtransformed it is found that the fit of the ECC data to the experiment is again much better in the range o-5 Å⁻¹ compared to the fit of the full charges simulation. It is this improvement in the fitting of the lower Q data in all three comparisons that is the most robust indicator that the ECC model is performing significantly better in emulating the structure of the experimental aqueous guanidinium carbonate solution than the full charges model. This also suggests that in full charge, non-polarizable simulations the ion-protein interactions, such as those found between the guanidinium side-chains of arginine and the carbonate ions in proteins such as lactoferrins, may be overestimated.

Conclusions

Simulations were performed for aqueous Gdm₂CO₃ using full charges, as well as scaled charges within the ECC approach which implicitly accounts for electronic polarization effects. The predictions of the structure from the simulations were then compared to previous neutrons scattering measurements of the same system.¹⁷ This comparison suggests that the ion-ion interactions are significantly stronger in the full charges simulations leading to a more extensive ion pairing and clustering than in the ECC model. It also indicates that the ECC model is performing significantly better than the full charges model, emulating quantitatively the structure obtained via the experimental measurement. This result is also suggestive for similar

ion-pairing structures found in proteins,^{3, 29, 30} which may be overemphasized in full charges simulations in nonpolarizable water and where ECC may represent a simple way how to improve the results.

A video summary of this paper is available at:

http://www.youtube.com/watch?v=8qi9ufOg97c

AUTHOR INFORMATION

Corresponding Author

• philip.mason@uochb.cas.cz (P. E. M.)

ACKNOWLEDGMENT

Support from the Czech Science Foundation (grant P208/12/G016) is gratefully acknowledged. P.J. thanks the Academy of Sciences for the Praemium Academie award.

REFERENCES

1. Puglisi, J. D.; Chen, L.; Frankel, A. D.; Williamson, J. R., Role of rna structure in arginine recognition of tar rna. Proceedings of the National Academy of Sciences of the United States of America 1993, 90, 3680-3684.

2. Rothbard, J. B.; Kreider, E.; Vandeusen, C. L.; Wright, L.; Wylie, B. L.; Wender, P. A., Arginine-rich molecular transporters for drug delivery: Role of backbone spacing in cellular uptake. Journal of Medicinal Chemistry 2002, 45, 3612-3618.

3. Schneider, C. P.; Shukla, D.; Trout, B. L., Arginine and the Hofmeister Series: The Role of Ion-Ion Interactions in Protein Aggregation Suppression. Journal of Physical Chemistry B 2011, 115, 7447-7458.

4. Boudon, S.; Wipff, G.; Maigret, B., Monte-carlo simulations on the like-charged guanidinium guanidinium ion-pair in water. Journal of Physical Chemistry 1990, 94, 6056-6061.

5. Soetens, J. C.; Millot, C.; Chipot, C.; Jansen, G.; Angyan, J. G.; Maigret, B., Effect of polarizability on the potential of mean force of two cations. The guanidinium-guanidinium ion pair in water. Journal of Physical Chemistry B 1997, 101, 10910-10917.

6. Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M. L.; Dempsey, C. E.; MacKerell, A. D.; Brady, J. W., The structure of aqueous guanidinium chloride solutions. Journal of the American Chemical Society 2004, 126, 11462-11470.

7. Vondrasek, J.; Mason, P. E.; Heyda, J.; Collins, K. D.; Jungwirth, P., The Molecular Origin of Like-Charge Arginine-Arginine Pairing in Water. Journal of Physical Chemistry B 2009, 113, 9041-9045.

8. Vazdar, M.; Vymetal, J.; Heyda, J.; Vondrasek, J.; Jungwirth, P., Like-Charge Guanidinium Pairing from Molecular Dynamics and Ab Initio Calculations. Journal of Physical Chemistry A 2011, 115, 11193-11201.

9. Vazdar, M.; Uhlig, F.; Jungwirth, P., Like-Charge Ion Pairing in Water: An Ab Initio Molecular Dynamics Study of Aqueous Guanidinium Cations. The Journal of Physical Chemistry Letters 2012, 3, 2021-2024.

10. Neves, M. A. C.; Yeager, M.; Abagyan, R., Unusual Arginine Formations in Protein Function and Assembly: Rings, Strings, and Stacks. Journal of Physical Chemistry B 2012, 116, 7006-7013.

1. Pednekar, D.; Tendulkar, A.; Durani, S., Electrostaticsdefying interaction between arginine termini as a thermodynamic driving force in protein-protein interaction. Proteins-Structure Function and Bioinformatics 2009, 74, 155-163. 12. Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L., The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. Biochemistry 1993, 32, 9668-9676.

13. Olson, C. A.; Spek, E. J.; Shi, Z. S.; Vologodskii, A.; Kallenbach, N. R., Cooperative helix stabilization by complex Arg-Glu salt bridges. Proteins-Structure Function and Genetics 2001, 44, 123-132.

14. Baker, H. M.; Anderson, B. F.; Baker, E. N., Dealing with iron: Common structural principles in proteins that transport iron and heme. Proceedings of the National Academy of Sciences of the United States of America 2003, 100, 3579-3583.

15. Ward, P. P.; Paz, E.; Conneely, O. M., Multifunctional roles of lactoferrin: a critical overview. Cellular and Molecular Life Sciences 2005, 62, 2540-2548.

16. Adams, T. E.; Mason, A. B.; He, Q. Y.; Halbrooks, P. J.; Briggs, S. K.; Smith, V. C.; MacGillivray, R. T. A.; Everse, S. J., The position of arginine 124 controls the rate of iron release from the N-lobe of human serum transferrin - A structural study. Journal of Biological Chemistry 2003, 278, 6027-6033.

17. Mason, P. E.; Neilson, G. W.; Kline, S. R.; Dempsey, C. E.; Brady, J. W., Nanometer-scale ion aggregates in aqueous electrolyte solutions: Guanidinium carbonate. Journal of Physical Chemistry B 2006, 110, 13477-13483.

18. Leontyev, I. V.; Stuchebrukhov, A. A., Electronic Continuum Model for Molecular Dynamics Simulations of Biological Molecules. Journal of Chemical Theory and Computation 2010, 6, 1498-1508.

19. Schroder, C., Comparing reduced partial charge models with polarizable simulations of ionic liquids. Physical Chemistry Chemical Physics 2012, 14, 3089-3102.

20. Leontyev, I.; Stuchebrukhov, A., Accounting for electronic polarization in non-polarizable force fields. Physical Chemistry Chemical Physics 2011, 13, 2613-2626.

21. Leontyev, I. V.; Stuchebrukhov, A. A., Electronic Polarizability and the Effective Pair Potentials of Water. Journal of Chemical Theory and Computation 2010, 6, 3153-3161.

22. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The missing term in effective pair potentials. Journal of Physical Chemistry 1987, 91, 6269-6271.

23. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh ewald method. Journal of Chemical Physics 1995, 103, 8577-8593.

24. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular-dynamics with coupling to an external bath. Journal of Chemical Physics 1984, 81, 3684-3690.

25. Vangunsteren, W. F.; Berendsen, H. J. C., Algorithms for macromolecular dynamics and constraint dynamics. Molecular Physics 1977, 34, 1311-1327.

26. Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A., In AMBER 11.

27. Mason, P. E.; Wernersson, E.; Jungwirth, P., Accurate Description of Aqueous Carbonate Ions: An Effective Polarization Model Verified by Neutron Scattering. Journal of Physical Chemistry B 2012, 116, 8145-8153.

28. Mason, P. E.; Neilson, G. W.; Dempsey, C. E.; Brady, J. W., Neutron diffraction and simulation studies of CsNO3 and Cs2CO3 solutions. Journal of the American Chemical Society 2006, 128, 15136-15144.

29. Shukla, D.; Schneider, C. P.; Trout, B. L., Complex Interactions between Molecular Ions in Solution and Their Effect on Protein Stability. Journal of the American Chemical Society 2011, 133, 18713-18718.

30. McLain, S. E.; Soper, A. K.; Daidone, I.; Smith, J. C.; Watts, A., Charge-Based Interactions between Peptides Ob-

served as the Dominant Force for Association in Aqueous Solution. Angewandte Chemie-International Edition 2008, 47, 9059-9062.