Metal cation complexation with natural organic matter in aqueous solutions: Molecular dynamics simulations and potentials of mean force

E. Iskrenova-Tchoukova, Andrey G. Kalinichev, R. James Kirkpatrick

Revised version submitted: *Langmuir,* September 3, 2010

Metal Cation Complexation with Natural Organic Matter in Aqueous Solutions: Molecular Dynamics Simulations and Potentials of Mean Force

Eugenia Iskrenova-Tchoukova,[1] Andrey G. Kalinichev,[1,2,*,†] R. James Kirkpatrick[3]

[1] Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA,

[2] Department of Geological Sciences, Michigan State University, East Lansing, MI 48824, USA

[3] College of Natural Science, Michigan State University, East Lansing, MI 48824, USA

[†] Present address: Laboratoire SUBATECH, Ecole des Mines de Nantes, Université de Nantes, CNRS-IN2P3, 4 rue A. Kastler, BP 20722, F-44307 Nantes Cedex 03, France

[*] To whom correspondence should be addressed: kalinich@subatech.in2p3.fr

Abstract

Natural organic matter (NOM, or humic substance) has a known tendency to form colloidal aggregates in aqueous environments, with the composition and concentration of cationic species in solution, pH, temperature, and the composition of the NOM itself playing important roles. Strong interaction of carboxylic groups of NOM with dissolved metal cations is thought to be the leading chemical interaction in NOM supramolecular aggregation. Computational molecular dynamics (MD) study of the interactions of Na^+, Mg^{2+}, and Ca^{2+} with the carboxylic groups of a model NOM fragment and acetate anions in aqueous solutions provides new quantitative insight into the structure, energetics, and dynamics of the interactions of carboxylic groups with metal cations, their association and the effects of cations on the colloidal aggregation of NOM molecules. Potentials of mean force and the equilibrium constants describing overall ion association and the distribution of metal cations between contact ion pairs and solvent separated ions pairs were computed from free MD simulations and restrained umbrella sampling calculations. The results provide insight into the local structural environments of metal-carboxylate association and the dynamics of exchange among these sites. All three cations prefer contact ion pair to solvent separated ion pair coordination, and Na^+ and Ca^{2+} show a strong preference for bidentate contact ion pair formation. The average residence time of a Ca^{2+} ion in a contact ion pair with the carboxylic groups is of the order of 0.5 ns, whereas the corresponding residence time of a Na^+ ion is only between 0.02 and 0.05 ns. The average residence times of a Ca^{2+} ion in a bidentate coordinated contact ion pair vs. a monodentate coordinated contact ion pair are about 0.5 ns and about 0.08 ns, respectively. On the 10-ns time scale of our simulations, aggregation of the NOM molecules occurs in the presence of Ca^{2+} but not Na^+ or Mg^{2+}. These results agree with previous experimental observations and are explained

by both Ca^{2+} ion-bridging between NOM molecules and decreased repulsion between the NOM molecules due to the reduced net charge of the NOM-metal complexes. Simulations on a larger scale are needed to further explore the relative importance of the different aggregation mechanisms and the stability of NOM aggregates.

I. INTRODUCTION

Natural organic matter (NOM) is an important and chemically active "soft matter" component of soils, groundwater, and sediments that forms principally by weathering and decay of plant material.[1-3] It interacts with and changes the solubility and toxicity of trace elements in natural aqueous environments.[4,5] NOM can also easily adsorb onto mineral surfaces, forming an organic coating which significantly alters the electrostatic properties of the surface and strongly affects the transport and availability of toxic species.[6-11] The formation of NOM coatings is also a significant factor controlling the stability of viruses in aquatic systems[12] and the performance characteristics of inorganic and polymeric nanofiltration membranes.[13,14] Common cations such as Na^+, Mg^{2+}, and Ca^{2+} are thought to play an important role in the supramolecular aggregation of NOM molecules into colloidal particles,[2,3,15-20] the formation of NOM surface coatings,[10-13] and the aggregation of mineral particles.[21,22] These background cations also successfully compete for binding sites with toxic metals, thus effecting their mobility in the environment.[23,24] Knowledge of the molecular mechanisms of the selective interaction of metal ions with chemically active groups on NOM molecules is important for understanding a wide variety of industrial, environmental, and geochemical processes.

Despite the ubiquitous nature and importance of NOM, its complex chemical structure and composition are difficult to characterize and are poorly known at the fundamental molecular level. Whether NOM is chemically a true macromolecular entity or a supramolecular assemblage of smaller molecular fragments held together by relatively weak non-covalent interactions (primarily electrostatic and hydrogen-bonding) is still under discussion.[2,3,25-28] Recent experimental results strongly support the supramolecular view .[29-31]

Modeling approaches relying only on an averaged structural and compositional

representation of NOM are currently successful in accurately predicting the ion binding and aggregation capabilities of NOM, as well as its other functional characteristics.[5,32-36] However, important questions about the specific molecular mechanisms responsible for the physical-chemical properties and behavior of NOM can only be addressed adequately by molecular scale approaches. Computational molecular modeling is a powerful tool for studying molecular-scale phenomena in the chemical, biological, and materials sciences.[37] Despite the well recognized uncertainties of the NOM composition and structure, such methods have been successfully used over the last decade to investigate NOM in molecular-scale detail.[38-47]

The TNB (Temple-Northeastern-Birmingham) model of an NOM molecular fragment[38-40] used in our work provides a good structural and compositional analog of the Suwannee River NOM commonly used in experimental studies.[43-45] Carboxylic groups are the most important interaction sites for metal binding to these molecules, as they are for many other organic and bio-organic molecules. Experimental and theoretical studies of the hydration of metal ions[48-53] and carboxylic functional groups[54-57] show that carboxylic groups are sites of strong association of divalent ions with NOM[24,58-60] and that increasing the cation charge density increases the tendency to form a contact ion pair.[44,47]

Here we present a computational molecular dynamics (MD) study of the interaction of Na^+, Ca^{2+}, and Mg^{2+} with the carboxylic groups of a TNB NOM fragment and with acetate anion that provides an improved quantitative understanding of the structure and energetics of these interactions. This work is an extension of our previous research on metal-NOM interaction.[43-45] A combined experimental ^{133}Cs NMR and computational MD study of the Cs^+ complexation with Suwannee River NOM provided a detailed molecular-scale picture of these interactions and showed that Cs^+ associates only weakly with NOM via outer sphere complexation and that the

frequency of exchange between the NOM and the bulk solution is greater than 10^2 Hz.[43] A later study[44] extended this molecular modeling approach to Na^+, Mg^{2+}, and Ca^{2+}, and showed that Ca^{2+} forms the strongest inner-sphere complexes with NOM, Na^+ forms weak outer-sphere complexes, and that Mg^{2+} does not interact with NOM due to much stronger interaction with H_2O molecules in their first hydration shell. A combined experimental and molecular modeling study of fouling of polyethersulfone water filtration membranes caused by NOM[45] confirmed earlier observations[13] that Mg^{2+} and Na^+ should cause much less membrane fouling than Ca^{2+}. Those results suggested that divalent ions may cause fouling primarily by promoting the aggregation of NOM molecules in solution rather than by forming "ionic bridges" between negatively charged sulfonate functional groups of the membranes and the carboxylic groups of NOM. The work presented here further quantifies these observations by determining the potentials of mean force for the interactions of the metal ions and carboxylic groups, by relating the molecular-scale energetics and dynamics to the local structure, and by quantitatively investigating aggregation of the fragments. The results for acetate anion provide an important comparison, especially since small carboxylate species are often considered as simplified proxies of the larger NOM molecules in theoretical studies.[61,62]

The model systems presented here are almost an order of magnitude larger than in the previous studies, allowing us to approach the issues of supramolecular complexation quantitatively. The choice of cations explores the effects of the interplay of ion charge and size: Na^+ and Mg^{2+} are of approximately the same size, whereas Mg^{2+} and Ca^{2+} have the same charge. In addition to computing the radial distribution functions and potentials of mean force, we calculate the cation residence times in contact ion pair (CIP) and solvent separated ion pair (SSIP) configurations with the carboxylic groups of the simulated anions and estimate the ion

association and equilibrium constants for each ion pair.

II. COMPUTATIONAL METHODS AND DETAILS

NOM is a complex material with a wide range of structures and compositions.[2,3,26-28] The TNB model fragment used in our simulations contains three carboxylic groups, three carbonyl groups, two phenolic groups, two amine groups, and four other R-OH alcohol groups and has a total molecular weight of 753 Da.[38-40] The size, composition, molecular weight, degree of aromaticity, and total charge density of this model are in good agreement with available experimental characterization of NOM[26,27,29], the results of computer assisted 3-D structure elucidation,[42] and stochastic biogeochemical modeling.[33] The composition of the TNB model is also quite close to the composition of Suwannee River NOM, which is often used experimentally.[43-45] To model near-neutral pH conditions in our simulations, we consider the three carboxylic groups of the NOM fragment as completely deprotonated and the hydroxyl and amine groups as always protonated. Following our earlier work on the molecular modeling of metal-NOM interactions,[43-45] we focus on the quantitative characterization of site-specific interactions between metal cations and each of the three structurally distinct carboxylic groups of the TNB fragment (Figure 1). For comparison, similar aqueous solutions containing acetate anions instead of NOM were also simulated.

We used the TIP3P water model,[63] and all other interatomic interactions were modeled using the AMBER FF03 force field, which for the atom types considered here is identical to FF99.[64, 65] The partial charges carried by each atom of the NOM model molecule were derived with a utility in AMBER Tools[66] based on the AM1-BCC charge generation method.[67,68] Thus, the partial charges (in electron charge units, *e*) on each carboxylate carbon in NOM or acetate

were +0.3, and the charge on each of the carboxylate oxygens was –0.6. The two oxygen atoms of each $–COO^-$ group are arbitrarily labeled O1 and O2 (Figure 1). For comparison, the charges from the R.E.D. database[69] are +0.8 for carboxylate carbon and –0.6 and –0.7 for the two carboxylate oxygens. Since the the carboxylic groups in our simulations are considered always deprotonated, and the two oxygen atoms are statistically equivalent in aqueous solution, we opted for equal charges of –0.6 on the two oxygen atoms of the carboxylic groups. Placing different charges on the two oxygen atoms would create an artificial bias on the interaction of the metal cation with the carboxylic group which does not contribute to any deeper physical insight on the metal ion – NOM interaction of interest.

The model systems were prepared by combining eight anions (TNB or acetate) with the appropriate number of Na^+, Ca^{2+}, or Mg^{2+} cations to provide charge neutrality (Table 1). The neutral metal-NOM systems were solvated with 4215 water molecules and the metal-acetate systems with 2200 water molecules. Classical MD simulations were run using the AMBER 9.0 software package[66] in cubic simulation boxes (box size: approximately 41 Å for the smaller acetate systems and approximately 51 Å for NOM systems) under periodic boundary conditions using Ewald summation to account for the long-range electrostatic interactions.

All simulations were performed at ambient temperature (T = 300 K). The equilibration of the simulated systems consisted of three steps: a 100-ps MD run at constant temperature and a constant, relatively low density, followed by compression of the system in a 100-ps simulation run at a constant pressure of 100 MPa to achieve liquid water density, and an additional 100-ps run at a pressure of 0.1 MPa to ensure good structural relaxation of the system under ambient conditions. All the equilibrium MD runs were then performed at constant volume and temperature. Each production MD run lasted for a total simulation time of 10 ns. The Newtonian

equations of atomic motion were integrated with a time step of 1 fs. For all systems, the models were first built with randomized initial positions of the metal ions. For the systems with Mg^{2+}, we also performed a second set of 10-ns production runs with a different starting configuration in which all Mg^{2+} ions were initially placed at least 5 Å away from any of the carboxylic groups. The equilibrium MD trajectories resulting from all production runs were processed to extract the radial distribution functions (RDFs), potentials of mean force (PMFs), and other structural, dynamical and energetic properties of the simulated solutions.

One important goal of this study was to quantify the metal-carboxylate interactions by deriving the potentials of mean force. The PMF characterizes the change in the free energy of the system due to the changes in its configuration.[37,70,71,] The free energy of a system in thermodynamic equilibrium in an *NVT* statistical ensemble, constant number of particles (*N*), temperature (*T*), and volume (*V*) is defined as:[37]

$$F = -k_B T \ln Z , \tag{1}$$

where Z is the canonical partition function and k_B is the Boltzmann constant. For two interacting species in solution, the PMF, $W(r)$, is a function of the reaction coordinate or the separation distance r, and can be obtained by integrating the mean force between these two species along the reaction coordinate.[37,71,72] Thus, the PMF represents the free energy profile of the system as a function of the separation distance between the two species. It can be shown[71,72] that:

$$W(r) = -k_B T \ln g(r), \tag{2}$$

where $g(r)$ is the corresponding RDF for this pair of species with the standard normalization for large separation, $g(r) \rightarrow 1$ as $r \rightarrow \infty$. Thus, $W(r)$ asymptotically approaches zero for large separation distances. We will refer to the PMF calculated from eq. (2) as one derived from free (unrestrained) MD simulations.

However, this straightforward method does not perform well for systems with high potential barriers in their energy landscapes. A high potential barrier in the free energy profile corresponds to a low probability region in the phase space of the system. The free MD trajectories of the system will, naturally, avoid those regions leading to poor sampling. The sampling statistics can be improved by introducing specific restraints into the MD simulations using the so-called umbrella sampling algorithm.[72,73] After the simulations, the effect of the restraints on the final results can be removed with a variety of methods.[73,74] Trzesniak *et al.*[72] provides a recent, comparative review of the methods to compute PMFs.

In our umbrella sampling simulations, the distance r between the metal ion and the carbon atom of a carboxylic group was chosen as the reaction coordinate and was restrained by means of a harmonic potential:

$$V_{\text{harm}} = \frac{1}{2} k (r - r_c)^2 , \tag{3}$$

where k is the harmonic force constant and r_c is the center of the restraining potential. A series of restrained MD runs (umbrella windows) were performed with restraining potentials centered at different interatomic distances from 2 Å to 10 Å, using small steps of 0.25 Å or 0.1 Å, depending on the strength of the restraining force constant, as described in the next paragraph. In order to extract the PMF from the biased simulations, the umbrella integration (UI) method[75,76] was applied.

In setting up the umbrella sampling simulations, one has control over two parameters: the strength of the restraining harmonic potential and the spacing of the umbrella windows. The harmonic constant k should be large enough to allow for adequate sampling of the potential energy barrier region. Following the recommendations of Kästner and Thiel,[76] we set the restraining constants at 50 kcal/(mol·Å^2) for Ca^{2+} – acetate and Mg^{2+} – acetate and at 100 kcal/(mol·Å^2) for Ca^{2+} – TNB and Mg^{2+} – TNB umbrella sampling simulations. The spacing of the umbrella windows is to be determined according to the strength of the harmonic constant to ensure the adequate sampling of all distances. For k = 50 kcal/(mol·Å^2), we spaced the umbrella windows every 0.25 Å and for k = 100 kcal/(mol·Å^2) at 0.1 Å. The restrained umbrella sampling simulations were run for 550 ps for each umbrella window. A snapshot of the simulated system was output every 0.05 ps (every 50 simulation steps). The sizes of the simulated systems are shown in Table 1. The statistics of the umbrella sampling simulations were monitored by checking different segments of the MD trajectories for equilibration as well as by making sure that the frequency of the output was adequate. Consequently, the first 150 ps of each umbrella sampling simulation was discarded, and the umbrella integration results were calculated for a total of 400 ps. Umbrella sampling was not used for the systems with Na^+, because its lower charge allowed it to be mobile enough to efficiently probe the phase space of the system in the free MD runs. The Jacobian correction[77] was added to all PMFs derived from umbrella sampling simulations.

The computed radial distribution functions and potentials of mean force were then used to calculate bulk ion-pair association constants and the equilibrium constants describing the partitioning of metal ions between contact ion pair (CIP) and solvent-separated ion pair (SSIP) configurations. CIPs have direct contact between the metal ion and one or both oxygens of a

carboxylic group, whereas SSIPs have a single layer of water molecules separating the cation and the oxygens of the carboxylic group. CIPs correspond to the first peak in the radial distribution function, and SSIPs to the second. The position of the first minimum in the radial distribution function, R_L, is the threshold distance separating CIPs and SSIPs. For simplicity, we assume that the ions are fully dissociated at distances larger than the second minimum of the RDF, R_U. The equilibrium constant, K_{eq}, describing the distribution of the carboxyl-associated metal ions between the SSIPs and the CIPs can be calculated as the ratio of their densities:[78,79]

$$K_{eq} = \frac{\rho_{SSIP}}{\rho_{CIP}} = \frac{4\pi \int_{R_L}^{R_U} g(r)\, r^2 dr}{4\pi \int_{0}^{R_L} g(r)\, r^2 dr} = \frac{4\pi \int_{R_L}^{R_U} \exp[-W(r)/k_B T]\, r^2 dr}{4\pi \int_{0}^{R_L} \exp[-W(r)/k_B T]\, r^2 dr} \quad (4)$$

where, as described above, R_L and R_U are the lower and the upper distance bounds, respectively, for the solvent separated ion pair. Similarly, the ion-pair association constant, K_a, describing the distribution of metals between the bulk solution and all ion pairs is calculated as

$$K_a = 4\pi N_A \int_{0}^{R_U} \exp[-W(r)/k_B T]\, r^2 dr \quad (5)$$

where N_A is Avogadro's number and R_U is the upper bound distance for SSIPs.

To quantify aggregation of NOM molecules during the simulations, we calculated the radius of gyration[80] of pairs NOM molecules as implemented in AMBER Tools.[64] In this definition, for a system of N particles the radius of gyration is given by:

$$R_{gyr} = \frac{1}{N}\sqrt{\sum_i \left(\vec{r}_i - \vec{R}_0\right)^2}\ , \tag{6}$$

where the position of each particle is described by a vector $\vec{r}_i$, and $\vec{R}_0 = \frac{1}{N}\sum_i \vec{r}_i$.

III. RESULTS AND DISCUSSION

Metal cations can associate with carboxylic groups via different coordination geometries.[46,81-84] Figure 1 schematically illustrates the most common configurations that were observed in our simulations. In *bidentate* coordination, the metal cation M is coordinated simultaneously with the two oxygen atoms of the deprotonated carboxylic group (C29 in Figure 1). In this relatively stable inner sphere CIP coordination, the ion stays, on average, equidistant from the two oxygen atoms on the bisector plane orthogonal to the plane of the carboxylic group. In a *monodentate* CIP coordination, the metal ion is coordinated with only one of the oxygen atoms (oxygen O1 of C23 in Figure 1). A recent survey of the known metal ion – carboxylic group coordinations combined with quantum chemical calculations using density functional theory[82,83] showed that preference for monodentate or bidentate coordination depends on the ion charge and the local ligand environment.

1. Metal – carboxylate structure and dynamics from free MD simulations

The radial distribution functions of all systems modeled by free MD simulations show a complex local ion association structure involving contact and solvent separated ion pairs (Figure 2; Table 2). In all cases, the contact ion pair region contains two distinct maxima corresponding to bidentate and monodentate coordinations, CIP(b) and CIP(m). To clarify this

structure, we analyzed the angular distributions of the metal cations around the carboxylic group of the anion at selected interatomic distances. The angular coordinates are defined as follows (Figure 3): the origin of the coordinate system is at the position of the carbon atom, C, of the carboxylic group (see also Figure 1), $\vec{R}$ is the vector connecting C with the metal ion, M, and angle φ is measured in the plane defined by the three atoms of the carboxylic group. The angle φ is defined as the angle that the projection of $\vec{R}$ onto the plane of the carboxylic group makes with one of the bonds, arbitrarily labeled C–O1. Thus, the second oxygen atom, O2, is at $\varphi \approx 120°$ and the bisector of the (O1-C-O2) angle is at $\varphi \approx 60^{o}$. The azimuthal angle, θ, is measured from the plane of the carboxylic group and takes values between -90° and +90°, with $\theta = 0$ when the ion is located in the plane of the carboxylic group.

Figures 4(a-c) show the angular distributions of the orientation of the Ca^{2+}–C vector in the Ca^{2+}–TNB NOM system. The data in each plot are accumulated from all three carboxylic groups of the NOM molecule at three ranges of distances. At Ca^{2+}–C distances between 3.0 and 3.5 Å, corresponding to the first maximum in the RDF, only $\theta \approx 0 \pm 30°$ and $\varphi \approx 60°$ are observed, demonstrating that only bidentate coordination is stable at this distance (Figure 4a). The results also clearly show greater metal mobility perpendicular to the plane of the carboxylic group than in the plane. At larger Ca^{2+}–C separations between 3.5 and 4.0 Å, only monodentate coordination is observed with the angular distribution centered at $\theta \approx 0$ and $\varphi \approx 0$ or $\varphi \approx 120°$ (Figure 4b). The metal mobility in the plane of the carboxylic group is greater in monodentate coordination than in bidentate coordination. Finally, at even larger Ca^{2+}–C distances of 4.0 to 6.5 Å, corresponding to solvent separated ion pairs, there is much greater metal mobility than for the contact ion pairs and the orientation of the Ca^{2+}–C vector covers wide ranges of angular

space characteristic of both monodentate and bidentate coordination (Figure 4c). Thus, the global maxima of the radial distribution functions for the systems containing Na^+ and Ca^{2+} (Figure 2) correspond to CIP(b) coordination, and the secondary maxima correspond to CIP(m) coordination. For the systems containing Mg^{2+} ions, within the accuracy of our calculations, both bidentate and monodentate coordinations appear to be approximately equally probable with a slight preference for the monodentate coordination. These two maxima in the RDFs correspond to the global and secondary minima, respectively, of the potentials of mean force, as we discuss in more detail below.

All characteristic distances for each cation – carboxylic group pair in our simulations depend only on the nature of the cation and for a given cation are virtually the same for TNB and acetate. Therefore, we list only one value for each cation in Table 2. These values are in generally good agreement with similar structural parameters for metal cation binding to the carboxylic groups of other organic molecules.[81-86] For comparison, in Table 2 we show characteristic distances between the metal ion and the C atom of the carboxylic groups of 2-keto-3-deoxyoctanoate, galacturonate, and carbonate obtained from the recent quantum chemical (DFT level) calculations of Selvarengan et al.[46] Note, however, that our simulations reflect thermodynamic averages at a finite temperature of 300 K over many equilibrium structures obtained by classical force field-based MD simulations for bulk aqueous solutions. In contrast, the distances obtained by the quantum chemical calculations represent single optimized geometries (one for each species) of relatively small molecular clusters at 0 K.

The results of the free MD simulations provide information about the dynamical behavior of the metal cations through the characteristic times a cation spends in different coordinations to the carboxylic groups. The average lifetimes the metal cations in the three different coordination

coordinations – CIP(b), CIP(m), and SSIP – are estimated from the time autocorrelation functions (ACFs) of the configuration populations in a manner similar to the one usually employed in the studies of hydrogen bonding lifetimes.[48,87,88] For example, for CIP(m) coordination we define two population functions: $h(t)$ and $H(t; t^*)$ as follows:. $h(t) = 1$ if the metal ion is in a monodentated CIP coordination at time t, otherwise, $h(t) = 0$;. $H(t; t^*) = 1$ if the metal ion was in monodentate coordinated CIP at time $t = 0$, it still is in the same coordination with the same carboxyl carbon at time t, and meanwhile, it did not leave the coordination for times longer than t^*; otherwise $H(t; t^*) = 0$. The parameter t^* is usually taken to be 2 ps – a typical lifetime of a hydrogen bond in liquid water at ambient conditions,[88] which also reflects the rate of exchange of water molecules in the first hydration shell of metal ions.[48] For a metal ion to form a CIP with the carboxylic group, at least one water molecule must leave the first hydration shell to allow its place can be taken by one of the carboxyl oxygens. Therefore, we adopt the same value of $t^* = 2$ ps in our calculations.

The autocorrelation function of $h(t)$, averaged over all ion pairs, is a measure of the survival probability of an *intermittent* CIP(m). Analogously, the autocorrelation function of $H(t; t^*)$, averaged over all ion pairs, reflects the time-dependence of the survival probability of a *persistent* CIP(m). Figure 5 shows the evolution of the survival probabilities for Ca^{2+} and Na^+ computed from the free MD trajectories. The average lifetimes obtained from exponential fits of these survival probabilities are collected in Table 3. Results for Mg^{2+} are not shown because it underwent no transitions between different coordinations during the free MD simulations.

The differences in the dynamical behavior of the three cations studied here are also well illustrated by the time evolution of the metal–carboxylic carbon distances in NOM systems (Figure 6). Na^+ shows much higher mobility and rapid switching between different

coordinations, as indicated by its much shorter residence times. Note the shorter timescale for the Na^+–NOM system in Figure 6. Ca^{2+} ions also show relatively frequent transitions between bidentate and monodentate CIP coordination and CIP(m) residence times similar to those of Na^+. The transitions of Ca^{2+} between CIP and SSIP coordination, on the other hand, are much slower, which is reflected in the longer residence times. Again, Mg^{2+} shows no transitions between different coordinations.

Na^+ does not coordinate as strongly with carboxylic groups as Ca^{2+} or Mg^{2+}. For example, during the 10-ns simulation about half of the Na^+ ions in the simulated systems are never coordinated to an acetate molecule. The average coordination times in CIP(b) coordination were relatively short: 30 ps with acetate and 50 ps with NOM (Table 3). The differences in these values probably reflect not only the different nature of the anionic binding sites but also the statistical uncertainty of such estimates. The accuracy of the estimate of the average lifetime of Na^+ in a monodentate coordinated CIP is very low, because the monodentate CIP is not well defined for Na^+. The radial distribution functions in Figure 2 show that the maximum corresponding to monodentate coordination within the CIP is not well resolved for the systems involving Na^+, indicating a lower energy barrier between bidentate and monodentate CIP coordinations for this ion.

The RDFs for the Ca^{2+}–C pairs show distinct bidentate and monodentate maxima within the CIP coordination range. The average residence time of CIP(b) coordinated Ca^{2+} is about 530 ps with acetate and 450 ps with NOM. These values contrast with average lifetimes in CIP(m) coordination of the order of 10 ps and 80 ps, respectively, reflecting the tendency of Ca^{2+} to form long-lasting inner-sphere complexes with carboxylic groups in solution. The average Ca^{2+} residence times in SSIP coordination (outer-sphere complexing) are approximately 20 ps with

acetate and 200 ps with NOM.

The charge to radius ratio of Mg^{2+} is the largest of the three cations considered in this study, and it binds the most tightly to the carboxylic groups. The absence of transitions between coordination environments during the 10-ns equilibrium MD runs (Figure 6) is reflected in the distinct gap near 2.5 Å between bidentate and monodentate coordination represented by the first two RDF maxima of the Mg^{2+}–C pairs (middle row in Figure 2). This observation also indicates a likely source of statistical uncertainty due to insufficient sampling of the phase space of Mg^{2+}–C ion-pair in all simulations with Mg^{2+}. In the simulations with fully randomized initial Mg^{2+} positions, each Mg^{2+} that started out coordinated to a carboxylic group in CIP coordination stayed in that coordination with the same carboxylic group for the entire duration of the simulation. Therefore, we can only put a lower bound of 10 ns on the residence time of Mg^{2+} in CIP with carboxylate. The second set of 10-ns simulations with initial Mg^{2+}–C separations of at least 5 Å, corresponding approximately to SSIP coordination, showed no evidence of CIP formation for the entire duration of the simulations. The observation that no new ion pairs were formed indicates that the exchange of H_2O molecules and carboxyl oxygen atoms in the first solvation shell of Mg^{2+} is a very slow process characterized by high potential energy barriers. These results confirm previous findings that Mg^{2+} ions remain predominately hydrated in aqueous solutions for times longer than the 10-ns timescale of the present simulations and are consistent with the notion that the desolvation rate of Mg^{2+} ion is about three orders of magnitude slower than that for Ca^{2+} due to the greater surface charge density.[48,49,52,85,86] Such differences in the size and strength of the hydration shells of Na^+, Cs^+, Ca^{2+}, and Mg^{2+} (reflected in the corresponding metal-H_2O RDFs) and their effects on the metal-NOM complexation are also discussed in our earlier publication.[44]

2. Potentials of mean force, association and equilibrium constants for individual ion pairs

The structural and dynamical results above demonstrate the need for restrained sampling simulations of the interaction of carboxylic groups with Ca^{2+} and Mg^{2+} in aqueous systems, because obtaining the potentials of mean force from radial distribution functions derived from free MD simulations relies on adequate sampling of the configuration space of the system. The PMFs calculated from the RDFs for Na^{+} and by umbrella sampling for Ca^{2+} and Mg^{2+} using the methods described in Section II (Figure 7) are consistent with the computed RDFs and residence lifetimes. For each cation, the PMF curves are generally similar for acetate and NOM, and the depth of the potential minimum increases from Na^{+} to Ca^{2+} to Mg^{2+}. Thus, the specific features of the short-range interaction are controlled predominately by the metal cation.

For Mg^{2+}, the minimum of the PMF corresponding to monodentate CIP coordination is somewhat deeper than the one corresponding to bidentate CIP coordination with a very high (~10 kcal/mol) energy barrier between the two. This is consistent with previous results[46,82,83] that monodentate coordination is more stable for Mg^{2+} than bidentate coordination in aqueous solution. This is because the water molecules in the first hydration shell bind to it more strongly than to the other cations studied here. Thus, it is energetically more difficult for hydrated Mg^{2+} to lose more than one nearest neighbor water molecule to accommodate coordination with both oxygen atoms of the —COO^{-} group. For NOM molecules in aqueous solution, this means that Mg^{2+} does not easily associate with carboxylic groups. At the same time, should Mg^{2+}–NOM association occur, Mg^{2+} has difficulty leaving either of the CIP states or exchanging between monodentate and bidentate coordinations.

Water molecules are less tightly held to Ca^{2+} ions, because the charge-to-radius ratio of

Ca^{2+} is smaller than for Mg^{2+}. Thus, calcium ions have a better capacity for association and formation of relatively long-lived contact ion pairs with carboxylic groups in aqueous solutions. The CIP minimum in the PMFs is deeper for Ca^{2+} than for Na^{+}, which corresponds to much longer-lived Ca^{2+}–NOM complexes than Na^{+}–NOM complexes in aqueous solutions. The relative stability of such Ca^{2+}–NOM complexes enhances the supramolecular aggregation of the NOM molecules, which is observed experimentally.[12,13,15,17,18] Na^{+} is the least strongly held to carboxylic groups and, thus, is likely to have little effect on NOM aggregation.

For all TNB NOM systems, we consider the metal interaction with the three structurally distinct carboxylic groups individually. The site-specific potentials of mean force shown in Figure 8 are different for each caboxylate site due to steric constraints imposed by the structure of the TNB model. These steric effects also manifest themselves as prominent RDF and PMF features at relatively large distances near 9 Å, well beyond the range where strong site-site correlations might be expected (Figure 9). These features arise because the C23 and C25 carbons are nearly rigidly held at about this distance from each other by the aliphatic carbon backbone of the TNB molecular fragment (Figure 1). Therefore, when a cation is strongly associated with one of these carboxylic groups, the distance from that cation to the other carboxylic group is also nearly fixed, creating a false impression of strong metal-carboxylate correlations at unusually large distances.

In an analogous way the large and flexible TNB molecule affects the shape of the PMF curve at long distances beyond the SSIP range, although this is difficult to account for. Confirmation that this effect is due to the extent of the NOM molecule comes from comparing the PMF curves with those for acetate. The acetate molecule is smaller and almost spherical, and the steric effects are much less pronounced. These observations warn of the limitations of the

one-dimensional umbrella sampling method when applied to large molecules. Umbrella sampling using a single reaction coordinate provides the benefits of simplicity and clarity of interpretation, but it relies on uniform sampling of the angular coordinates of the system. For larger molecules, parts of the angular space are not accessible, and this introduces uncertainties in the computed PMFs. Another issue pertinent to the computation of all potentials is the requirement that the PMF should vanish at large separation. In all practical cases, however, one is forced to choose a finite distance where the PMF becomes zero. In our systems, we chose the distance of 8.5 Å. For large molecules this may not be sufficient, and our results show that at this distance, the PMF is significantly different from zero for all NOM systems.

The characteristic energies for all metal cation – carboxylic group interactions from our PMF calculations presented in Table 4 are consistent with the structural and dynamic results. Systems with Mg^{2+} show the lowest energy for bidentate and monodentate CIP coordination with a significant energy barrier between them, as well as the highest energy barrier separating the CIP and SSIP coordinations (see also Figure 7).

Calculation of the ion-pair association constants, Eq. (5), and the CIP/SSIP distribution coefficients, Eq. (4), from the computed PMFs is a sensitive test of the methods and provides one of the few ways to probe CIP/SSIP distributions in solution. The calculated association constants K_a of the cation – carboxylic group ion pairs are listed in Table 5, and the values of the equilibrium constants K_{eq} are listed in Table 6. The calculated ion association constant of the Na^+ – acetate ion pair is in a reasonable agreement with experimental results,[89-91] and the computed values of the association constant are consistent across different TNB carboxylic sites.

The association constants of Ca^{2+} for the different carboxylic sites of the TNB model of NOM are similar to each other but different for Ca^{2+}– acetate. For Ca^{2+}– acetate, there is a

considerable disagreement among the experimental results[89-93] with which the computed results should be compared (Table 5). The evaluation of these experimental differences is beyond the scope of this paper.

The computed values of the CIP/SSIP equilibrium constants, K_{eq}, (Table 6), for Na^+ are similar for all carboxylic sites in acetate and TNB and show the preference of Na^+ for CIP coordination. For Mg^{2+} and Ca^{2+}, all computed values show a preference for CIP over SSIP, and the values are of the same order of magnitude. There are no experimental values for the CIP/SSIP equilibrium constants, K_{eq}, with which to compare the computed results.

The differences among the computed K_a and K_{eq} values can arise from several sources. The results of the free MD simulations can be very sensitive to the initial setup, as described in Section III.1. This is most conspicuous for Mg^{2+}, for which the barrier crossing between CIP and SSIP coordinations occurs on a timescale longer than the length of the simulations. Thus, adequate sampling of the phase space of the Mg^{2+}–carboxylate ion pairs would require much longer simulations, and we do not report K_a values from free MD calculations. Another source of uncertainty comes from the fact that the umbrella sampling simulations restrained only the distance between the metal ion and the carboxylic group but not the angular variables and they assume that the angular space is probed uniformly. For Ca^{2+}, the results from free MD and umbrella sampling yield similar K_a values for the relatively small and almost spherical acetate anion. For the larger and more complex TNB molecule, however, the differences are significant. These differences reflect the uncertainties in the PMFs due to insufficiently uniform sampling of the angular variables, because some TNB conformations could make certain regions of the phase space temporarily or permanently inaccessible to the metal ions. This effect is difficult to quantify, and it is stronger in the case of *K*a calculations, because the steric obstruction is more

severe at shorter distances. The effect is less pronounced for K_{eq}, because this parameter probes the relative distribution between CIP and SSIP association.

The third significant source of uncertainty in the calculated K_a and K_{eq} values arises from the choice of the cutoff distances for the different coordinations, i.e. the values of R_U and R_L in Eq. (5) and Eq. (4). In MD simulations, these values come from the analysis of the positions of the minima and maxima of the corresponding radial distribution functions and are, ultimately, affected by the choice of the force field describing the interatomic interactions in the system. We address this issue in more detail elsewhere.[65,94] Thus, the computed K_a and K_{eq} values should be considered only as semi-quantitative estimates. To help evaluate the results, it will be important to expand and resolve the discrepancies among the experimental K_a values. It will also be necessary to develop better theoretical approaches to PMF calculations involving large and complicated molecules which significantly change their conformation during the simulation and for which the phase space cannot be adequately sampled using umbrella sampling with a single reaction coordinate.

3. Aggregation of the NOM molecules in the presence of Ca^{2+}

Ca^{2+} is thought to play an important role in causing supramolecular aggregation of relatively small NOM molecules (molecular weights from a few hundred to a few thousand Daltons) into larger colloidal particles.[15,17,18,23,31,44,45] The results of our MD simulations confirm that aggregation can occur even on the 10-ns time scale in systems with Ca^{2+} but not in systems with Na^+ or Mg^{2+}. The time evolution of the radii of gyration (eq. 6) for all pairs of NOM molecules from the Ca^{2+}–NOM free MD simulations (Figure 10) demonstrates aggregation through a decrease in the magnitude and convergence of the pair wise radii of gyration for

several pairs of NOM molecules. For example, the left panel of Figure 10 shows the evolution of R_{gyr} for all pairs of TNB molecules involving TNB1. The results show that TNB1 aggregates with TNB2 and TNB6, with the radii of gyration of these pairs decreasing from initial values of about 15-20 Å to about 9 Å. Similarly, the time evolution of R_{gyr} for all pairs of molecules involving TNB6 (right panel of Figure 10) shows its aggregation with TNB2 and TNB4 during the last 4 ns of the simulation. Analysis of the radii of gyration of all pairs of NOM fragments, shows that aggregates as large as four TNB fragments form at some time during the simulation. Figure 11 is a snapshot of a 3-molecule cluster with the three NOM molecules indicated by a different color for clarity. In this cluster calcium ion Ca2 is coordinated to two carboxylic groups of the same TNB molecule, and Ca4 is coordinated simultaneously to three carboxylic groups belonging to three different molecules. The latter structure confirms the earlier hypothetical picture of Ca^{2+} being able to accept up to four NOM carboxylic groups in its inner-sphere coordination shell.[32] The radii of gyration for the systems with Na^{+} or Mg^{2+} did not indicate any aggregation of NOM fragments. We have also observed noticeable conformational changes of the relatively large and flexible TNB molecules upon complexation with Ca^{2+}. These changes are especially prominent in situations when two carboxylic groups of the same NOM molecule are coordinated with the same metal ion. This conformational flexibility definitely contributes to easier aggregation of the NOM-metal ion complexes. However, a detailed quantitative exploration of this effect is beyond the scope of our present study.

All these observations suggest that the Ca^{2+} complexation may affect colloidal aggregation of NOM through a direct mechanism of cationic bridging between carboxylic groups from different NOM molecules, effectively bringing and holding them together (see Ca4 coordination in Figure 11). Calcium ions can also contribute to a second, indirect mechanism of

aggregation involving simultaneous Ca^{2+} complexation with two or more carboxylic groups of the same NOM molecule (see, for example, Ca2 coordination in Figure 11). This coordination produces a complex with a reduced net charge that is not repelled as strongly by other cation–NOM complexes, thus allowing the complexes to approach each other more readily. The size of the simulated systems used here is not large enough to assess the relative importance of these two aggregation mechanisms. Questions about the size distribution and lifetimes of the aggregated NOM supramolecular structures can be answered quantitatively only with MD simulations of much larger systems over longer times.

The systems involving acetate showed no evidence of anion aggregation beyond the ion-pair formation. There were, however, only four calcium ions available to coordinate with eight carboxylate molecules in the simulated system (Table 1). Thus, although the scale of the simulations was adequate for deriving the radial distribution functions and studying the contact and solvent separated ion pairs, it was not large enough to evaluate the degree and possible mechanisms of ion clustering in the acetate systems.

IV. CONCLUSIONS

Computational molecular dynamics study of the interactions of Na^{+}, Mg^{2+}, and Ca^{2+} and the carboxylic groups of a model NOM fragment and acetate anions in aqueous solutions provides a new quantitative insight into the structure, energetics, and dynamics of the interactions of metal cations with organic acids and the effects of cations on the colloidal aggregation of NOM molecules in solution. Site-specific and average potentials of mean force for the interaction of metal cations and carboxylic groups were derived from free MD simulations and restrained umbrella sampling simulations. These calculations provide a

quantitative description of the local structural environments of metal-carboxyl association and the dynamics of exchange among these sites in terms of the equilibrium constants describing overall ion association and the distribution of the metal cations between contact ion pairs and solvent separated ions pairs. All three cations prefer contact ion-pairs to solvent separated ion-pairs, and Na^+ and Ca^{2+} show a strong preference for bidentate contact ion pair coordination. The average residence time of a Ca^{2+} ion in a contact ion pair with the carboxylic groups is of the order of 1 ns, whereas the corresponding residence time of a Na^+ ion is only about 0.1 ns. The average residence times of a Ca^{2+} ion in a bidentate coordinated contact ion pair vs. a monodentate coordinated contact ion pair are about 0.5 ns and about 0.08 ns, respectively. Mg^{2+} shows no transitions between different coordination states during the simulations. On the 10-ns time scale of our simulations, the aggregation of the NOM molecules is observed in the presence of Ca^{2+} but not Na^+ or Mg^{2+}. Our results agree with previous experimental findings and are explained by both Ca^{2+} ion-bridging between NOM molecules and decreased repulsion between the NOM molecules due to the reduced net charge of the NOM-metal complexes. Larger scale simulations of Ca^{2+}–containing systems with different and more structurally diverse NOM models are needed to further explore and quantify the relative importance of the different aggregation mechanisms and the stability and dynamics of NOM aggregates.

Acknowledgments

This work was supported by the US Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences (grant number DE-FG02-08ER-15929). The computational resources were provided by the US DOE National Energy Research Scientific Computing Center (NERSC) and the National Science Foundation

TeraGrid facilities provided by the National Center for Supercomputing Applications (grant number TG- EAR000002) E. I.-T. is grateful to Dr. Alan Grossfield and Dr. O. Akin-Ojo for useful discussions.

References

(1) Stevenson, F. J. *Humus Chemistry: Genesis, Composition, Reactions*, 2nd ed. John Wiley, New York, 1994.

(2) Wershaw, R. L. Evaluation of conceptual models of natural organic matter (humus) from a consideration of the chemical and biochemical processes of humification. *U.S. Geological Survey Scientific Investigations Report 2004-5121*, Reston, VA, 2004.

(3) Leenheer, J. A. Systematic approaches to comprehensive analyses of natural organic matter. *Annals of Environmental Science* **2009,** *3*, 1-130.

(4) Buffle, J. *Complexation Reactions in Aquatic Systems: An Analytical Approach*. Ellis Horwood Ltd.: Chichester, 1988; p 692.

(5) Tipping, E. *Cation Binding by Humic Substances*. Cambridge University Press: Cambridge, 2002; p 434.

(6) Au, K.-K.; Penisson, A. C.; Yang, S.; O'Melia, C. R. *Geochim. Cosmochim. Acta* **1999,** *63*, 2903-2917.

(7) Mylon, S. E.; Chen, K. L.; Elimelech, M. *Langmuir* **2004**, *20*, 9000-9006.

(8) Tombácz, E.; Libor, Z.; Illés, E.; Majzik, A.; Klumpp, E. *Org. Geochem.* **2004,** *35*, 257-267.

(9) Yoon, T. H.; Johnson, S. B.; Brown, G. E. *Langmuir* **2005,** *21*, 5002-5012.

(10) Feng, X.; Simpson, A. J.; Simpson, M. J. *Org. Geochem.* **2005,** *36*, 1553-1566.

(11) Sutton, R.; Sposito, G. *Geochim. Cosmochim. Acta* **2006,** *70*, 3566-3581.

(12) Mylon, S. E.; Rinciog, C. I.; Schmidt, N.; Gutierrez, L.; Wong, G. C. L.; Nguyen, T. H. *Langmuir* **2010**, *26*, 1035-1042.

(13) Hong, S. K.; Elimelech, M. *J. Membr. Sci.* **1997,** *132*, 159-181.

(14) Kim, J.; Shan, W.; Davies, S. H. R.; Baumann, M. J.; Masten, S. J.; Tarabara, V. V. *Environ. Sci. Technol.* **2009,** *43*, 5488-5494.

(15) Wall, N. A.; Choppin, G. R. *Appl. Geochem.* **2003,** *18*, 1573-1582.

(16) Diallo, M. S.; Glinka, C. J.; Goddard, W. A.; Johnson, J. H. *Journal of Nanoparticle Research* **2005,** *7*, 435-448.

(17) Baalousha, M.; Motelica-Heino, M.; Coustumer, P. L. *Colloids and Surfaces A: Physicochem. Eng. Aspects* **2006**, *272*, 48-55.

(18) Brigante, M.; Zanini, G.; Avena, M. *Colloids and Surfaces A: Physicochem. Eng. Aspects* **2007,** *294*, 64-70.

(19) Schaumann, G. E. *Journal of Plant Nutrition and Soil Science-Zeitschrift Fur Pflanzenernahrung Und Bodenkunde* **2006,** *169*, 145-156.

(20) Yang, R.; van den Berg, C. M. G. *Environ. Sci. Technol.* **2009**, *43*, 7192-7197.

(21) Majzik, A.; Tombácz, E. *Org. Geochem.* **2007,** *38*, 1319-1329.

(22) Majzik, A.; Tombácz, E. *Org. Geochem.* **2007,** *38*, 1330-1340.

(23) Pinheiro, J. P.; Mota, A. M.; Benedetti, M. F. *Environ. Sci. Technol.* **1999,** *33*, 3398-3404.

(24) Rey-Castro, C.; Mongin, S.; Huidobro, C.; David, C.; Salvador, J.; Garcés, J. L.; Galceran, J.; Mas, F.; Puy, J. *Environ. Sci. Technol.* **2009,** *43*, 7184-7191.

(25) Wershaw, R. L. *Soil Science* **1999,** *164*, 803-813.

(26) Piccolo, A. *Soil Science* **2001,** *166*, 810-832.

(27) Leenheer, J. A.; Croué, J. P. *Environ. Sci. Technol.* **2003**, *37*, 18A-26A.

(28) Sutton, R.; Sposito, G. *Environ. Sci. Technol.* **2005,** *39*, 9009-9015.

(29) Simpson, A.; Kingery, W.; Hayes, M.; Spraul, M.; Humpfer, E.; Dvortsak, P.; Kerssebaum, R.; Godejohann, M.; Hofmann, M. *Naturwissenschaften* **2002,** *89*, 84-88

(30) Šmejkalová, D.; Piccolo, A. *Environ. Sci. Technol.* **2008,** *42*, 699-706.

(31) Nebbioso, A.; Piccolo, A. *Environ. Sci. Technol.* **2009,** *43*, 2417-2424.

(32) Leenheer, J. A.; Brown, G. K.; MacCarthy, P.; Cabaniss, S. E. *Environ. Sci. Technol.* **1998,** *32*, 2410-2416.

(33) Cabaniss, S.; Madey, G.; Leff, L.; Maurice, P.; Wetzel, R. *Biogeochemistry* **2007,** *86*, 269-286.

(34) Cabaniss, S. E. *Environ. Sci. Technol.* **2008**, *42*, 5210-5216.

(35) Cabaniss, S. E. *Environ. Sci. Technol.* **2009,** *43*, 2838-2844.

(36) Orsetti, S.; Andrade, E. M.; Molina, F. V. *Langmuir* **2010***, 26*, 3134-3144.

(37) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*. Oxford University Press: New York, 1987; p 385.

(38) Jansen, S. A.; Malaty, M.; Nwabara, S.; Johnson, E.; Ghabbour, E.; Davies, G.; Varnum, J. M. *Materials Science & Engineering C-Biomimetic Materials Sensors and Systems* **1996,** *4*, 175-179.

(39) Davies, G.; Fataftah, A.; Cherkasskiy, A.; Ghabbour, E. A.; Radwan, A.; Jansen, S. A.; Kolla, S.; Paciolla, M. D.; Sein, L. T.; Buermann, W.; Balasubramanian, M.; Budnick, J.; Xing, B. S. *J. Chem. Soc. Dalton Trans.* **1997**, 4047-4060.

(40) Sein, L. T.; Varnum, J. M.; Jansen, S. A. *Environ. Sci. Technol.* **1999,** *33*, 546-552.

(41) Kubicki, J. D.; Apitz, S. E. *Org. Geochem.* **1999,** *30*, 911-927.

(42) Diallo, M. S.; Simpson, A.; Gassman, P.; Faulon, J. L.; Johnson, J. H.; Goddard, W. A.; Hatcher, P. G. *Environ. Sci. Technol.* **2003,** *37*, 1783-1793.

(43) Xu, X.; Kalinichev, A. G.; Kirkpatrick, R. J. *Geochim. Cosmochim. Acta* **2006,** *70*, 4319-4331.

(44) Kalinichev, A. G.; Kirkpatrick, R. J. *Eur. J. Soil Sci.* **2007,** *58*, 909-917.

(45) Ahn, W. Y.; Kalinichev, A. G.; Clark, M. M. *J. Membr. Sci.* **2008,** *309*, 128-140.

(46) Selvarengan, P.; Kubicki, J. D.; Guégan, J. P.; Châtellier, X. *Chem. Geol.* **2010,** *273*, 55-75.

(47) Sutton, R.; Sposito, G.; Diallo, M. S.; Schulten, H. R. *Environ. Toxicol. Chem.* **2005,** *24*, 1902-1911.

(48) Impey, R. W.; Madden, P. A.; McDonald, I. R. *J. Phys. Chem.* **1983,** *87*, 5071-5083.

(49) Ohtaki, H.; Radnai, T. *Chem. Rev.* **1993,** *93*, 1157-1204.

(50) Carrillo-Tripp, M.; Saint-Martin, H.; Ortega-Blake, I. *J. Chem. Phys.* **2003,** *118*, 7062-7073.

(51) Hofer, T. S.; Tran, H. T.; Schwenk, C. F.; Rode, B. M. *J. Comp. Chem.* **2004,** *25*, 211-217.

(52) Rode, B. M.; Schwenk, C. F.; Hofer, T. S.; Randolf, B. R. *Coord. Chem. Rev.* **2005,** *249*, 2993-3006.

(53) Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. *J. Phys. Chem. B* **2006,** *110*, 18553-18559.

(54) Masamura, M. *J. Phys. Chem.* **1993,** *97*, 3157-3159.

(55) Leung, K.; Rempe, S. B. *J. Amer. Chem. Soc.* **2004,** *126*, 344-351.

(56) Liu, X.; Lu, X.; Wang, R.; Zhou, H.; Xu, S. *Geochim. Cosmochim. Acta* **2008,** *72*, 5896-5907.

(57) Gojło, E.; Śmiechowski, M.; Panuszko, A.; Stangret, J. *J. Phys. Chem. B* **2009,** *113*, 8128-8136.

(58) Denecke, M. A.; Reich, T.; Pompe, S.; Bubner, M.; Heise, K. H.; Nitsche, H.; Allen, P. G.; Bucher, J. J.; Edelstein, N. M.; Shuh, D. K.; Czerwinski, K. R. *Radiochimica Acta* **1998,** *82*, 103-108.

(59) Jeong, C. Y.; Young, S. D.; Marshall, S. J. *Soil Sci. Soc. Am. J.* **2007,** *71*, 515-528.

(60) Armbruster, M. K.; Schimmelpfennig, B.; Plaschke, M.; Rothe, J.; Denecke, M. A.; Klenze, R. *Journal of Electron Spectroscopy and Related Phenomena* **2009,** *169*, 51-56.

(61) Aquino, A. J. A.; Tunega, D.; Pasalic, H.; Haberhauer, G.; Gerzabek, M. H.; Lischka, H. *Chem. Phys.* **2008,** *349*, 69-76.

(62) Aquino, A. l. J. A.; Tunega, D.; Schaumann, G. E.; Haberhauer, G.; Gerzabek, M. H.; Lischka, H. *J. Phys. Chem. C* **2009,** *113*, 16468-16475.

(63) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.,*J. Chem. Phys.* **1983,** *79*, 926-935.

(64) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000,** *21*, 1049-1074.

(65) We have recently explored the effects of using several different force fields on the modeling of metal cation – NOM complexation in aqueous solutions (see ref. 94 for details). One encouraging result of that comparison is that the main qualitative and quantitative characteristics of such complexation are quite robust and are also consistent with available experimental data.

(66) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER 9*; University of California: San Francisco, 2006.

(67) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000,** *21*, 132-146.

(68) Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002,** *23*, 1623-1641.

(69) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Lelong, R. ; Grivel, N. ; Cieplak, P., http://q4md-forcefieldtools.org/REDDB/projects/W-46/ (2005)

(70) Kollman, P. *Chem. Rev.* **1993,** *93*, 2395-2417.

(71) McQuarrie, D. A. *Statistical Mechanics*. 2 ed.; University Science Books: Sausalito, CA, 2000; 641 pp.

(72) Trzesniak, D.; Kunz, A.-P. E.; van Gunsteren, W. F. *ChemPhysChem* **2007,** *8*, 162-169

(73) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187-199.

(74) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1339-1350.

(75) Kästner, J.; Thiel, W. *J. Chem. Phys.* **2005**, *123*, 144104-5.

(76) Kästner, J.; Thiel, W. *J. Chem. Phys.* **2006**, *124*, 234106-7.

(77) Khavrutskii, I.; Dzubiella, J.; McCammon, J. *J. Chem. Phys.* **2008**, *128*, 044106-13.

(78) Guàrdia, E.; Rey, R.; Padró, J. A. *Chem. Phys.* **1991**, *155*, 187-195.

(79) Chialvo, A. A.; Cummings, P. T.; Cochran, H. D.; Simonson, J. M.; Mesmer, R. E. *J. Chem. Phys.* **1995,** *103*, 9379-9387.

(80) Khokhlov, A. R.; Grosberg, A. Y.; Pande, V. J. *Statistical Physics of Macromolecules (Polymers and Complex Materials)*, American Institute of Physics, 2002, 384 pp.

(81) Lu, Y.; Miller, J. D. *J. Coll. Interf. Sci.* **2002,** *256*, 41-52.

(82) Dudev, T.; Lim, C. *Chem. Rev.* **2003,** *103*, 773-787.

(83) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2004,** *108*, 4546-4557.

(84) McGill, P. R.; Idriss, H. *Surface Science* **2008,** *602*, 3688-3695.

(85) Falke, J. J.; Drake, S. K.; Hazard, A. L.; Peersen, O. B.. *Quart. Rev. Biophys.* **1994,** *27*, 219-290.

(86) Cates, M. S.; Teodoro, M. L.; Phillips Jr, G. N. *Biophysical Journal* **2002,** *82*, 1133-1146.

(87) Rapaport, D. *Molecular Physics* **1983**, 50, 1151.

(88) Luzar, A. *J. Chem. Phys.* **2000**, 113, 10663.

(89) Archer, D. W.; Monk, C. B. *J. Chem. Soc.* **1964**, 3117-3122.

(90) Daniele, P. G.; De Robertis, A.; De Stefano, C.; Sammartano, S.; Rigano, C. *J. Chem. Soc. Dalton Trans.* **1985**, 2353-2361.

(91) Garrels, R. M.; Thompson, M. E. *Amer. J. Sci.* **1962,** *260*, 57-66.

(92) Shock, E. L.; Koretsky, C. M. *Geochim. Cosmochim. Acta* **1993,** *57*, 4899-4922.

(93) Colman-Porter, C. A.; Monk, C. B. *J. Chem. Soc.* **1952**, 4363-4368.

(94) Kalinichev, A. G.; Iskrenova-Tchoukova, E.; Ahn, W.-Y.; Clark, M. M.; Kirkpatrick, R. J. *Geoderma* **2010**, submitted.

Table 1. Simulated systems: size, number of atoms, number of TIP3P water molecules.

System	H_2O molecules	Total atoms
Free MD simulations		
4 Ca^{2+} – 8 ACE	2200	6660
4 Mg^{2+} – 8 ACE	2200	6660
12 Ca^{2+} – 8 TNB	4215	13385
12 Mg^{2+} – 8 TNB	4215	13385
24 Na^{+} – 8 TNB	4215	13397
Umbrella sampling simulations		
Mg^{2+} – ACE	753	2268
Mg^{2+} – Na^{+} – TNB	1440	4413
Ca^{2+} – Na^{+} – TNB	1440	4413

Table 2. Characteristic distances (in Å) of the cation – carboxylic group interactions: CIP(b) – location of the PMF minimum in the bidentate coordinated contact ion pair; CIP(m) – location of the PMF minimum in the monodentate coordinated CIP; SSIP – location of the PMF minimum of the solvent separated ion pair. R_L and R_U define the lower and upper threshold distances used to define the CIP and SSIP in Eqs. (4) and (5), while R_{bident} defines the boundary between the bidentate and monodentate coordinated contact ion pair. For comparison, three characteristic distances between the metal ion and the carbon of the carboxyl groups of other molecules from the quantum chemical calculations of Selvarengan *et al.*[46] are given in parenthesis for 2-keto-3-deoxyoctanoate ([a]), galacturonate ([b]), and carbonate ([c]) species.

	CIP(b)	R_{bident}	**CIP(m)**	R_L	**SSIP**	R_U
			Na^+			
Na^+	2.7	-	-	3.9	5.3	6.4
			Mg^{2+}			
Mg^{2+}	2.3 (2.48)[a] (2.55)[b]	2.5	3.0 (3.19)[a] (3.01)[b] (3.25)[c]	3.8	4.8 (5.05)[a] (4.23)[b]	5.9
			Ca^{2+}			
Ca^{2+}	3.1 (2.78)[a] (3.09)[b]	3.4	3.6 (3.42)[a] (3.40)[b] (3.37)[c]	4.2	5.7 (4.62)[a] (4.48)[b]	6.7

Table 3. Lifetimes (decorrelation times), in picoseconds, for persistent: CIP(b) – bidentate coordinated CIP, CIP(m) – monodentate coordinated CIP, and SSIP, for the four systems with Ca^{2+}, Na^{+}, NOM, and acetate (ACE).

	CIP(b)	CIP(m)	SSIP
Ca^{2+} – NOM	450	80	210
Ca^{2+} – ACE	500	10	20
Na^{+} – NOM	50	20	20
Na^{+} – ACE	30	10	10

Table 4. Characteristic energies (in kcal/mol) from the potentials of mean force calculations for all metal cation – carboxylic group interactions. "TS" stands for the transition state or the maximum barrier hight between the CIP and SSIP coordinations.

	ACE	**TNB**	**TNB–C23**	**TNB–C25**	**TNB–C29**
Na^+					
CIP (b)	–1.5	–1.6	–1.5	–1.3	–1.0
CIP(m)	–	–	–	–	–
TS	2.1	0.9	1.0	0.5	1.5
SSIP	–0.5	–0.4	–0.5	–0.5	–0.2
Mg^{2+}					
CIP (b)	–19.6	–9.3	–32.4	–1.9	–1.0
CIP (m)	–20.4	–11.7	–29.6	–6.8	2.1
TS	6.2	9.4	4.5	8.3	23.8
SSIP	–0.04	–0.5	0.8	–2.9	2.5
Ca^{2+}					
CIP (b)	–3.5	–2.8	–3.2	–2.8	–1.8
CIP (m)	0.8	–0.7	–0.4	–1.1	–0.7
TS	5.8	1.6	2.4	1.3	2.0
SSIP	0.2	–0.9	–0.7	–1.1	–1.1

Table 5. Ion association constants, K_a, of metal cations with individual binding sites (Eq. 5). The experimental results are taken from Ref. 92.

	ACE	TNB	TNB(C23)	TNB(C25)	TNB(C29)
Na^+					
Free MD	1.43	1.38	1.42	1.50	1.19
Exp.[a]	0.66				
Exp.[b]	0.78				
Ca^{2+}					
UI	1.58	0.41	0.15	0.14	0.24
Free MD	1.90	2.05	1.81	2.19	1.99
Exp.[a]	13.16				
Exp.[b]	8.55				
Exp.[c]	5.88				

[a] Exp: Archer and Monk, 1964 (ref. 89)

[b] Exp: Daniele *et al.*, 1985 (ref. 90)

[c] Exp: Coleman-Porte and Monk, 1952 (ref. 93)

Table 6: Equilibrium constants, K_{eq} (Eq. 4).

	ACE	TNB	TNB(C23)	TNB(C25)	TNB(C29)
Na^{+}	4.0	4.1	3.8	4.2	4.4
Mg^{2+}	4.4	9.3	5.3	1.1	1.4
Ca^{2+} (UI)	1.3	2.1	4.3	1.5	1.4
Ca^{2+} (free MD)	4.6	3.3	2.7	3.2	4.9

Figure captions

Figure 1: TNB model[38-40] of the NOM fragment used in this work. The colors of the atoms are as follows: carbon – dark gray , oxygen – red, hydrogen – white, and nitrogen – blue. The three different carboxyl groups are labeled as C23, C25, and C29. Bidentate and monodetate coordinated contact ion pairs are illustrated for C29 and C23 carboxylic groups, respectively.

Figure 2: Radial distribution functions of all metal ion – carboxyl carbon ion pairs. Left column – acetate; right column – NOM.

Figure 3: Definitions of the angular components of the radial vector describing the position of the metal ion with respect to the carboxylic group.

Figure 4: Angular distribution of the calcium ions around the carboxyl C atoms. The number of events for each pair of angles θ and φ is plotted. **(a)** Bidentate contact ion pair (CIP) coordination in the Ca^{2+} – TNB system (distances between 3.0 and 3.5 Å); **(b)** Monodentate CIP coordination in the Ca^{2+} – TNB system (distances between 3.5 and 4.0 Å); **(c)** Solvent separated ions pairs in the Ca^{2+} – TNB system (distances between 4.5 and 6.0 Å.

Figure 5: Survival probabilities of intermittent and persistent metal ion – carboxylate carbon coordination in **(a)** Ca2+ – NOM and **(b)** Na+– NOM. Notation in the legend: 1 –

intermittent CIP(b); 2 – intermittent CIP(m); 3 – intermittent SSIP; 4 – persistent CIP(b); 5 – persistent CIP(m); 6 – persistent SSIP.

Figure 6: Evolution of the distance between Na^+, Mg^{2+}, and Ca^{2+} ions and the coordinated to them C atoms of the TNB carboxyl groups. The distances on the vertical axes are in Å. Note the different timescale for the Na^+ dynamics (top panel); Na^+ is much more mobile than Mg^{2+} and Ca^{2+}. Over the 10-ns scale of our simulations, the Mg^{2+} ions never change their coordination with respect to the carboxylic groups. Three different Mg^{2+} ions are shown (middle panel) that are coordinated in bidentate CIP (2.3 Å), monodentate CIP (3.03 Å), and SSIP (4.7 Å).

Figure 7: Potentials of mean force for all metal cations – carboxylic group systems. The energy scales are different for each metal ion.

Figure 8: Site-specific potentials of mean force for all metal cations – carboxylic group systems.

Figure 9: Radial distribution function and the potential of mean force for Ca^{2+} – TNB interaction. The minimum in PMF (maximum in RDF) at about 9 Å is due to the specific geometry of the TNB molecule and the fact that two of the three carboxylic groups, namely C23 and C25 (Fig.1), remain at about that distance from each other due to the relatively rigid molecular structure of the TNB fragment.

Figure 10: Evolution of the radii of gyration during the simulation. (a) Radii of gyration of all pairs of molecules involving TNB1, The radii of gyration of pairs: TNB1-TNB2 and TNB1-TNB6 systematically decrease during the simulation. We identify three NOM fragments: TNB1, TNB2, and TNB6, that aggregate. (b) Evolution of the radii of gyration of the four pairs of NOM fragments that we have identified as clustering: TNB1-TNB2, TNB1-TNB6, TNB2-TNB6, and TNB4-TNB6.

Figure 11: A snapshot from the free MD simulation illustrating the NOM (TNB) aggregation in the presence of Ca^{2+} ions. A long-lived cluster of three TNB molecules and three Ca^{2+} ions is observed. H_2O molecules and other ions are removed for clarity. Different TNB molecules are represented by different colors, whereas calcium ions are represented by blue spheres (not in scale). Ca4 is simultaneously coordinated by the 1C25 and 2C29 carboxilic groups, both in bidentate configuration, and by the 6C23 carboxilic group in a monodentate configuration. Ca2 is coordinated by 1C23 and 1C29 carboxilic groups of the same molecule in bidentate innersphere and .outersphere coordination, respectively.

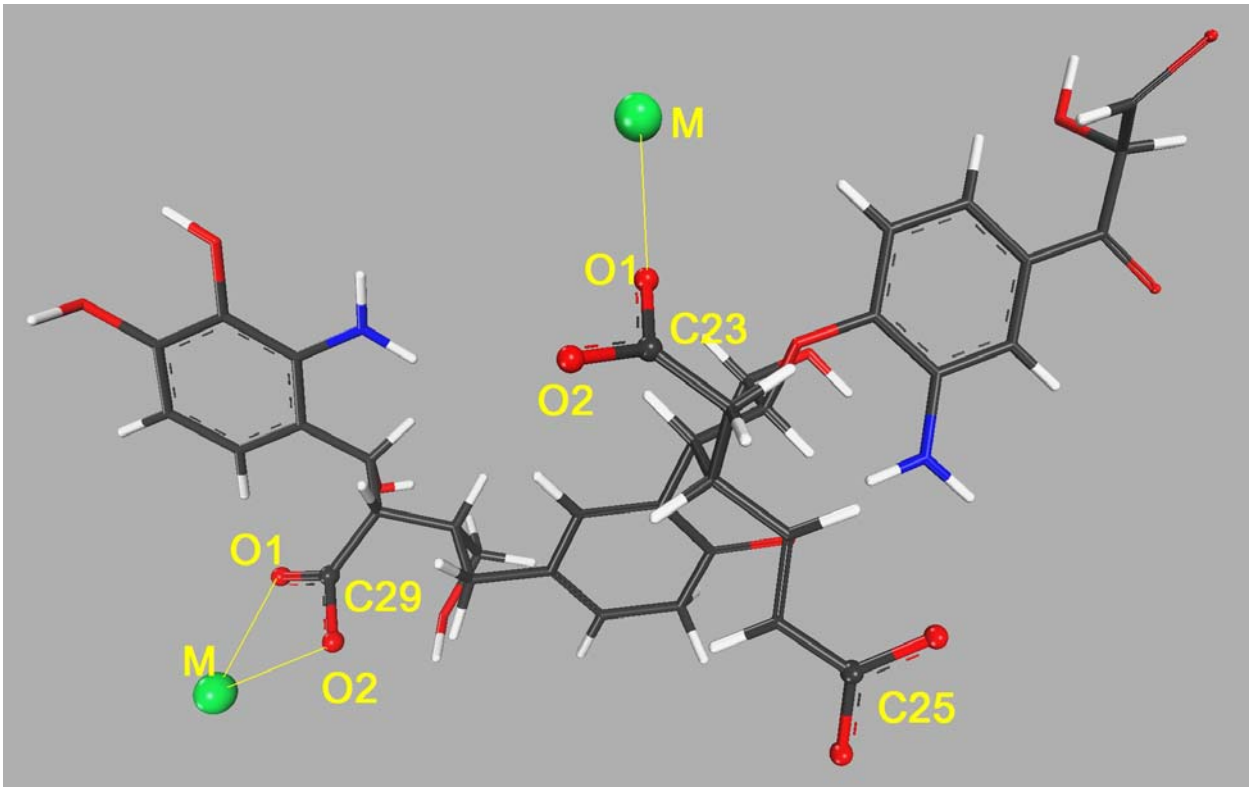

FIG. 1: TNB model[38-40] of the NOM fragment used in this work. The colors of the atoms are as follows: carbon – dark gray , oxygen – red, hydrogen – white, and nitrogen – blue. The three different carboxyl groups are labeled as C23, C25, and C29. Bidentate and monodetate coordinated contact ion pairs are illustrated for C29 and C23 carboxylic groups, respectively.

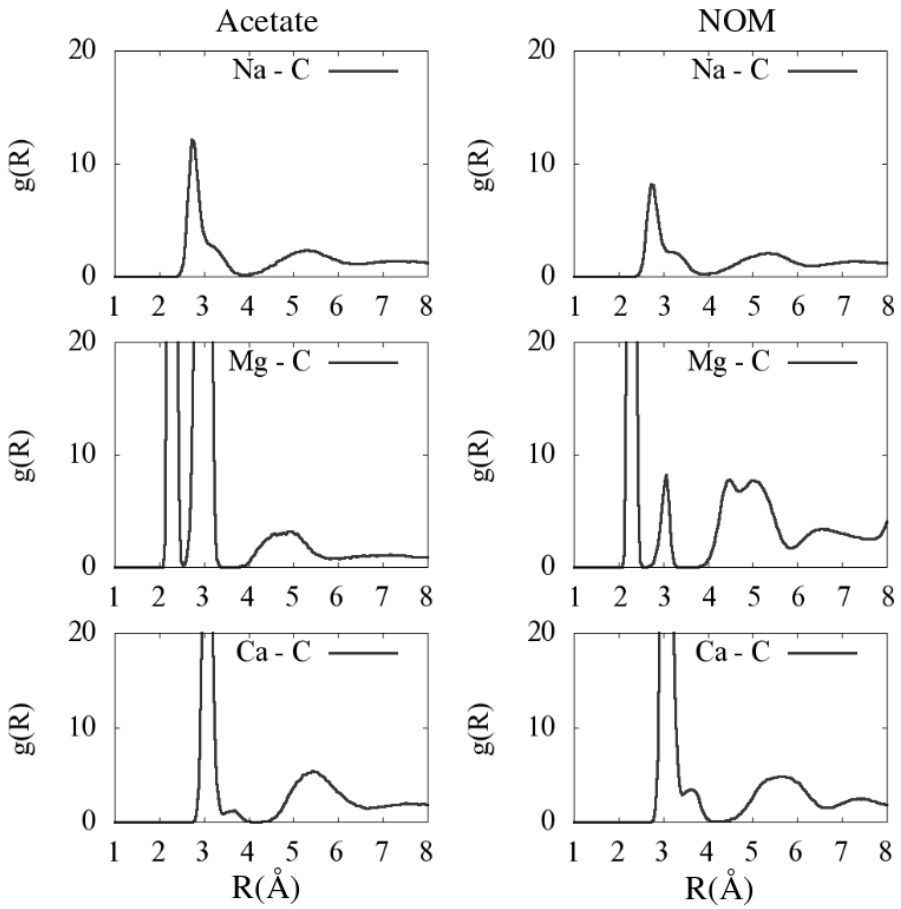

FIG. 2. Radial distribution functions of all metal ion – carboxyl carbon ion pairs. Left column – acetate; right column – NOM.

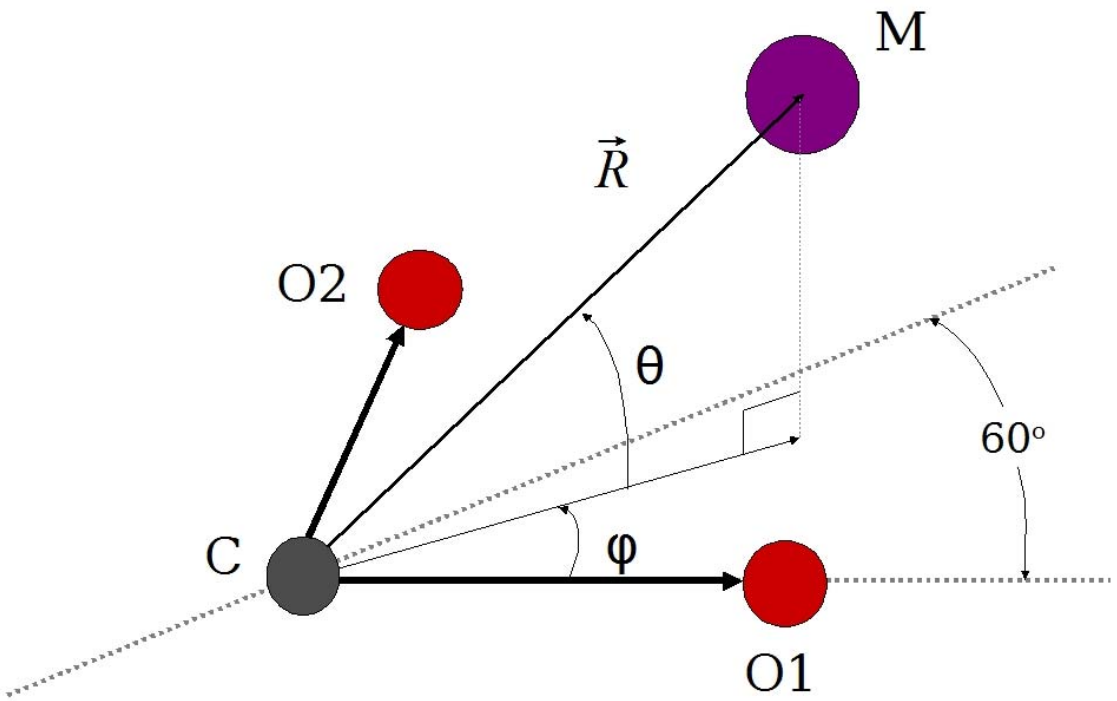

FIG. 3: Definitions of the angular components of the radial vector describing the position of the metal ion with respect to the carboxylic group.

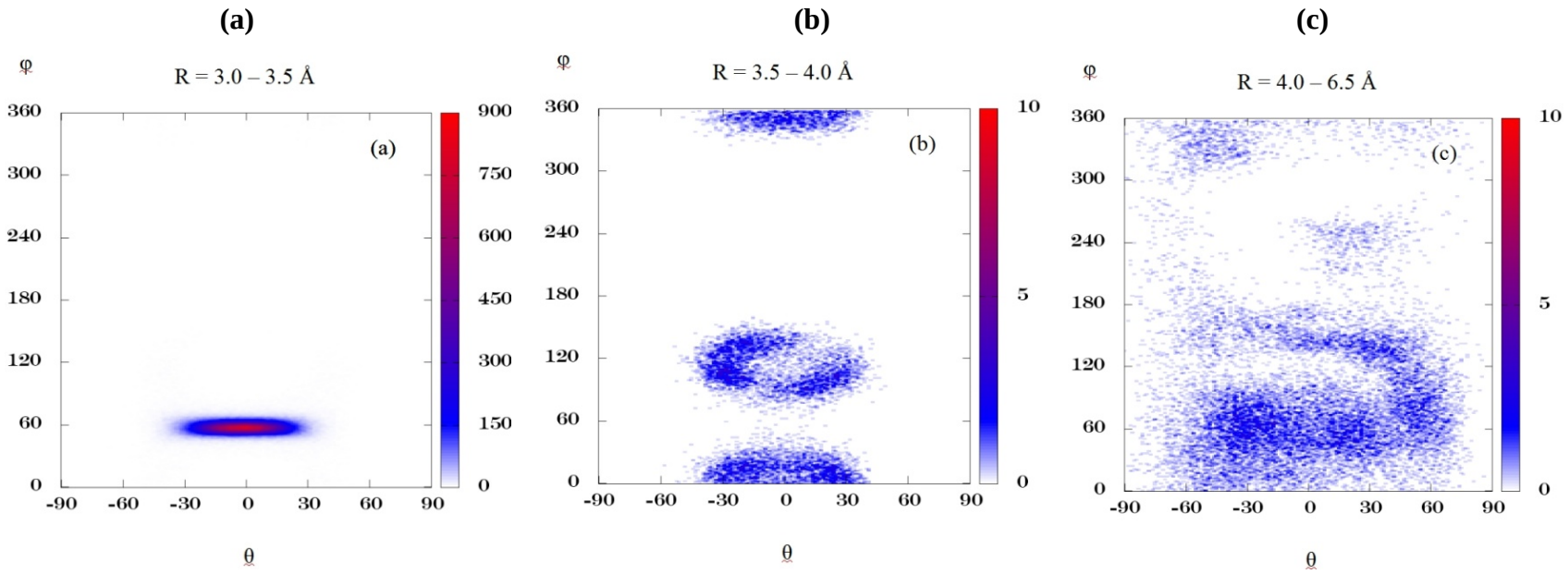

FIG. 4: Angular distribution of the calcium ions around the carboxyl C atoms. The number of events for each pair of angles θ and φ is plotted. **(a)** Bidentate contact ion pair (CIP) coordination in the Ca^{2+} – TNB system (distances between 3.0 and 3.5 Å); **(b)** Monodentate CIP coordination in the Ca^{2+} – TNB system (distances between 3.5 and 4.0 Å); **(c)** Solvent separated ions pairs in the Ca^{2+} – TNB system (distances between 4.5 and 6.0 Å). Both mono- and bi-dentate coordinations are observed

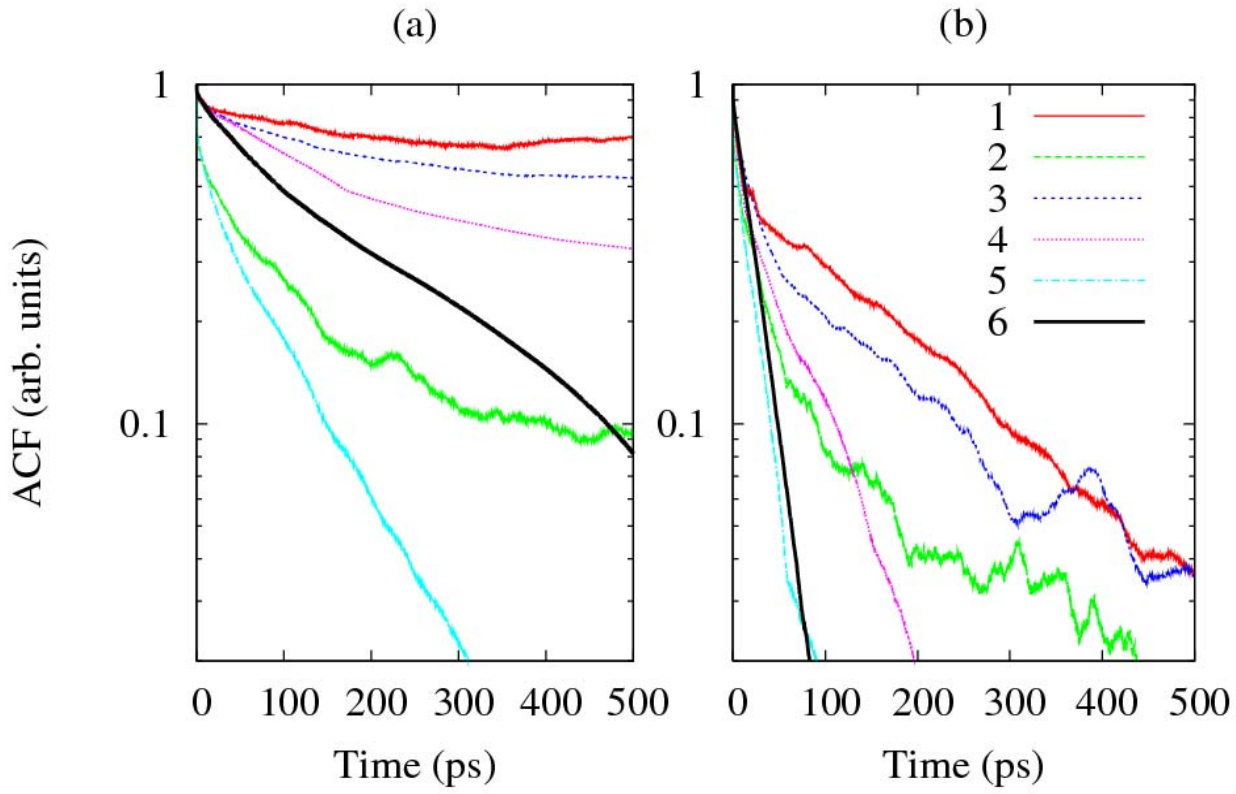

FIG. 5: Survival probabilities of intermittent and persistent metal ion – carboxylate carbon coordination in (a) Ca^{2+} – NOM and (b) Na^{+}– NOM. Notation in the legend: 1 – *intermittent* CIP(b); 2 – *intermittent* CIP(m); 3 – *intermittent* SSIP; 4 – *persistent* CIP(b); 5 – *persistent* CIP(m); 6 – *persistent* SSIP.

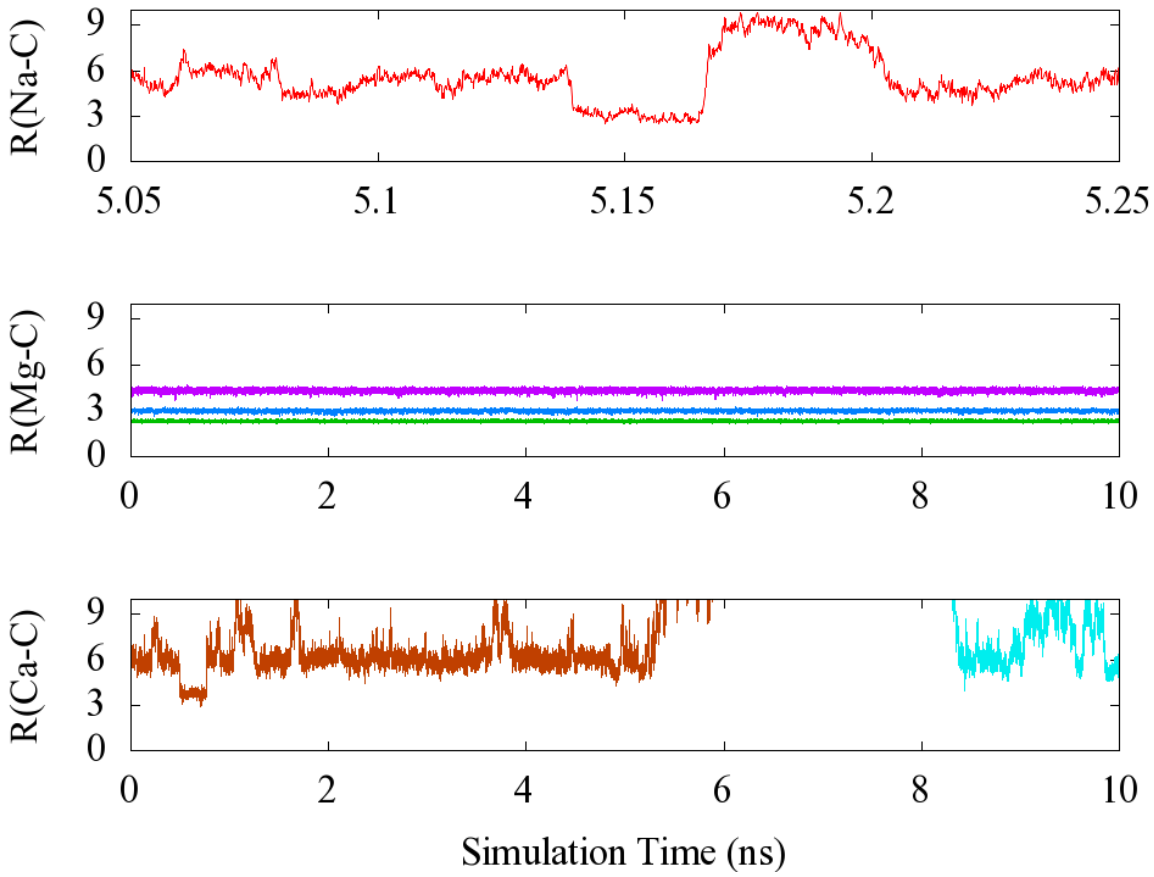

FIG. 6: Evolution of the distance between Na^{+}, Mg^{2+}, and Ca^{2+} ions and the coordinated to them C atoms of the TNB carboxyl groups in unrestrained MD simulations. The distances on the vertical axes are in Å. Note the different timescale for the Na^{+} dynamics (top panel), illustrating that Na^{+} is much more mobile than Mg^{2+} and Ca^{2+}. Over the 10-ns scale of our simulations, the Mg^{2+} ions never change their coordination with respect to the carboxylic groups. Three different Mg^{2+} ions are shown (middle panel) that are coordinated in bidentate CIP (2.3 Å), monodentate CIP (3.03 Å), and SSIP (4.7 Å) configuration..

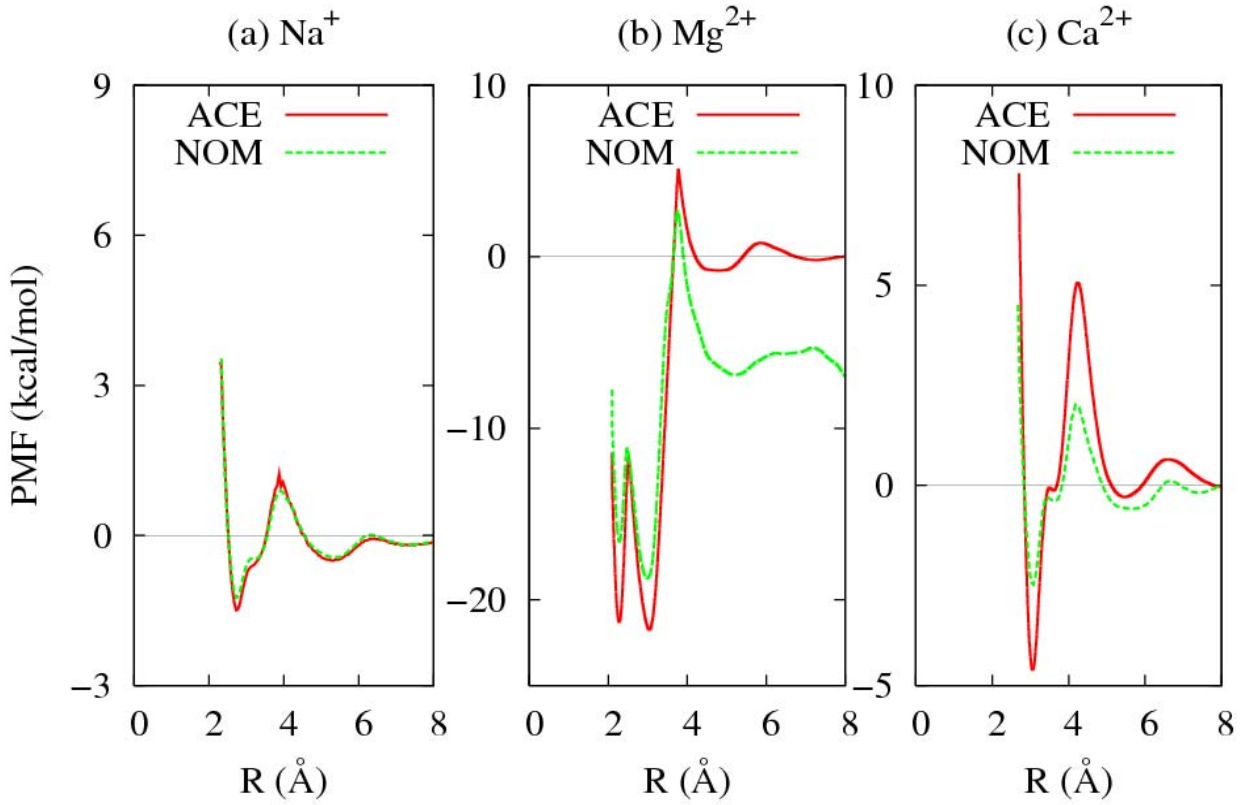

FIG. 7: Potentials of mean force for all metal cations – carboxylic group systems. The energy scales are different for each metal ion.

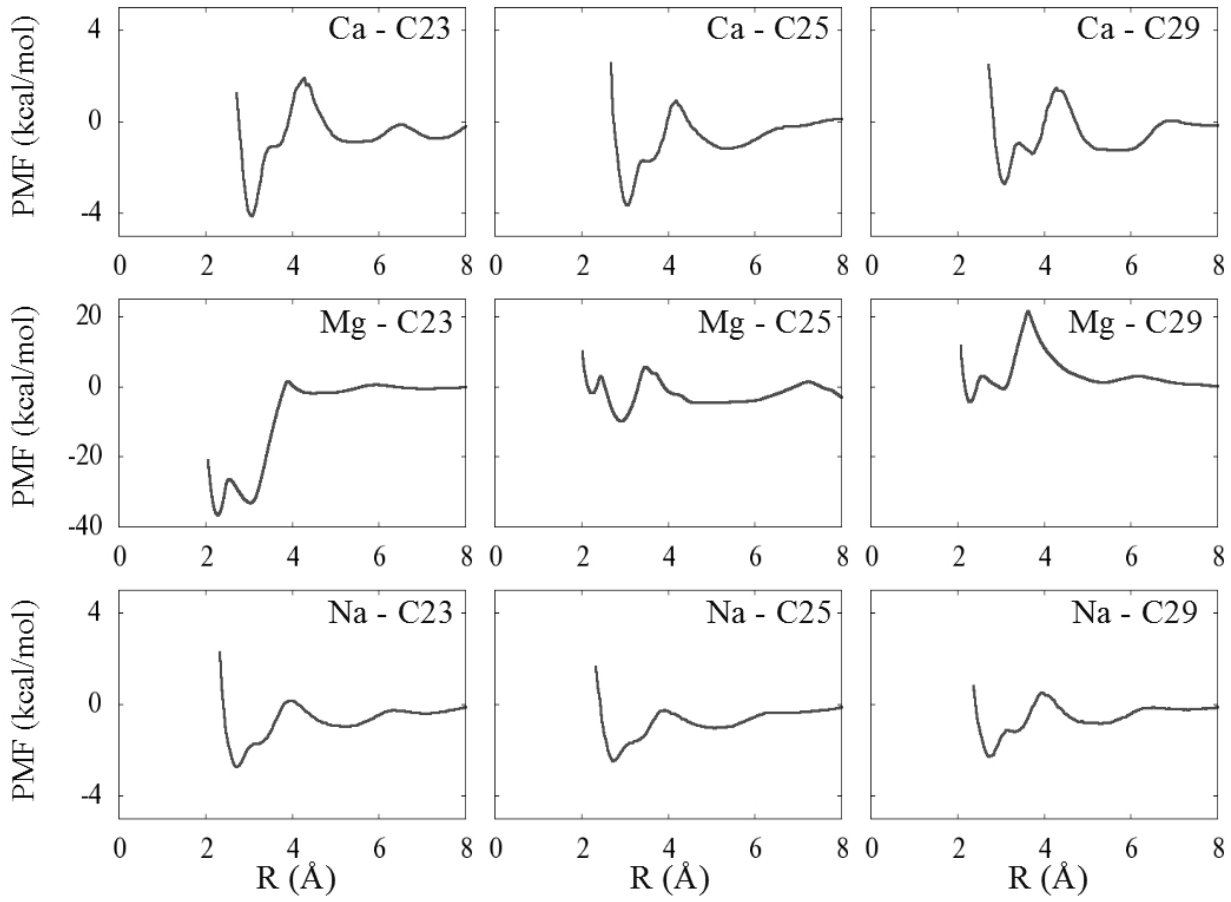

FIG. 8: Site-specific potentials of mean force for all metal cations – carboxylic group systems.

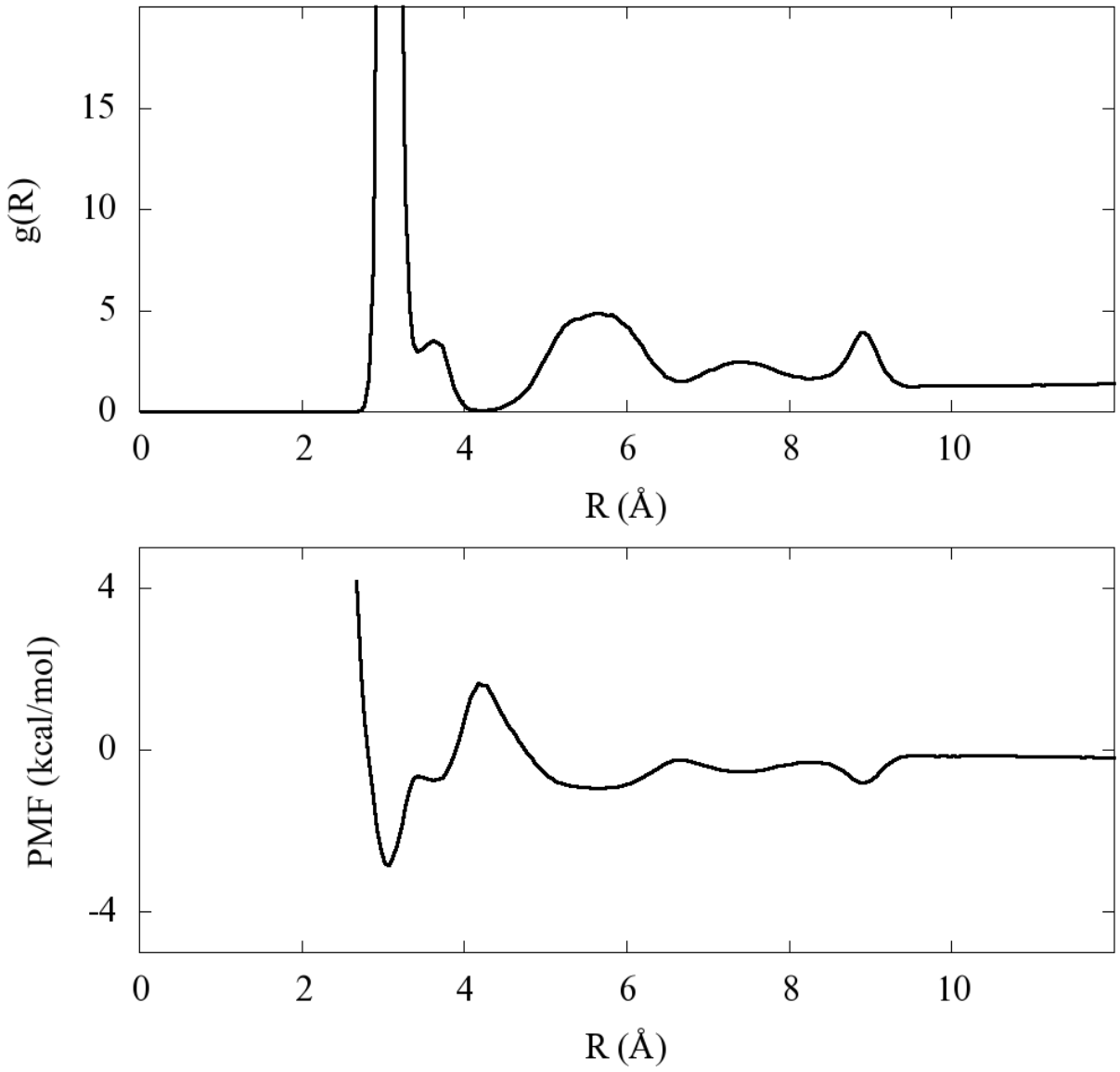

FIG. 9: Radial distribution function and the potential of mean force for Ca^{2+} – TNB interaction. The minimum in PMF (maximum in RDF) at about 9 Å is due to the specific geometry of the TNB molecule and the fact that two of the three carboxylic groups, namely C23 and C25 (Fig.1), remain at about that distance from each other due to the relatively rigid molecular structure of the TNB fragment.

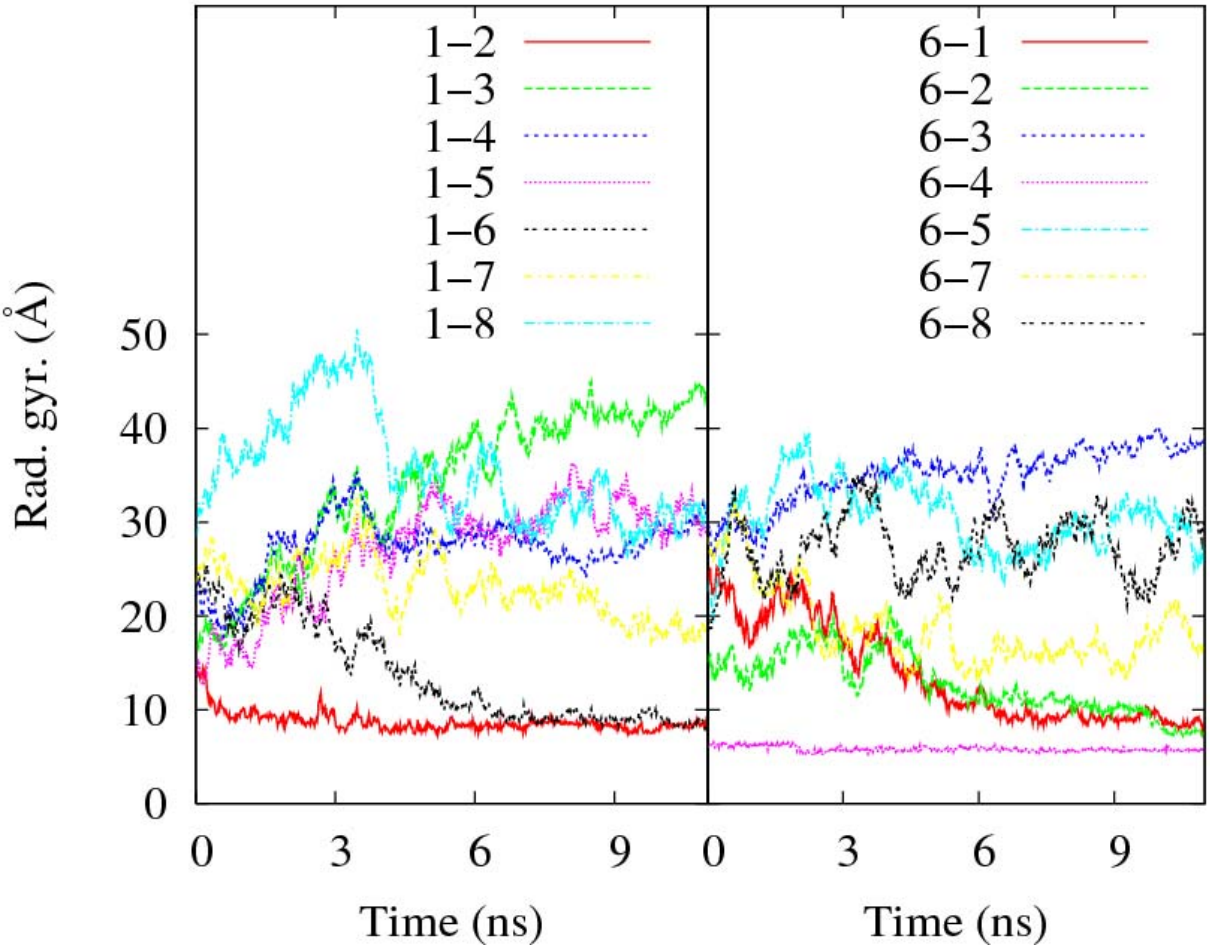

FIG. 10: Evolution of the radii of gyration during the simulation. *Left panel* – Radii of gyration of all pairs of molecules involving TNB1, The radii of gyration of pairs: TNB1-TNB2 and TNB1-TNB6 systematically decrease during the simulation. We identify three NOM fragments: TNB1, TNB2, and TNB6 that aggregate. *Right panel* – Evolution of the radii of gyration of all pairs of molecules involving TNB6. We can identify the four pairs of NOM fragments that aggregate during the last four picoseconds of the simulation: TNB1-TNB2, TNB1-TNB6, TNB6-TNB2, and TNB6-TNB4.

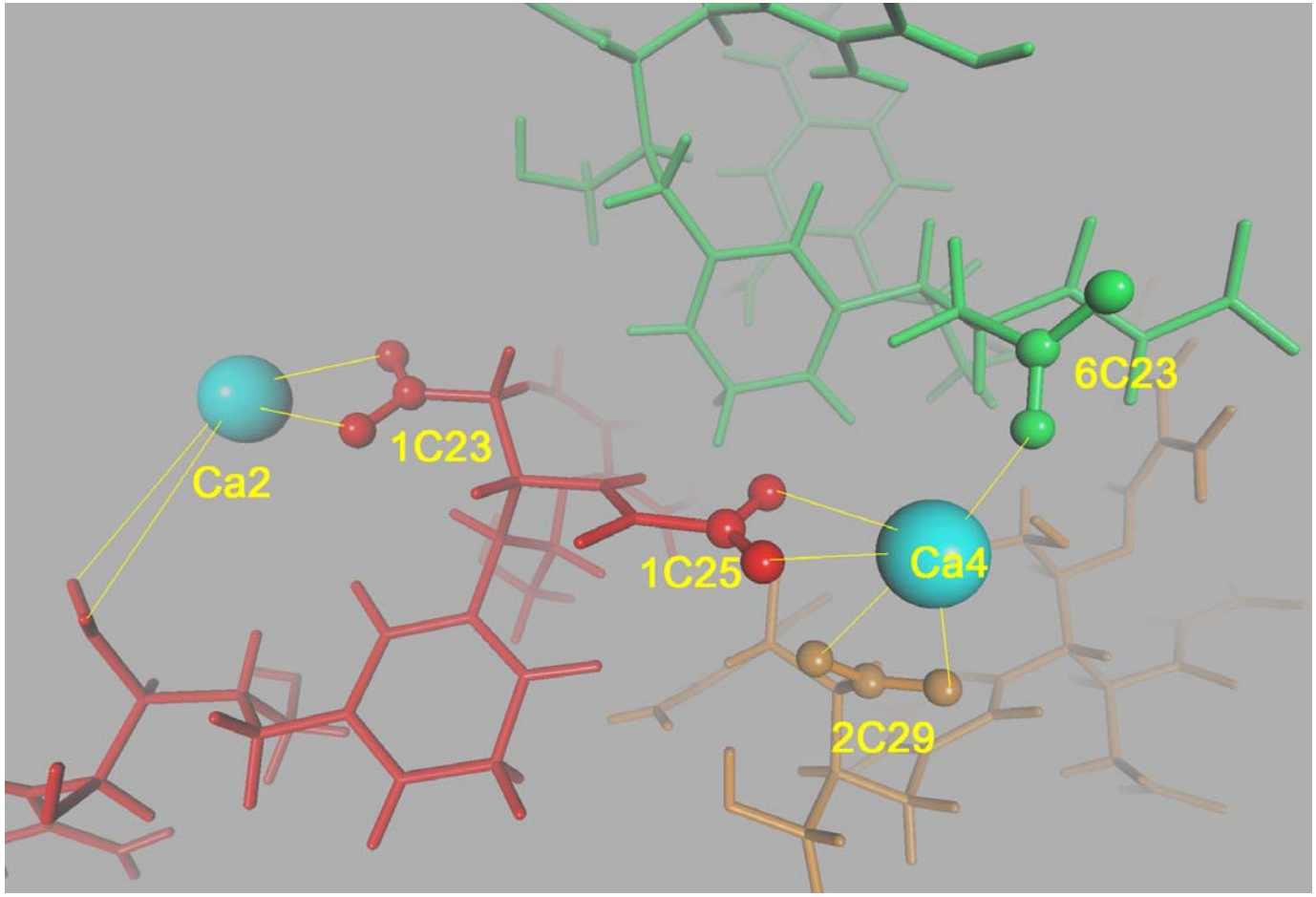

FIG. 11: A snapshot from the free MD simulation illustrating the NOM (TNB) aggregation in the presence of Ca^{2+} ions. A long-lived cluster of three TNB molecules and three Ca^{2+} ions is observed. H_2O molecules and other ions are removed for clarity. Different TNB molecules are represented by different colors, whereas calcium ions are represented by blue spheres (not in scale). Ca4 is simultaneously coordinated by the 1C25 and 2C29 carboxilic groups, both in bidentate configuration, and by the 6C23 carboxilic group in a monodentate configuration. Ca2 is coordinated by 1C23 and 1C29 carboxilic groups of the same molecule in bidentate innersphere and outersphere coordination, respectively.

For table of contents only