PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Physical Chemistry Chemical Physics
CREATED USING THE RSC LaTeX PCCP ARTICLE TEMPLATE – SEE www.rsc.org/electronicfiles FOR DETAILS

Ab initio and Metadynamics Studies on the Role of Essential Functional Groups in Biomineralization of Calcium Carbonate and Environmental Situations

Moumita Saharay,*[a] R. James Kirkpatrick[b]

Received Xth XXXXXXXXXX 2013, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

The interactions of proteins, polysaccharides and other biomolecules with Ca^{2+}, CO_3^{2-}, and water are central to the understanding of biomineralization and crystallization of calcium carbonate ($CaCO_3$), and their association with the natural organic matter (NOM) in the environment. A molecular-level investigation of how such interactions and thermodynamic forces drive the nucleation and growth of crystalline $CaCO_3$ in living organisms remains elusive. This paper presents *ab initio* and metadynamics studies of the interactions of Ca^{2+}, CO_3^{2-}, and water with the essential amino acids/functional groups, *e.g.* arginine/NH_2^+, aspartate or glutamate/COO^-, aspartic or glutamic acid/COOH, and serine/OH, of protein/organic molecules that are likely to be critical to the biomineralization of $CaCO_3$. These functional groups were modeled as guanidinium (Gdm^+), acetate (AcO^-), acetic acid (AcOH), and ethanol (EtOH) molecules, respectively. The Gdm^+-Ca^{2+}-CO_3^{2-} and AcO^--Ca^{2+}-CO_3^{2-} systems were found to form stable ion-complexes irrespective of the presence of near neighbor water molecules. The strong electrostatic interactions of these functional groups with their counterions significantly affect the fundamental vibrational frequencies of the functional groups, mainly the NH_2 stretching (str.) and degenerate (deg.) scissors modes of Gdm^+ and -C=OO, CC, and CO str. modes of AcO^-. The free-energy calculations reveal that EtOH forms weakly bound molecular complexes with the Ca^{2+}-CO_3^{2-} ion pairs in water. However, the interaction strength of EtOH with crystalline $CaCO_3$ can increase significantly due to combined effect of H-bond and electron donor acceptor (EDA) type of interactions. These results indicate that -NH_2^+ and -COO^- bearing molecules serve as potential nucleation sites promoting crystallization of $CaCO_3$ phases while -OH bearing molecules are likely to control the morphology of the crystalline phases by attaching to the growing crystal surfaces.

The interactions of proteins, polysaccharides and other biomolecules with Ca^{2+}, CO_3^{2-}, and water molecules are central to understanding the biomineralization of $CaCO_3$ and are also fundamental to the interactions among naturally occurring organic matter, mineral surfaces, groundwater, and minerals [1–5]. $CaCO_3$ biomineralization is an essential process in the formation of the exoskeletons and other structural components of many living organisms, including crustaceans, algae, sea urchins, and avian eggshells. Crystalline $CaCO_3$ formed both biogenically and inorganically occurs as one of the three polymorphs, calcite, aragonite, and less commonly vaterite. The crystallization pathway for these phases depends greatly on the physical and chemical conditions [6–11]. The biologically produced phases have a wide range of morphologies and physical properties due to the influence of the proteins and enzymes present [12–15] and these properties can be utilized to develop 'smart' materials [7,16]. Proteins generally carry a net non-zero surface charge that facilitates interaction and binding with either positively charged Ca^{2+} or negatively charged CO_3^{2-}. Although it is not yet known how the binding of proteins to these ions controls nucleation, it has been speculated that the calcium-binding proteins delay nucleation and reduce the rate of crystal growth [17] by increasing the activation energy of the process.

Calcite is formed in systems with proteins that containThe interactions of proteins, polysaccharides and other biomolecules with Ca^{2+}, CO_3^{2-}, and water molecules are central to understanding the biomineralization of $CaCO_3$ and are also fundamental to the interactions among naturally occurring organic matter, mineral surfaces, groundwater, and minerals [1–5]. $CaCO_3$ biomineralization is an essential process in the formation of the exoskeletons and other structural components of many living organisms, including crustaceans, algae, sea urchins, and avian eggshells. Crystalline $CaCO_3$ formed both biogenically and inorganically occurs as one of the three

[a] *Center for Modelling Simulation and Design, University of Hyderabad, Hyderabad 500046, Andhra Pradesh, India; E-mail: saharaym@msu.edu*
[b] *College of Natural Science, Michigan State University, East Lansing, Michigan 48824, USA*

polymorphs, calcite, aragonite, and less commonly vaterite. The crystallization pathway for these phases depends greatly on the physical and chemical conditions[6–11]. The biologically produced phases have a wide range of morphologies and physical properties due to the influence of the proteins and enzymes present[12–15] and these properties can be utilized to develop 'smart' materials[7,16]. Proteins generally carry a net non-zero surface charge that facilitates interaction and binding with either positively charged Ca^{2+} or negatively charged CO_3^{2-}. Although it is not yet known how the binding of proteins to these ions controls nucleation, it has been speculated that the calcium-binding proteins delay nucleation and reduce the rate of crystal growth[17] by increasing the activation energy of the process.

Calcite is formed in systems with proteins that contain principally aspartic acid/aspartate[18,19] or have negatively charged regions on their surface[11]. The interaction of individual amino acid residues with Ca^{2+}-CO_3^{2-} ion pairs are also important in biomimetics, where additives with selective charged functional groups are used to control the crystallization pathway [20] and crystal morphologies[21]. For example, poly(aspartic acid) and poly(acrylic acid), both contain COOH functional groups and stabilize a polymer-induced liquid phase (PILP) prior to the formation of amorphous calcium carbonate (ACC)[22]. In contrast, positively charged additives[21] with amine (NH_2^+) groups induce phase separation driven by the interaction with negatively charged carbonate ions forming hydrated charge-separated (Ca^{2+}/additive/CO_3^{2-}) droplets. This phase separation induces dramatic changes in the crystal morphologies of $CaCO_3$. In all these processes, charge-dipole and dipole-dipole interactions play a major role in the association and dissociation of ions from the protein surface, synthetic polymers, or liquid organic additives. Thus, understanding the binding affinities and interaction strengths of amino acid residues with Ca^{2+} and CO_3^{2-} is an important step in understanding and predicting the crystallization behavior of biogenic $CaCO_3$ phases.

Despite great progress in experimental studies of the nucleation and growth of biogenic and natural $CaCO_3$, molecular scale understanding of the influence of biomolecules on the crystallization pathways has remained elusive. Computational modeling studies have the potential to provide significant insight into this question. Recent studies using classical simulation methods have focused principally on the interaction and adsorption of proteins[23], polysaccharides[15], and synthetic peptides[24] on the surfaces of crystalline $CaCO_3$ phases. Peter et al.,[24] for instance, studied the adsorption of polystyrene sulfonate (PSS) on calcite using classical molecular dynamics (MD) simulations. PSS is an additive to control crystal growth in vitro. They observed that the binding of the acidic sulfonate group on the polar (001) calcite surface is much stronger than on the non-polar (104) surface due to the dominant interaction between sulfonate functional groups and the positively charged Ca^{2+} on the (001) surface. In addition to this charge-dipole interaction, solvent-mediated interaction between the acidic functional groups and Ca^{2+} is also crucial in this surface binding, with free energy calculations showing that the binding energies for direct and solvent-mediated adsorption on (001) surfaces are essentially the same. In contrast, on the non-polar (104) surface, solvent separated acidic groups form strong hydrogen bonds with the ordered water molecules on the surface.

Freeman *et al.*[23] studied the binding of the eggshell protein Ovocleidin-17 on calcite surfaces. This protein has a net +7 charge at neutral pH. These results show that the residues alanine (functional group: CH_3), arginine (functional group: NH_2^+), aspartic acid (functional group: COOH), and serine (functional group: OH^-) are able to help bind the protein on the non-polar (104) calcite surface and that the interaction of arginine is dominant. The strong interaction between the calcite surface and water molecules results in the formation of an ordered water layer on the surface, and the most favorable binding of the protein is achieved by solvent-mediated interactions that minimize displacement of water molecules from the surface layer.

In natural and anthropogenic environments, aqueous fluids rich in Ca^{2+} and carbonate species are common in near surface environments and deeper in the earth's crust. These fluids are the medium from which $CaCO_3$ minerals crystallize, and as in biomineralization, bio- and organic- molecules appear to play a significant role. Ca^{2+} binding is also thought to be central to the structure and behavior of NOM, its interaction with mineral surfaces, and even fouling of water purification membranes[25–27]. The relevant functional groups[2,28] are often the same as in biomineralization, and the molecular scale interactions are also incompletely understood.

Modeling results concerning calcium carbonate biomineralization published to date have focused on the interaction of macromolecules and preexisting calcite surfaces, but to understand the effects of these molecules on $CaCO_3$ nucleation and growth it is also necessary to study the interaction of the important amino acid residues with Ca^{2+} and CO_3^{2-} in aqueous solution. Ca^{2+} and CO_3^{2-} form strong ion pairs in solution[29–31], and it is thus also important to understand these interactions not just with the individual ions but also with both ions present. Similarly, the published computational studies of the role of Ca^{2+} in NOM structure and binding have not included CO_3^{2-} or HCO_3^-, despite the near ubiquitous presence of carbonate species in natural waters.

We present here a computational study of the interaction of Ca^{2+} and CO_3^{2-} with functional groups relevant to $CaCO_3$ biomineralization and NOM using DFT, classical molecular dynamics (MD) and metadynamics methods to evaluate the structural environments, energetics of these interactions in small molecular clusters, changes in the vibrational frequen-

cies due to ion-association, and the effects of solvation of the clusters in bulk water. The results illustrate quantitatively the role of solvating H_2O molecules in changing the structural arrangements of the gas phase clusters.

There have been several previously published computational studies relevant to our results. Hamm *et al.*[32] performed classical MD simulations of hydrated Ca^{2+} in the presence of negatively charged aspartate (Asp) using CHARMM[33] force field parameters. They showed that the strong interaction between these species triggers dehydration of Ca^{2+} by reducing the energy barrier of Ca^{2+} dehydration. Nielsen *et al.*[34] considers this to be the rate-limiting step for the calcium carbonate nucleation and growth. Gale *et al.*[35] used modified CHARMM parameters to perform similar calculations to study the free energy landscape for the binding of Asp/Aspartic acid to the surfaces of amorphous and crystalline $CaCO_3$. Freeman *et al.*[23,36] used AMBER[37] parameter set in their study of Ovocleidin-17 on calcite surfaces, although this parameter set underestimates the hydration free energies of Ca^{2+} and CO_3^{2-}.

For the classical MD and metadynamics calculations here, we use CHARMM parameters for Ca^{2+}, CO_3^{2-}, and the organic molecules and TIP3P[38] for water molecules. The functional groups modeled are NH_2^+ (guanidinium, CN_3H_6, Gdm^+), COO- (acetate, AcO^-), COOH (acetic acid, AcOH) and OH (ethanol, EtOH). These groups are relevant models for arginine, aspartate/glutamate, aspartic acid/glutamic acid, and serine, respectively and are common functional groups in NOM.

1 Simulation Details

To study amino acid-$CaCO_3$-water interactions in the gas phase, we performed quantum-chemical calculations using the Gaussian09[39] suite of programs at the B3LYP/6-311+G(d,p) level. The C_α atoms of alanine, aspartic/glutamic acid, and serine were modeled as methyl groups. We will refer to these residues as ethanol (EtOH), acetate (AcO^-) and acetic acid (AcOH). The arginine side chain was modeled as a positively charged guanidinium ion (Gdm^+). Vibrational analysis of the optimized geometries do not show any imaginary modes that represent a stable geometry.

Classical MD simulations of aqueous solutions containing a formula unit of $CaCO_3$ and a model residue, Re (Re = EtOH, AcO^-, AcOH, Gdm^+), were carried out using the NAMD software[40] package, the CHARMM 27 all-atom force-field[33] parameters, and the TIP3P water model[38]. The simulations were performed in the NPT ensemble at 310 K and 1 atm pressure. The system pressure was controlled using a Langevin piston with a period of 100 fs. The temperature was controlled by a Langevin thermostat. The cubic box dimensions of the EtOH-, AcO^--, AcOH-, and Gdm^+-$CaCO_3$ systems were 29.1, 22.8, 22.3, and 19.1 $Å^3$, respectively. The AcO^- and Gdm^+ systems were neutralized by one Na^+ and one Cl^-, respectively. The number of water molecules in these systems were 826, 400, 285, and 233, respectively. The non-bonded interactions were smoothed using a switching function between 10 and 12Å. Electrostatic interactions were calculated using the particle mesh Ewald (PME) method with a distance cutoff of 12Å. Each system was run for total 4 ns including 2 ns of initial equilibration.

The metadynamics methods implemented in the NAMD software package were used to study the free energy profiles for the association of the ion pairs in bulk solution using collective variable (CV) module. Each CV is the measure of distance between a particular pair of atoms. The CVs for each system are described in the respective sections of this paper. For these calculations, each system was run for total 22 ns including 2 ns of initial equilibration. During the simulation, Gaussian hills of width 2^o were added at a regular interval of 100 fs.

2 Validation of Force Field Parameters

The applicability of MD models involving organic and inorganic species depends greatly on the quality of the force field parameters. To validate the suitability of the CHARMM force field parameters to study the interactions between organic molecules and $CaCO_3$ we calculated the hydration free energies of the Ca^{2+} and CO_3^{2-} and the local hydration shell around Ca^{2+} using the NAMD software package. These values were compared with the existing theoretical and experimental values. The hydration free energy of a particular ion was calculated by the Free-Energy-perturbation (FEP) method, implemented in NAMD package, where the interaction strength of the ion with the was gradually increased or decreased. For example, if the interaction strength (λ) is 0 then the ion is considered to be non interacting and is placed infinitely far apart from the solvent molecules. In contrast, $\lambda = 1$ indicates favorable interaction between the ion and the solvent. Thus, the potential energy of a system containing solute (here, the ion) and solvent is given by

$$E_{Total} = E_{solute-solute} + E_{solvent-solvent} + \lambda E_{solute-solvent} \quad (1)$$

The change in free energy between two λ states (i, i+1) is

$$\Delta G_i = -\frac{1}{\beta} ln < exp[-\beta(E_{\lambda_{i+1}} - E_{\lambda_i})] > \quad (2)$$

and the hydration free energy is

$$\Delta G_{hyd} = \sum_{i=1}^{N} \Delta G_i \quad (3)$$

where N=total number of λ points, $\lambda_1 = 0$ and $\lambda_N = 1$. During the FEP calculations, the λ value was varied in steps of 0.5 for both the forward (λ= 0→ 1) and reverse (λ= 1→ 0) processes. Each model system was generated by solvating an ion (Ca^{2+} or CO_3^{2-}) in a cubic water box of dimension 30Å and was equilibrated for 1 ns before initializing the FEP calculations. At each λ value, the equilibrium run length was 60 ps and the production run length was 1.4 ns. To overcome the singularity issue at the terminal windows, λ=0 and 1, a soft-core potential was used.

Table 1 lists the hydration free energies of the ions from the existing literature and the results from our calculations. The ΔG_{hyd} of Ca^{2+} and CO_3^{2-} calculated using the CHARMM parameter set are -1384 kJ/mol and -1247.34 kJ/mol, respectively. Compared to the experimental values, our simulation results yield a slightly lower ΔG_{hyd}. Previous simulation studies of calcium carbonate biomineralization using the AMBER[37] force field parameters for the biomolecules and Pavese *et al.*'s parameters[44] for $CaCO_3$ show much lower values of the ΔG_{hyd} and underestimate the solubilities of both Ca^{2+} and CO_3^{2-}. In a recent publication[42], Gale *et al.* provided a new set of force field parameters for $CaCO_3$ that yield better agreement with the experimental ΔG_{hyd} values. However, the compatibility of these parameters with the existing force field parameters for biomolecules are not yet verified. Moreover, mixing two different types of force fields may lead to systematic errors in the calculations. Thus, we chose to use CHARMM parameters for both organic and inorganic species.

Table 1 Hydration free energies (kJ mol^{-1}) of Ca^{2+} and CO_3^{2-}

Ion type	Theoretical	Experimental
Ca^{2+}	-1384. (this work) -1333.[36,42] -1447.[42]	-1446.4[41]
CO_3^{2-}	-1247.34 (this work) -1175.[36,42] -1301.[42]	-1314.[43]

To further test the applicability of the CHARMM parameters, we also studied the near neighbor hydration shell around Ca^{2+}. The radial distribution function of atom pair Ca,O_{water} (g_{Ca-O_W}, Figure 1) shows a nearest neighbor peak at 2.26Å, a minimum at 2.56Å, a second neighbor peak at 4.5Å, and a second minimum at 5.5Å. The running coordination number at the first minimum is 6.8. These values compare well with the X-ray diffraction studies[45] which show the position of the first minimum is at ∼2.5 Å and a corresponding running coordination number of ∼7. Based on the agreement between the calculated and experimental hydration free-energies and nearest neighbor distance and coordination number, we conclude that the force field parameters used here are suitable for our purposes.

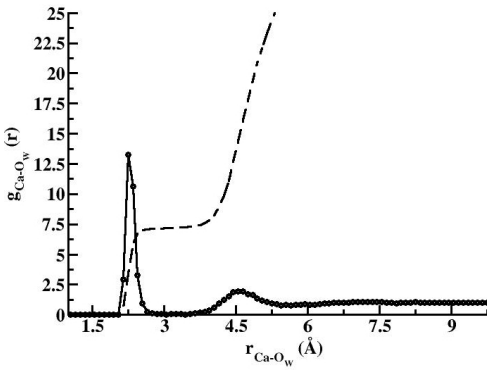


Fig. 1 Radial distribution function (solid line) and running coordination number (dashed line) for the atom pair Ca^{2+} and oxygen of water (O_W) molecules.

3 Results and Discussion

For each model residue, we first discuss the ground state properties of the gas-phase systems with and without neighboring water molecules. Then we discuss the solvation effects on these systems from the MD and metadynamics results.

3.1 Gdm^+

3.1.1 Gas-phase

The optimized gas phase structures of the Gdm^+ clusters show that H_2O has a significant effect on the equilibrium structural relationships among the Gdm^+, Ca^{2+} and CO_3^{2-} (Figure 2). For the system without H_2O, two NH_2 groups of the Gdm^+ are located directly perpendicular to the CO_3^{2-} ion forming two hydrogen bonds with two oxygens of CO_3^{2-} (O_{CO_3}) forming a 'twin nitrogen-twin oxygen' (TNTO) geometry[46–48] (Figure 2A). This is a common structural arrangement for Gdm^+ and COO^- containing residues. The Ca^{2+} is in bidentate coordination with two O_{CO_3} and is not directly coordinated to the Gdm^+. In general, this geometric relationship between Ca^{2+} and CO_3^{2-} is the most stable gas phase structure[29], and all the Gdm^+-containing clusters here have this geometry, irrespective of the number of near-neighbor water molecules. The strong interaction between Ca^{2+} and CO_3^{2-} results in changes in the lengths of the C-O bonds in CO_3^{2-} relative to the bare CO_3^{2-}. In isolated CO_3^{2-}, the C-O bond length is 1.31Å. This value increases to 1.32 and 1.36Å for the two O_{CO_3} coordinated to Ca^{2+} and decreases to 1.23Å for the O_{CO_3} not coordinated to Ca2+. The presence of the Ca^{2+} also causes the

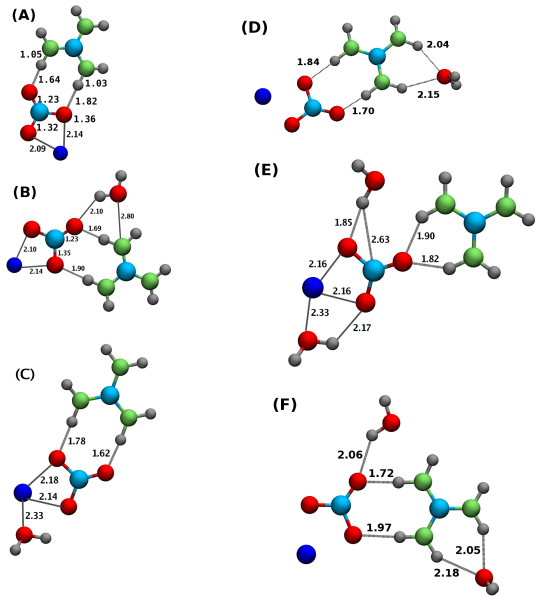


Fig. 2 Geometry optimized structures of Gdm^+ containing clusters with (A) Ca^{2+} and CO_3^{2-}, (B,C,D) Ca^{2+}, CO_3^{2-} and one H_2O, (E,F) Ca^{2+}, CO_3^{2-} and two H_2O. Color scheme: carbon (cyan), nitrogen (green), oxygen (red), calcium (blue), hydrogen (gray).

Table 2 Binding energies of the optimized ion-complexes.

Superscript on left indicates the corresponding Figure number for the optimized geometry of the complex. The energy difference, ΔE, for $CaCO_3$ containing complex ($Re[CaCO_3][H_2O]_n$) is ΔE = E(complex) - E(Re) - E($CaCO_3$) - nE(H_2O) and that of Ca^{2+} / CO_3^{2-} containing complex is ΔE = E(complex) - E(Re) - E(Ca^{2+}/CO_3^{2-}) - nE(H_2O), where Re and n are the model functional group and number of water molecules, respectively.

Complex	ΔE (kcal/mol)
$^{2a}Gdm[CaCO_3]^+$	-46.24
$^{2b}Gdm[CaCO_3][H_2O]^+$	-55.50
$^{2c}Gdm[CaCO_3][H_2O]^+$	-75.30
$^{2d}Gdm[CaCO_3][H_2O]^+$	-59.45
$^{2e}Gdm[CaCO_3][2H_2O]^+$	-85.43
$^{2f}Gdm[CaCO_3][2H_2O]^+$	-67.90
$^{6a}AcO[Ca]^+$	-321.99
$^{6b}AcO[CaCO_3]^-$	-89.22
$^{6d}AcO[CaCO_3][H_2O]^-$	-107.92
$^{6e}AcOH[Ca]^{2+}$	-83.38
$^{6f}AcOH[Ca]^{2+}[H_2O]$	-109.87
$^{13a}EtOH[CaCO_3]$	-8.61
$^{13b}EtOH[CaCO_3][H_2O]$	-45.63

H-bonds donated to the O_{CO_3} by the Gdm+ to have different lengths.

In systems with one water molecule, the Gdm^+ and CO_3^{2-} are coplanar, and the TNTO geometry between them remains present. There are three optimized structures with one water molecule in which the H_2O interacts with, (i) Gdm^+ and CO_3^{2-} (Figure 2B), (ii) Ca^{2+} (Figure 2C), and (iii) Gdm^+ (Figure 2D). The energetically most stable arrangement (Figure 2C, Table 2) is the one in which H_2O is coordinated to a Ca^{2+}. The positions of the water molecule in these clusters are very different from that in an isolated Gdm^+-water complex[49], in which the oxygen of water (O_W) lies in the plane of Gdm^+ and forms two hydrogen bonds with the hydrogens of two adjacent NH_2^+ groups. Moreover, the distance between the Gdm^+ and CO_3^{2-} is significantly changed by the presence of a water molecule that interacts with both of these ions (Figure 2B) and in the situation where it interacts with one of them (Figure 2C,D). The closest distances between O_{CO_3} and H_{Gdm} in structure B increase by ∼0.12Å than the structure in Figure 2C. The strong H-bond interaction between Gdm^+ and H_2O in the structure in Figure 2D increases these distances even more by pulling the Gdm^+ away from CO_3^{2-}.

For systems with two water molecules, the minimum energy structure (Figure 2E) has one water that interacts with both CO_3^{2-} and Gdm^+ at essentially the same position as in the minimum energy one-water (Figure 2B) system and a second water molecule that donates an H-bond to the CO_3^{2-}. This structure does not show the TNTO geometry. Rather, the CO_3^{2-} is coordinated to the Gdm^+ by two H-bonds donated to a single O_{CO_3} from two different NH_2 groups. We will refer this arrangement as 'twin-nitrogen-single-oxygen' (TNSO). In the less stable complex with two water molecules (Figure 2F), Gdm^+ is in TNTO geometry with CO_3^{2-} and the second water molecule interacts only with Gdm^+ and forms two H-bonds with the NH_2 group. Thus, the position of the nearest neighbor water molecule significantly affects the strength of H-bond interaction between Gdm^+ and CO_3^{2-}. Systems with more water molecules could not be optimized. Our MD simulation of the bulk system shows that the association of more than two water molecules with the ion-complex requires an H-bond network among the neighboring water molecules.

3.1.2 Vibrational Spectra

The presence of near-neighbor water molecules has a significant effect on the vibrational modes of Gdm^+ (Figure 3). The NH stretching modes at 3592 cm^{-1} (symmetric str.) and 3714 cm^{-1} (asymmetric str.) and the NH_2 degenerate scissors modes at 1713 cm^{-1} are most sensitive to the presence of CO_3^{2-} in the TNTO geometry. The establishment of strong H-bonds shifts the str. modes of the participating NH bonds

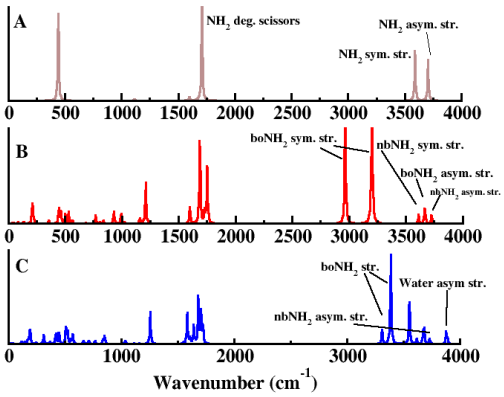


Fig. 3 Power spectra of (A) Gdm^+, (B) Gdm-$CaCO_3$, and (C) Gdm-$CaCO_3$-$2H_2O$ clusters.

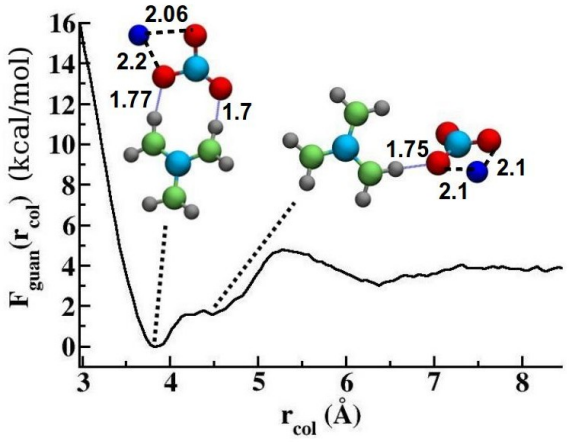


Fig. 4 Potential of mean force for the binding of Gdm^+ to $CaCO_3$ ion-pair.

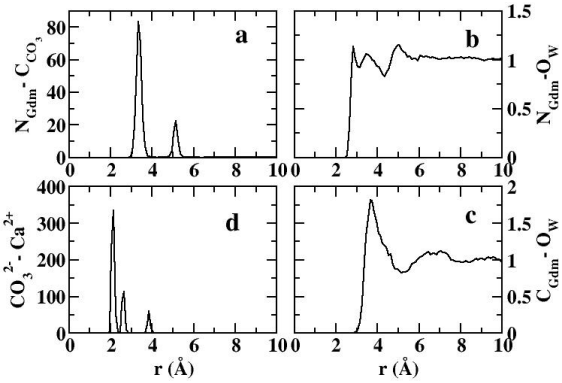


Fig. 5 Radial distribution functions for (a) nitrogen (N_{Gdm}) of Gdm^+ and carbon (C_{CO_3}) of CO_3, (b) N_{Gdm} and oxygen of water (O_W), (c) carbon (C_{Gdm}) of Gdm^+ and O_W, (d) CO_3^{2-} and Ca^{2+}.

to much lower frequencies that appear at 2965 and 3200 cm^{-1} for N-H...O-C distances of 1.64 and 1.8Å, respectively. However, this effect is less prominent for the non-donating NH bond. This restricted motion also splits the NH_2 deg. scissors modes into lower frequency modes. *Ab initio* studies[50] of solvated and isolated Gdm^+ show that the presence of the first hydration shell has a negligible effect on the vibrational modes of the ion. One of the reasons for this could be that the N-H (in Gdm^+) and O-H (in water) stretching frequencies are almost the same, and it is difficult to identify the effect of water that is H-bonded with the ion and with other water molecules. In the two-H_2O complex here, the transformation from TNTO geometry in the dry complex to TNSO geometry with two waters relaxes the NH vibration modes of the NH_2 groups that are bonded to the CO_3^{2-} and shifts them towards higher frequencies, although these modes are still at lower frequencies than the non-bonded NH_2 (Figure 3C). The deprotonated state of the arginine side chain is also important in this context. Braiman *et al.*[51] show that deprotonation of the arginine side chains will red-shift the CN stretching mode by ∼50-100 cm^{-1} relative to the protonated state. In a hydrophobic environment, the association of an arginine side chain with a counter-ion depends on the position and characteristics of the counter-ion and the strength of the H-bond[51].

3.1.3 Free Energies

The free energy profile ($F_{guano}(r_{col})$) of the Gdm^+-Ca-CO_3 ion-complex solvated in bulk water as a function of the distance between carbon atoms of Gdm^+ and CO_3^{2-} (r_{col}) contains a global minimum at 3.83Å that is 4 kcal/mol more stable than the non-interacting, dissociated state at 8Å and a secondary minimum at 4.5Å that is 3 kcal/mol more stable than the dissociated state (Figure 4). The arrangement of the Gdm^+, Ca^{2+} and CO_3^{2-} at the global energy minimum is generally similar to that of gas phase complex without water molecules (TNTO) except that the Gdm^+ and CO_3^{2-} are not coplanar. At this minimum, the CO_3^{2-} receives two H-bonds from the Gdm^+, and the Ca^{2+} is in bidentate coordination with the CO_3^{2-} throughout the simulation due to strong electrostatic interaction with a binding energy of nearly 8.84 kcal/mol[52]. At the secondary minimum, the CO_3^{2-} is almost perpendicular to the Gdm^+ and this twist in the θ angle leads to the formation of only one H-bond while maintaining the same H_{Gdm}-O_{CO_3} distance as at the global minimum.

3.1.4 Bulk Solution

The intermolecular radial distribution functions (RDFs) obtained from the classical MD simulations of Gdm^+, Ca^{2+}, and

CO_3^{2-} solvated in bulk water show that the Gdm^+ and CO_3^{2-} retain the TNTO structural arrangement (Figure 5a) and that the Ca^{2+} is mostly in bidendate coordination with CO_3^{2-} (Figure 5d). Previous MD calculations[53] of Gdm^+s and CO_3^{2-}s solvated in bulk water show that even at high concentration of Ca^{2+} and CO_3^{2-}, TNTO is the preferred geometry. As observed in our free-energy calculations, the strong H-bond interactions between Gdm^+ and CO_3^{2-} inhibits the dissociation of the Gdm-CO_3 ion-complex. This result indicates that guanidinium groups or, in general, arginine residues on protein surfaces and amine groups in NOM help bio- and organic-molecules recognize and bind to the CO_3^{2-} groups on calcium carbonate surfaces or CO_3^{2-} in solution. The RDF for the N_{Gdm}-C_{CO_3} atom pair (Figure 5a) shows two isolated peaks at 3.4 and 5.1Å, where the first peak corresponds to the nitrogen atoms H-bonded to CO_3^{2-} and the second peak corresponds to the unpaired nitrogen atoms. Compared to the gas-phase clusters, the shortest N-C distance is increased by ∼0.7Å in solution. This increase is due to the presence of the hydration shell around both Gdm^+ and CO_3^{2-} which weakens the H-bond strength. The RDF for the N_{Gdm}-O_W atom pair (Figure 5b) shows three prominent peaks at 2.8, 3.5, and 5.0Å. The running coordination numbers at each of these distances are 1.0, 3.6, and 15, respectively. This result shows that even in the bulk solution, the Gdm^+-$CaCO_3$ ion-complex can accommodate a maximum of one water molecule in the first hydration shell. The RDF for the C_{Gdm}-O_W atom pair (Figure 5c) shows a peak at 3.7Å and a broad peak at 6.9Å. Analysis of the arrangements related to these peaks indicate that there are two types of hydration waters around Gdm: (1) in-plane water molecules that receive H-bonds from Gdm^+, and (2) out-of-plane water molecules in the plane parallel to the plane of Gdm^+ interacting with the lone-electron pair of the N_{Gdm} atoms.

Classical MD study of the hydration of Gdm^+, Soetens *et al.*[49] used both polarizable and non-polarizable water molecules and found no dramatic differences in the solvent structure around the ions except that the number of nearest neighbor water molecules is more with the polarizable model. The structural properties extracted from our results match well with those of Soetens *et al.* using a polarizable water model. Thus, we believe that our results with the non-polarizable TIP3P water model are useful in understanding biomineralization.

3.2 AcO^-

3.2.1 Gas phase

The geometry optimized structures of gas phase systems containing AcO^-, Ca^{2+}, CO_3^{2-}, and H_2O show that in all cases coordination to the COO^- group occurs through the Ca^{2+} and

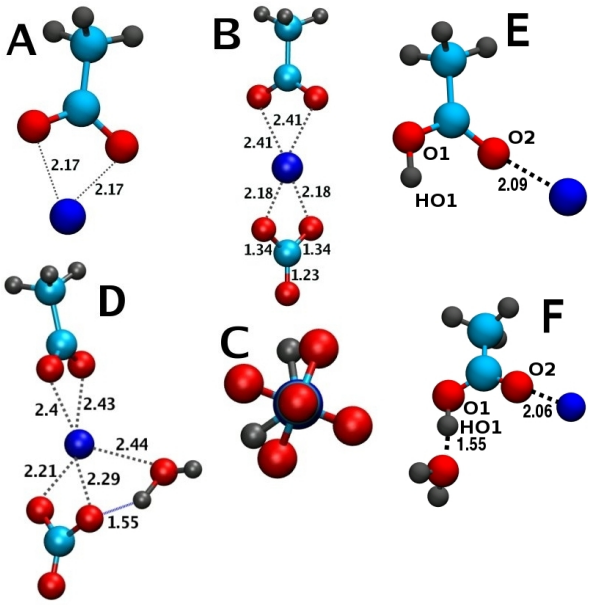


Fig. 6 Geometry optimized structures of (A,B,C,D) AcO^- and (E,F) AcOH containing clusters. Front and top views of the AcO^--Ca^{2+}-CO_3^{2-} complexes are shown in panels B and C, respectively. Color scheme: Carbon (cyan), Oxygen (red), Calcium (blue), hydrogen (gray).

that the presence of one water molecule changes the structure significantly (Figure 6). For the AcO^--Ca^{2+} ion complex (Figure 6A), Ca^{2+} is coplanar with the COO^- group and is equidistant from the two carboxyl oxygens at 2.17Å. The binding energies of the ion complexes are listed in Table 2. For the AcO^--Ca^{2+}-CO_3^{2-} complex (Figure 6B) Ca^{2+} bridges the two anions via coulombic interaction, with the Ca^{2+}-O_{CO_3} bond distances 0.23Å shorter than for Ca^{2+}-O_{AcO}. The COO planes of the AcO^- and CO_3^{2-} are perpendicular to each other (Figure 6C). The Ca^{2+}-O_{AcO} distances are larger than for the AcO^--Ca^{2+} complex (2.41Å vs. 2.17Å) due to strong interaction between Ca^{2+} and CO_3^{2-}.

As for the Gdm^+ clusters, in the AcO^--Ca^{2+}-CO_3^{2-} complex with one H_2O, the water molecule is coordinated to the Ca^{2+}. It donates an H-bond to O_{CO_3} and is not coordinated to the AcO^-, demonstrating the more hydrophilic nature of CO_3^{2-}-(Figure 6D) relative to COO^- of AcO^-. The Ca^{2+} in this complex is coordinated by a total of 5 oxygen atoms. We modelled several initial structures of the AcO^--Ca^{2+}-CO_3^{2-} complex with more than one water molecules but could never achieve a stable geometry-optimized structure other than the one in Figure 6D. This observation is in line with our MD results which show that in bulk water, the Ca^{2+} is in bidentate coordination by AcO^- and there is a maximum of one water molecule in the first hydration shell. Recent infrared spec-

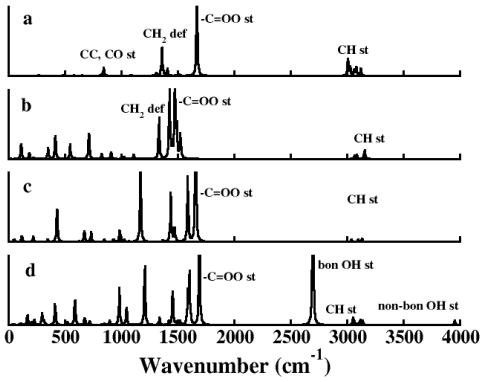


Fig. 7 Power spectra of (a) AcO^-, (b) AcO^--Ca^{2+}, (c) AcO^--$CaCO_3$, (d) AcO^--$CaCO_3$-H_2O

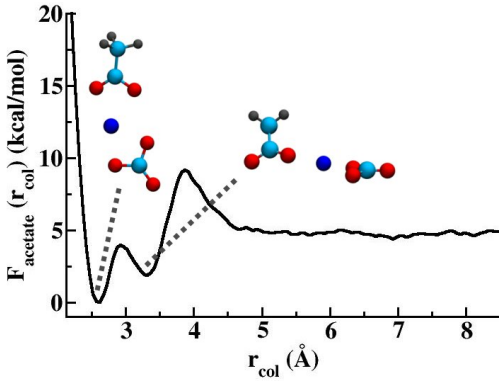


Fig. 8 Potential of mean force for the binding of AcO^- with Ca,CO_3 ion-pair. The most probable geometries of these ions at each energy minima are shown with ball-and-stick representation. Color scheme: Carbon (cyan), Oxygen (red), Calcium (blue).

troscopy experiments[54] show that even when the counter ion is Na^+ rather than Ca^{2+}, the hydration number per -COO^- is 1.0. Thus, in comparison to solvated AcO^- without any counter ion[55] where the O_{AcO} is coordinated by 3.4 O_W (oxygen of water), the presence of Ca^{2+} reduces the water coordination around the carboxyl group significantly. This inner sphere bridging of the carboxylic group and the CO_3^{2-} by Ca^{2+} is likely to play an important role in adsorption of bio- and organic molecules on the $CaCO_3$ surfaces at neutral and basic pHs and in helping these molecules structure the ions in solution during nucleation.

3.2.2 Vibrational Spectra

The calculated power spectra of isolated AcO^- and the AcO^--Ca^{2+}, AcO^--Ca^{2+}-CO_3^{2-}, and AcO^--Ca^{2+}-CO_3-H_2O complexes are significantly different (Figure 7). The -C=OO stretching mode at 1652 cm^{-1} and CC, CO stretching modes at 856 cm^{-1} are most sensitive to the presence of Ca^{2+} in the AcO^--Ca^{2+} complex (Figure 7b). The strong interaction in the ion pair splits and shifts the -C=OO stretching modes to 1477 and 1428 cm^{-1}, which are much lower frequencies than for isolated AcO^- (1664 cm^{-1}) (Figure 7a). There are similar changes[54] in a comparable system containing Na^+. This effect is however less prominent in the presence of CO_3^{2-}, because the carbonate significantly weakens the AcO^--Ca^{2+} binding. In the AcO^--$CaCO_3$-H_2O complex, the formation of strong H-bond between H_2O and CO_3^{2-} red shifts the OH stretching mode of the bonded OH (Figure 7d). Because the H_2O does not directly interact with AcO^-, the effect on the AcO^- modes is small. Previous IR spectroscopic[56,57] and theoretical[57] studies have shown that in bulk AcO^-/AcOH solution, there can be at least three types of hydrated species containing acetic acid with one or two nearest neighbor (NN) H_2O molecules and an acetate ion with two NN H_2O. In systems with protonated AcOH the mode at ~1700 cm^{-1} appears due to a Fermi resonance between the CO stretching mode and the first overtone of the CC stretching (885 cm^{-1}) mode[58]. The splitting of this mode in our AcO^--Ca^{2+} complex may be due to the absence of the Fermi resonance.

3.2.3 Free-energy Calculations

The free energy profile for the solvated AcO^--Ca^{2+}-CO_3^{2-} complex as a function of distance between C_{COO-} of AcO^- and Ca^{2+} (r_{col}) obtained from the metadynamics calculations shows a global energy minimum at r_{col}=2.62Å, a secondary minimum at 3.32Å that is 1.9 kcal/mol less stable, a transition state at 3.8Å, and a total energy barrier from the global minimum to the dissociated state of 9 kcal/mol (Figure 8). The energy barrier between the two minima is 3.9 kcal/mol. In the global minimum energy configuration, Ca^{2+} is in bidentate coordination with both anions, as in the gas phase structure. At the secondary energy minimum, Ca^{2+} is in monodentate coordination by the AcO^- and in bidentate coordination by O_{CO_3}. The high affinity of COO groups for Ca^{2+} and their selective binding is in agreement with the observation that acetate binding sites are in general Ca^{2+} philic. A recent study[28] of the complexation of Ca^{2+} with the carboxylic groups of a model molecule of NOM in the absence of CO_3^{2-} showed a similar free energy profile. This study also showed only small differences in the free energy profiles for carboxylic side chains and an isolated acetate group, confirming that the interactions between -COO^- groups and Ca^{2+} are strong irrespective of their end groups and the presence of other ions.

For a more detailed structural analysis of the AcO^--Ca^{2+}-CO_3^{2-} complex in solution, we define two angles, θ and ϕ

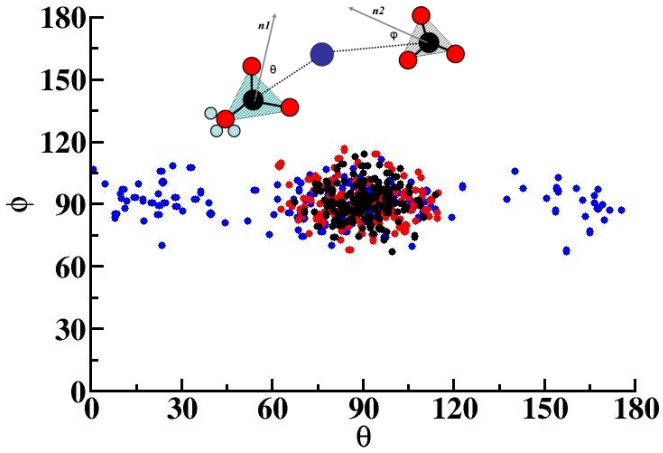


Fig. 9 Variation of θ and ϕ angles in metadynamics simulations at r_{col}=2.62 (the global energy minimum; black dots), 3.32 (secondary energy minimum; red dots), and 3.8Å (transition state; blue dots). Scheme: Definition of θ and ϕ angles. These two angles are between the C-Ca vector and the normal (n1 and n2 vectors) to the COO plane (shaded regions) for the CO_3^{2-} and AcO^- groups, respectively.

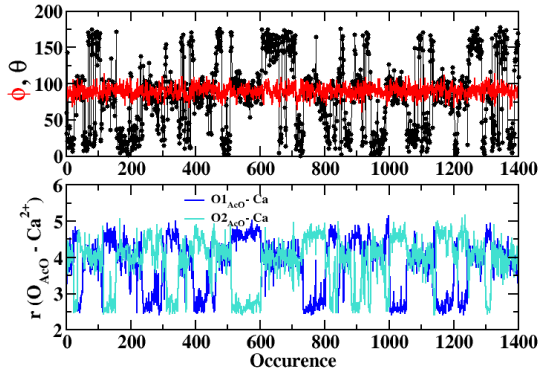


Fig. 10 Distribution of θ and ϕ angles (upper panel) and the corresponding O_{AcO}-Ca distances (lower panel) at the transition state of the free-energy profile. Atom-numbering is shown in Figure 6.

(Figure 9), and investigate their variation in the metadynamics simulations at r_{col}=2.62 (the global energy minimum), 3.32 (the secondary energy minimum), and 3.8Å (the transition state) of distance between C_{COO-} of AcO^- and Ca^{2+} (r_{col}) obtained from the metadynamics calculations shows a global energy minimum at r_{col}=2.62Å, a secondary minimum at 3.32Å that is 1.9 kcal/mol less stable, a transition state at 3.8Å, and a total energy barrier from the global minimum to the dissociated state of 9 kcal/mol (Figure 8). The energy barrier between the two minima is 3.9 kcal/mol. In the global minimum energy configuration, Ca^{2+} is in bidentate coordination with both anions, as in the gas phase structure. At the secondary energy minimum, Ca^{2+} is in monodentate coordination by the AcO^- and in bidentate coordination by O_{CO_3}. The high affinity of COO groups for Ca^{2+} and their selective binding is in agreement with the observation that acetate binding sites are in general Ca^{2+} philic. A recent study[28] of the complexation of Ca^{2+} with the carboxylic groups of a model molecule of NOM in the absence of CO_3^{2-} showed a similar free energy profile. This study also showed only small differences in the free energy profiles for carboxylic side chains and an isolated acetate group, confirming that the interactions between -COO^- groups and Ca^{2+} are strong irrespective of their end groups and the presence of other ions.

For a more detailed structural analysis of the AcO^--Ca^{2+}-CO_3^{2-} complex in solution, we define two angles, θ and ϕ (Figure 9), and investigate their variation in the metadynamics simulations at r_{col}=2.62 (the global energy minimum), 3.32 (the secondary energy minimum), and 3.8Å (the transition state). These two angles are between the C-Ca vector and the normal to the COO plane for the CO_3^{2-} and AcO^- groups, respectively. At the global and secondary energy minima, both the θ, ϕ distributions have maxima at 90^o with a distribution between 60 and 120^o. This geometry indicates that on average Ca^{2+} is co-planar with both AcO^- and CO_3^{2-}. These structures are similar to that of the gas phase complex. In the transition state configuration (Figure 9), the average ϕ value is 90^o, but there is a broad distribution of θ angles with maxima near 90 and 15^o. The distribution of distances between the carboxylic oxygens of AcO^- (O_{AcO}) and Ca^{2+} (Figure 10) shows that the transition state is composed of two structural arrangements. One is an in-plane monodentate structure in which the shortest O_{AcO}-Ca^{2+} distance is 2.5Å and θ=90^o. The other is an out-of-plane bidentate structure in which Ca^{2+} is equidistant from the O_{AcO} at ∼4Å and θ=15^o (Figure 10). At the transition state, the structural arrangement of the ions oscillates rapidly between these bidentate and monodentate geometries.

Throughout the metadynamics simulation, Ca^{2+} remains in in-plane bidentate coordination with the CO_3^{2-}. The 9 kcal/mol energy penalty for the transition from in-plane (at the global minima) to out-of-plane (at the transition state) geometry is much higher than the thermal energy at 310K (0.6 kcal/mol). Thus, the dissociation of this ion-complex at ambientThese two angles are between the C-Ca vector and the normal to the COO plane for the CO_3^{2-} and AcO^- groups, respectively. At the global and secondary energy minima, both the θ, ϕ distributions have maxima at 90^o with a distribution between 60 and 120^o. This geometry indicates that on average Ca^{2+} is co-planar with both AcO^- and CO_3^{2-}. These

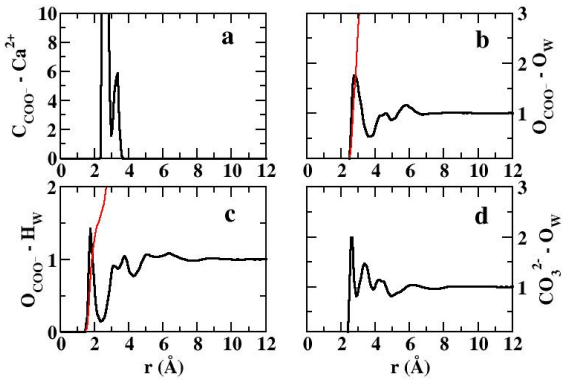


Fig. 11 Radial distribution functions for (a) carbon (C_{COO^-}) of AcO^- and Ca^{2+}, (b) oxygen (O_{COO^-}) of AcO^- and oxygen of water (O_W), (c) O_{COO^-} and H_W, (d) CO_3^{2-} and O_W.

structures are similar to that of the gas phase complex. In the transition state configuration (Figure 9), the average ϕ value is 90^o, but there is a broad distribution of θ angles with maxima near 90 and 15^o. The distribution of distances between the carboxylic oxygens of AcO^- (O_{AcO}) and Ca^{2+} (Figure 10) shows that the transition state is composed of two structural arrangements. One is an in-plane monodentate structure in which the shortest O_{AcO}-Ca^{2+} distance is 2.5Å and $\theta=90^o$. The other is an out-of-plane bidentate structure in which Ca^{2+} is equidistant from the O_{AcO} at ∼4Å and $\theta=15^o$ (Figure 10). At the transition state, the structural arrangement of the ions oscillates rapidly between these bidentate and monodentate geometries.

Throughout the metadynamics simulation, Ca^{2+} remains in in-plane bidentate coordination with the CO_3^{2-}. The 9 kcal/mol energy penalty for the transition from in-plane (at the global minima) to out-of-plane (at the transition state) geometry is much higher than the thermal energy at 310K (0.6 kcal/mol). Thus, the dissociation of this ion-complex at ambient temperature is a rare event, and binding of Ca^{2+}, CO_3^{2-} ion pairs to carboxylic groups may play an important role in facilitating calcium carbonate nucleation and growth.

3.2.4 Bulk Solution

The RDFs obtained from the classical MD simulations of the bulk solution of AcO^- show features that correspond to the global and secondary energy minima in the metadynamics calculations (Figure 11). The RDF for the C_{COO^-} and Ca^{2+} pairs (Figure 11a) shows two peaks at 2.55 and 3.3 Å that correspond to the global and secondary energy minima in the free-energy profile, respectively (Figure 8). The non-zero probability of the RDF in the range of 2.4 and 3.6 Å indicates the presence of a strongly bound undissociated ion complex throughout the molecular dynamics trajectory. The O_{COO^-}-O_W RDF (Figure 11b) shows a well-defined first hydration shell around AcO^- with a nearest neighbor coordination of 5 water molecules. However, analysis of the O_{AcO}-H_W RDF (Figure 11c) indicates the formation of only one H-bond between AcO^- and solvent water, in agreement with our gas-phase calculations in which the geometry optimized cluster of AcO^--$CaCO_3$-H_2O has only one water molecule H-bonded with AcO^-.

3.3 AcOH

3.3.1 Gas phase

For AcOH, protonation of the carboxylic group causes the interactions between Ca^{2+} and CO_3^{2-} to be drastically different than for deprotonated AcO^-. For the AcOH-Ca^{2+} cluster (Figure 6E), the geometry optimized structure has Ca^{2+} in monodentate coordination by the O_{AcOH} and colinear with the C=O bond at 2.09Å from the oxygen. The binding energy for this ion-complex is -83.38 kcal/mol. For the AcOH-Ca^{2+}-H_2O system , the only stable structure has the Ca^{2+} in monodentate coordination with the O-atom and the H_2O receiving an H-bond from the OH group (Figure 6F). The H_2O molecule does not coordinate the Ca^{2+}. The binding energy of this complex is -109.87 kcal/mol. For the AcOH-Ca^{2+}-CO_3^{2-} system there is no stable, energy minimized structure. Instead, the proton from the hydroxyl group of AcOH readily migrates to the carbonate forming a bicarbonate ion. Several initial structures were modeled by placing carbonate ion at different positions around the optimized structure of the AcOH-Ca^{2+} complex, but in all cases the proton migrates to the CO_3^{2-}. This result is expected, since the pKa of carboxylic groups are typically near 4, and H_2CO_3 is the stable carbonate species under these conditions. This result is also in line with experimental results that show rapid decomposition of $CaCO_3$ particles that are interacting with acetic acid[59–62].

Our results, thus, show that at low pHs or in a water-deficient region on a protein surface, aspartic or glutamic acids are likely to associate principally with Ca^{2+} rather than CO_3^{2-} [63]. Consequently, a large number of protonated side chains on a protein surface may help the protein adsorb onto the positively charged calcite (001) surface through direct interaction, whereas its adsorption on the neutral surface (104) may be due to solvent mediated interactions.

3.3.2 Free Energy

The free energy for the AcOH-Ca^{2+}-CO_3^{2-} system as a function of distance between the Cα atom of AcOH and C_{CO_3} rises sharply for low values of r_{col} where the planes of AcOH and CO_3^{2-} are almost parallel and the carbon and oxygen atoms of these species are perpendicular to the like atoms (Figure 12).

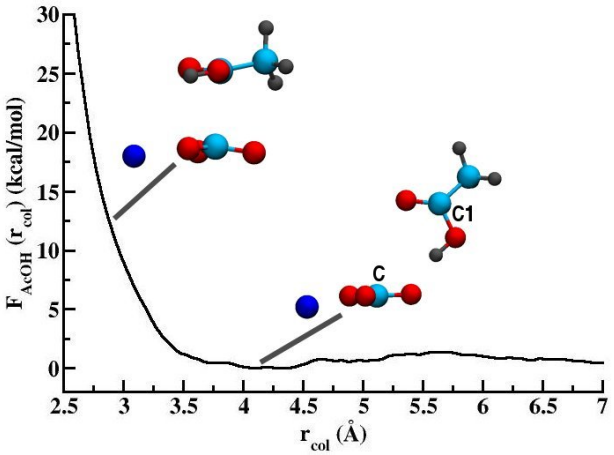


Fig. 12 Potential of mean force for the binding of AcOH with Ca,CO_3 ion-pair. The most probable geometries of these ions and molecule at each energy minima are shown with ball-and-stick representation. Color scheme: Carbon (cyan), Oxygen (red), Calcium (blue).

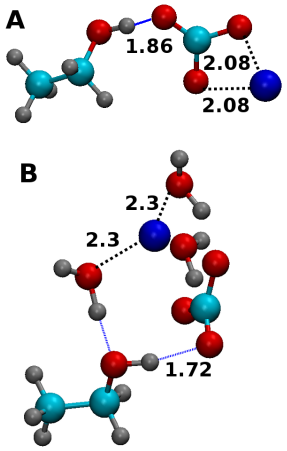


Fig. 13 Geometry optimized structures of EtOH containing clusters. Color scheme: Carbon (cyan), Oxygen (red), Calcium (blue). Blue line represents H-bond.

There is a weak global energy minimum at r_{col}=4.25Å where the hydroxyl group of AcOH points towards the oxygen of CO_3^{2-} forming a strong hydrogen bonded interaction. The free energy landscape is quite flat between 5 and 8Å and the activation barrier for dissociation of the Ca^{2+}-CO_3^{2-} ion pair from the AcOH is only 1.2 kcal/mol, of the order of the thermal energy at ambient conditions.

As AcOH in presence of Ca^{2+}, CO_3^{2-} ion-pair dissociates readily in a water-rich environment, we did not perform classical MD simulations of bulk solution containing these species.

3.4 EtOH

3.4.1 Gas phase

For the EtOH-$CaCO_3$ cluster (Figure 13A), the geometry optimized structure has the OH- group of EtOH donating an H-bond to one of the O_{CO_3} and the Ca^{2+} in in-plane bidentate coordination with CO_3^{2-}. The OH bond of EtOH is almost parallel to the CO_3^{2-} plane. The binding energy for this ion-complex is -8.61 kcal/mol. This geometry is different from that of the system containing EtOH and a neutral CO_2 molecule in which the most stable configuration is achieved through an electron donor-acceptor (EDA) type of interaction between the oxygen of CO_3 (O_{CO_2}) and the hydroxyl group of EtOH[64]. An EDA type of arrangement dominates even in a dense supercritical CO_2-EtOH mixture[64].

For the EtOH-$CaCO_3$-3H_2O system (Figure 13B), the presence of NN water molecules significantly affect the arrangement of EtOH, Ca^{2+}, and CO_3^{2-} relative to the anhydrous system. Here, EtOH donates an H-bond to O_{CO_3} where the OH bond is at 90^o to the CO_3^{2-} plane. The oxygen of EtOH (O_{EtOH}) receives an H-bond from the neighboring H_2O molecule, which is also in the first coordination shell of Ca^{2+}. The H-bond distances are 1.72 and 1.56Å, respectively. Ca^{2+} is in out-of-plane bidentate coordination by the CO_3^{2-} with all three water molecules coordinated to it at distances of about 2.3Å. Based on these results, in a wet system, protein side chains containing serine will form H-bonds with both water and CO_3^{2-}. Due to the hydrophilic nature of EtOH, or in general the hydroxyl group, the dominant interaction between these amino acids and oxy-anions could be through both solvent mediated and direct interactions, whereas interaction with cations should be through solvent or anion mediated interactions. The binding affinity of the alcohols also depends on the CH_2 chain length, with the affinity increasing with increasing chain length[65]. Proton transfer from the EtOH to the CO_3^{2-} did not occur in our calculations.

3.4.2 Free energy

The free energy profile as a function of separation (r_{col}) between O_{EtOH} and C_{CO_3} obtained from the metadynamics calculations (Figure 14) shows that the stable configuration is the dissociation of the EtOH and the Ca^{2+}-CO_3^{2-} ion pair. The free energy increases at low values of r_{col} up to r_{col}=4Å, and there is a metastable association of EtOH and the ions at r_{col}=3.45Å with an activation energy for dissociation of 0.5 kcal/mol. This energy minimum is composed of two types of complexes. In one of these, EtOH is H-bonded to the CO_3^{2-} as in the gas phase calculations, and in the other the electron lone pair of O_{EtOH} interacts with the C_{CO_3} forming an EDA complex. During the entire metadynamics simulation, the H-bonded structure is much more abundant than the EDA com-

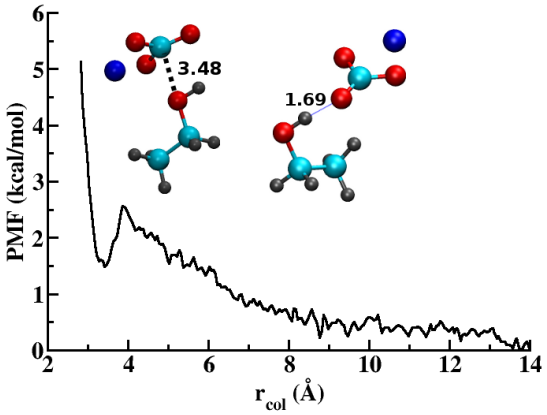


Fig. 14 Potential of mean force for the binding of EtOH with Ca^{2+}-CO_3^{2-} ion-pair. The most probable geometries of these ions and molecule at the local minimum at r_{col}= 3.45Å are shown with ball-and-stick representation. Color scheme: Carbon (cyan), Oxygen (red), Calcium (blue), Hydrogen (gray).

plex, supporting the results of our gas phase calculations. In comparison to the AcOH, EtOH shows higher binding affinity towards CO_3^{2-}. Experimental studies[65] show that the strong interaction with OH- containing groups can displace water molecules from a calcium carbonate surface facilitating direct interaction with the biomolecules that can control the crystal morphology and growth[66].

3.4.3 Bulk Solution

The RDFs between EtOH and the Ca^{2+}-CO_3^{2-} ion pair obtained from our classical MD calculations show that in a water-rich environment the H-bond interaction between EtOH and CO_3^{2-} is predominant and that the direct interaction with Ca^{2+} is rare (Figure 15). The minimum distance between the O_{EtOH} and Ca^{2+} is ∼3.5Å and the first and second maxima appear at 4.7 and at ∼6Å, respectively (Figure 15a). In contrast, there is a strong interaction between these species in alcohol-rich environments[66] with a well defined maximum at 2.5Å. The RDFs for the O_{EtOH}-O_{CO_3} (Figure 15b) and H_{EtOH}-O_{CO_3} (Figure 15c) atom pairs indicate strong H-bond formation between EtOH and CO_3^{2-}. However, during the adsorption of EtOH onto the charge-neutral (104)surface of calcite[65,66], the ethanol molecules bind strongly with the surface through both electron-donor-acceptor (EDA) interaction between O_{EtOH} and Ca^{2+} and H-bond interaction between H_{EtOh} and one of the oxygen atoms of CO_3^{2-} forming a layer of ethanol molecules. This EtOH layer transforms the polar calcite surface into a hydrophobic one due to the methyl groups of the ethanol molecules, which then facilitate the removal of water molecules from the surface and formation of second layer of ethanol through hydrophobic interactions. Thus, in a water-rich environment, OH- bearing amino acids are less likely to form $CaCO_3$ nucleation sites than during biomineralization where the concentration of Ca^{2+}-CO_3^{2-} ion pairs is generally high.

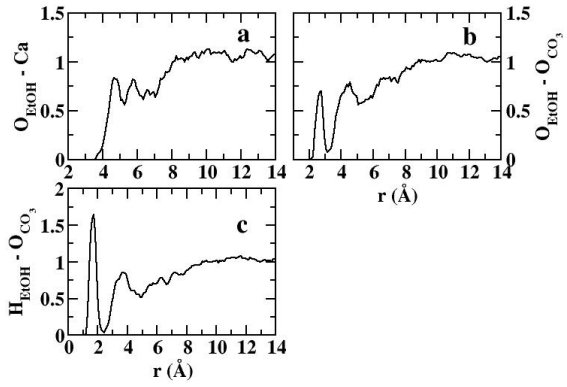


Fig. 15 Radial distribution functions for (a) oxygen (O_{EtOH}) of EtOH and Ca^{2+}, (b) O_{EtOH} and oxygen (O_{CO_3}) of CO_3^{2-}, (c) hydroxyl hydrogen (H_{EtOH}) of EtOH and O_{CO_3}.

4 Discussion

Insights from our study shed light on the atomistic details of how the NH_2^+, protonated and deprotonated COO^-, and C-OH functional groups of amino acids or organic molecules interact with Ca^{2+}-CO_3^{2-} ion pairs in water-rich and water-deficient environments relevant to biomineralization of calcium carbonate and natural and anthropogenic environmental situations. We modeled these functional groups as protonated guanidinium ion (Gdm^+), acetate (AcO^-), acetic acid (AcOH), and ethanol (EtOH), respectively, and used multiple computational techniques to study the structure and dynamics of these complexes in both isolated and bulk solvated systems. The gas-phase results indicate that the interactions of charged molecules (Gdm^+ and AcO^-) with Ca^{2+}, CO_3^{2-} ion-pair are significantly stronger than those of neutral molecules and that the presence of the first hydration shell around all these molecular-assemblies is energetically more stable than the dry state due to the formation of favorable H-bonds. Our free energy calculations show that although the free-energy differences between the bound and dissociated ion-assemblies containing Gdm^+ and AcO^- are almost equal at ∼5 kcal/mol, bound AcO^- is more stable with a barrier height of ∼9kcal/mol. Thus, the high affinity of arginine and other NH_2^+ bearing species and glutamate, aspartate, and other COO^- bearing species towards the Ca, CO_3 ion-pair indicates that these functional groups could be the potential locations of

the nucleation sites for the crystallization of calcium carbonate.

The activation barrier energy for the formation of EtOH-$CaCO_3$ is negligibly small, ∼0.5kcal/mol, and thus this molecular assembly is very unstable in a dilute, water-rich environment. However, ethanol or OH- bearing organic molecules may play an important role in controlling the morphology of the crystalline phase by stabilizing the already formed crystalline $CaCO_3$ surfaces through H-bond interactions with the CO_3^{2-} and electron-donor-acceptor interaction with the Ca^{2+}. In contrast, AcOH reacts with the neighboring CO_3^{2-} and dissociates into AcO^- and HCO^-. Experimental studies show that the presence of solvent water expedites this reaction, and thus the formation of a stable molecular assembly with AcOH is not feasible.

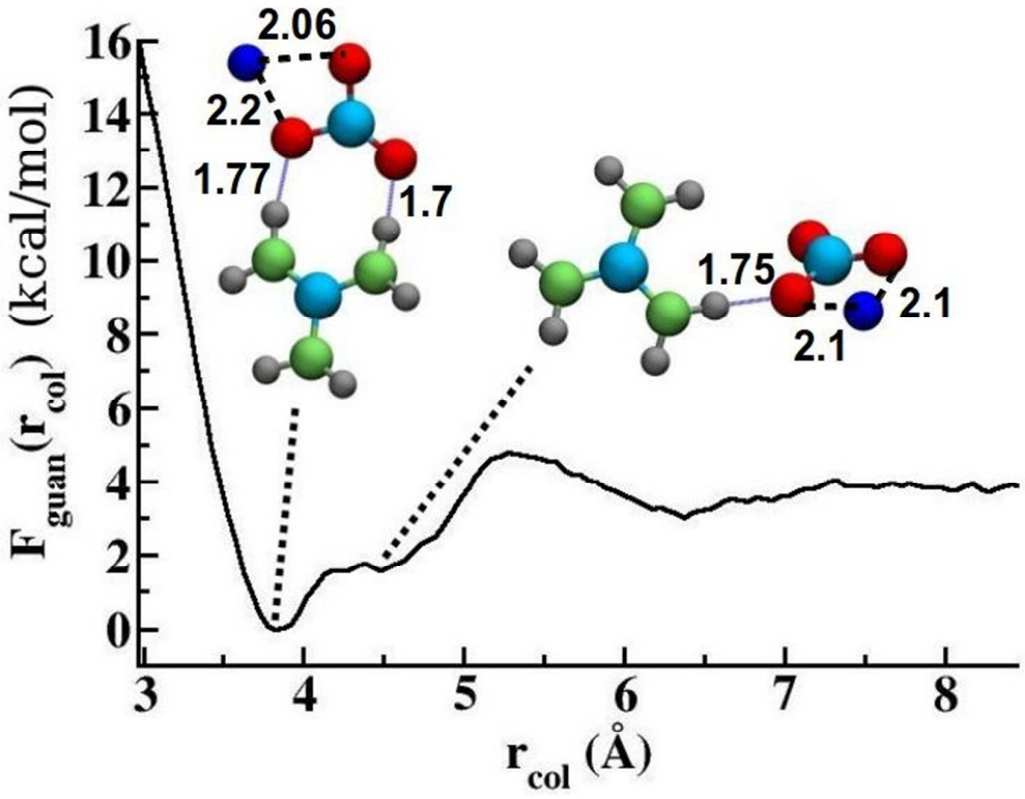
5 Acknowledgments

This work was supported by the United States Department of Energy, Office of Basic Energy Science through grants DEFG02-10ER16128 and DE-FG02-08ER15929 and also by the Science & Engineering Research Board, Department of Science and Technology, India through grant SERB/LS-2732011. We thank Drs. Geoffrey Bowers and Ozgur O. Yazaydin for helpful discussions.

References

1 N. A. Wall and G. R. Choppin, *Appl. Geochem.*, 2003, **18**, 1573–1582.
2 M. S. Diallo, C. J. Glinka, W. A. Goddard III and J. H. Johnson Jr, *J. Nanopart. Res.*, 2005, **7**, 435–448.
3 A. G. Kalinichev, E. Iskrenova-Tchoukova, W.-Y. Ahn, M. M. Clark and R. J. Kirkpatrick, *Geoderma*, 2011, **169**, 27–32.
4 W.-Y. Ahn, A. G. Kalinichev and M. M. Clark, *Geochim. Cosmochim. Ac. Suppl.*, 2008, **72**, 10.
5 A. Kalinichev and R. Kirkpatrick, *Eur. J. Soil Sci.*, 2007, **58**, 909–917.
6 S. Mann, B. R. Heywood, S. Rajam and J. D. Birchall, *Nature*, 1988, **334**, 692–695.
7 G. Xu, N. Yao, I. Aksay and J. Groves, *J. Am. Chem. Soc.*, 1998, **120**, 11977–11985.
8 J. Aizenberg, A. J. Black and G. M. Whitesides, *J. Am. Chem. Soc.*, 1999, **121**, 4500–4509.
9 Y.-J. Han and J. Aizenberg, *J. Am. Chem. Soc.*, 2003, **125**, 4032–4033.
10 K. Naka and Y. Chujo, *Chem. Mater.*, 2001, **13**, 3245–3259.
11 P. K. Ajikumar, R. Lakshminarayanan, B. T. Ong, S. Valiyaveettil and M. R. Kini, *Biomacromolecules*, 2003, **4**, 1321–1326.
12 C. Orme, A. Noy, A. Wierzbicki, M. McBride, M. Grantham, H. Teng, P. Dove and J. DeYoreo, *Nature*, 2001, **411**, 775–779.
13 H. Nagasawa, *Thalassas*, 2004, **20**, 15–24.
14 T. Samata, *Thalassas*, 2004, **20**, 25–44.
15 J. L. Arias and M. S. Fernández, *Chem. Rev.*, 2008, **108**, 4475–4482.
16 F. Meldrum, *Int. Mater. Rev.*, 2003, **48**, 187–224.
17 A. Wheeler, J. W. George and C. Evans, *Science*, 1981, **212**, 1397–1398.
18 J. Aizenberg, L. Addadi, S. Weiner and G. Lambert, *Adv. Mater.*, 1996, **8**, 222–226.
19 A. Wierzbicki, C. Sikes, J. Madura and B. Drake, *Calcified Tissue Int.*, 1994, **54**, 133–141.
20 J. Ihli, Y.-Y. Kim, E. H. Noel and F. C. Meldrum, *Adv. Funct. Mater.*, 2012, **23**, 1575–1585.
21 B. Cantaert, Y.-Y. Kim, H. Ludwig, F. Nudelman, N. A. Sommerdijk and F. C. Meldrum, *Adv. Funct. Mater.*, 2012, **22**, 907–915.
22 L. B. Gower, *Chem. Rev.*, 2008, **108**, 4551–4627.
23 C. L. Freeman, J. H. Harding, D. Quigley and P. M. Rodger, *J. Phys. Chem. C*, 2011, **115**, 8175–8183.
24 J.-W. Shen, C. Li, N. F. van der Vegt and C. Peter, *J. Phys. Chem. C*, 2013, **117**, 6904–6913.
25 J. A. Leenheer, *Ann. Env. Sci.*, 2009, **3**, 1–130.
26 S. E. Mylon, C. I. Rinciog, N. Schmidt, L. Gutierrez, G. C. Wong and T. H. Nguyen, *Langmuir*, 2009, **26**, 1035–1042.
27 A. Majzik and E. Tombácz, *Org. Geochem.*, 2007, **38**, 1319–1329.
28 E. Iskrenova-Tchoukova, A. G. Kalinichev and R. J. Kirkpatrick, *Langmuir*, 2010, **26**, 15909–15919.
29 G. Tribello, F. Bruneval, C. Liew and M. Parrinello, *J. Phys. Chem. B*, 2009, **113**, 11680–11687.
30 D. Gebauer, A. Völkel and H. Cölfen, *Science*, 2008, **322**, 1819–1822.
31 A. F. Wallace, L. O. Hedges, A. Fernandez-Martinez, P. Raiteri, J. D. Gale, G. A. Waychunas, S. Whitelam, J. F. Banfield and J. J. De Yoreo, *Science*, 2013, **341**, 885–889.
32 L. M. Hamm, A. F. Wallace and P. M. Dove, *The Journal of Physical Chemistry B*, 2010, **114**, 10488–10495.
33 A. D. MacKerell, D. Bashford, M. Bellott, R. Dunbrack, J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wirkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
34 A. E. Nielsen, *Journal of Crystal Growth*, 1984, **67**, 289–310.
35 P. Raiteri, R. Demichelis, J. D. Gale, M. Kellermeier, D. Gebauer, D. Quigley, L. B. Wright and T. R. Walsh, *Faraday Discussions*, 2012, **159**, 61–85.
36 C. L. Freeman, J. H. Harding, D. J. Cooke, J. A. Elliott, J. S. Lardge and D. M. Duffy, *The Journal of Physical Chemistry C*, 2007, **111**, 11943–11951.
37 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *Journal of the American Chemical Society*, 1995, **117**, 5179–5197.
38 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
39 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, . Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision D.01*, Gaussian Inc. Wallingford CT 2009.
40 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
41 F. David, V. Vokhmin and G. Ionova, *Journal of Molecular Liquids*, 2001, **90**, 45–62.

42 P. Raiteri and J. D. Gale, *J. Am. Chem. Soc.*, 2010, **132**, 17623–17634.
43 Y. Marcus, *Journal of the Chemical Society, Faraday Transactions*, 1991, **87**, 2995–2999.
44 A. Pavese, M. Catti, G. Price and R. Jackson, *Physics and chemistry of minerals*, 1992, **19**, 80–87.
45 T. Megyes, T. Grósz, T. Radnai, I. Bakó and G. Pálinkás, *The Journal of Physical Chemistry A*, 2004, **108**, 7261–7271.
46 X. Rozanska and C. Chipot, *J. Chem. Phys.*, 2000, **112**, 9691.
47 A. Sapse and C. Russell, *J. Mol. Struc-THEOCHEM*, 1986, **137**, 43–53.
48 J. B. Mitchell, J. M. Thornton, J. Singh and S. L. Price, *J. Mol. Biol.*, 1992, **226**, 251–262.
49 J.-C. Soetens, C. Millot, C. Chipot, G. Jansen, J. G. Ángyán and B. Maigret, *J. Phys. Chem. B*, 1997, **101**, 10910–10917.
50 A. Magalhaes and J. Gomes, *Int. J. Quantum Chem.*, 1997, **61**, 725–739.
51 M. S. Braiman, D. M. Briercheck and K. M. Kriger, *J. Phys. Chem. B*, 1999, **103**, 4744–4750.
52 F. Bruneval, D. Donadio and M. Parrinello, *J. Phys. Chem. B*, 2007, **111**, 12219–12227.
53 M. Vazdar, P. Jungwirth and P. E. Mason, *J. Phys. Chem. B*, 2013, **117**, 1844–1848.
54 J.-J. Max and C. Chapados, *J. Phys. Chem. A*, 2004, **108**, 3324–3337.
55 E. C. Meng and P. A. Kollman, *J. Phys. Chem.*, 1996, **100**, 11460–11470.
56 M. Banno, K. Ohta and K. Tominaga, *J. Phys. Chem. A*, 2008, **112**, 4170–4175.
57 G. Ruderman, E. R. Caffarena, I. G. Mogilner and E. J. Tolosa, *J. Solution Chem.*, 1998, **27**, 935–948.
58 F. Génin, F. Quilès and A. Burneau, *Phys. Chem. Chem. Phys.*, 2001, **3**, 932–942.
59 Q. Ma, Y. Liu, C. Liu and H. He, *Phys. Chem. Chem. Phys.*, 2012, **14**, 8403–8409.
60 H. A. Al-Hosney, S. Carlos-Cuellar, J. Baltrusaitis and V. H. Grassian, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3587–3595.
61 A. P. Prince, P. D. Kleiber, V. H. Grassian and M. A. Young, *Phys. Chem. Chem. Phys.*, 2008, **10**, 142–152.
62 A. P. Prince, P. Kleiber, V. Grassian and M. Young, *Phys. Chem. Chem. Phys.*, 2007, **9**, 3432–3439.
63 H. Endo, Y. Takagi, N. Ozaki, T. Kogure and T. Watanabe, *Biochem. J*, 2004, **384**, 159–167.
64 M. Saharay and S. Balasubramanian, *J. Phys. Chem. B*, 2006, **110**, 3782–3790.
65 D. V. Okhrimenko, J. Nissenbaum, M. P. Andersson, M. Olsson and S. L. Stipp, *Langmuir*, 2013, **29**, 11062–11073.
66 K. K. Sand, J. D. Rodriguez-Blanco, E. Makovicky, L. Benning and S. L. S. Stipp, *Cryst. Growth Des.*, 2011, **12**, 842–853.



A computational study of the interaction of Ca^{2+} and CO^{2-}_3 with functional groups relevant to $CaCO_3$ biomineralization and NOM using DFT, classical molecular dynamics and metadynamics methods to evaluate the structural environments, energetics of these interactions in small molecular clusters, changes in the vibrational frequencies due to ion-association, and the effects of solvation of the clusters in bulk water.