Ion Pairing and Multiple Ion Binding in Calcium Carbonate Solutions Based on a Polarizable Amoeba Force Field and *ab initio* Molecular Dynamics

Paolo Raiteri,* Alicia Schuitemaker, and Julian D. Gale*

Curtin Institute for Computation/The Institute for Geoscience Research (TIGeR), School of Molecular and Life Sciences, Curtin University, P.O. Box U1987, Perth, WA 6845, Australia

E-mail: p.raiteri@curtin.edu.au; j.gale@curtin.edu.au

Abstract

The speciation of calcium carbonate in water is important to the geochemistry of the world's oceans and has ignited significant debate regarding the mechanism by which nucleation occurs. Here it is vital to be able to quantify the thermodynamics of ion pairing versus higher order association processes in order to distinguish between possible pathways. Given that it is experimentally challenging to quantify such species, here we determine the thermodynamics for ion pairing and multiple binding of calcium carbonate species using bias-enhanced molecular dynamics. In order to examine the uncertainties underlying these results, we have derived a new polarizable force field for both calcium carbonate and bicarbonate in water based on the AMOEBA model to compare against our earlier rigid ion model, both of which are further benchmarked against *ab initio* molecular dynamics for the ion pair. Both force fields consistently indicate that the association of calcium carbonate ion pairs to form larger species is stable, though with an equilibrium constant that is lower than for ion pairing itself.

1

Introduction

Calcium carbonate ($CaCO_3$) is one of the most abundant minerals in the natural environment and is often formed as a product of biomineralization. Carbonate minerals more broadly are also important as a means for long-term carbon sequestration,[1] while also being utilized in a number of industrial processes. Given this range of contexts in which calcium carbonate is significant, there has been considerable interest in its' crystallization pathways, something that has been further heightened by the debate regarding so-called "non-classical" nucleation mechanisms.[2] There is already a growing body of evidence that formation of calcium carbonate is a complex process that can involve multiple steps, which may include formation of pre-nucleation clusters, liquid-liquid phase separation,[3] precipitation of polyamorphs of amorphous calcium carbonate,[4] and transient crystallization to disordered vaterite, depending on the specific conditions, prior to formation of the stable phase at ambient conditions, calcite.

One of the principal debates regarding calcium carbonate crystallization relates to the speciation that occurs in aqueous solution prior to the first nucleation event. Experiments in which a calcium-containing solution is gradually added to an excess of carbonate buffer reveal that a fraction of the calcium ions are contained within bound species, as detected by an ion-selective electrode. The traditional interpretation of this binding profile is to assign the linear curve wholly to formation of the calcium carbonate ion pair, $[CaCO_3]^0$. However, the work of Gebauer et al[5] based on analytic ultracentrifugation indicates two peaks that point to the existence of two separate species distributions, with approximate average dimensions in nanometer regime. Based on this evidence, they propose a multiple binding model for calcium to carbonate, which is also consistent with a linear-binding profile for ions in solution.

Given there exist two experimental models, pure ion pairing and multiple binding, that both have the potential to explain the observed initial speciation of calcium carbonate there have been a number of attempts to study this system using computer simulations as a means to resolving this dilemma. Tribello et al[6] used umbrella sampling to study both ion pairing and the dimerization of this species, leading to the finding that both processes were strongly exergonic with values of approximately 8 and 4 kcal/mol, respectively. Following an assessment of the thermodynamics of the available force fields, Raiteri et al[7] developed a series of calibrated new models that focused on free energies of hydration and solubility that led to ion pairing strengths that were in closer agreement to experiment. An important observation, as highlighted by Kellermeier et al,[8] is that ion pairing and indeed crystallization of calcium carbonate in

general is driven by the favorable entropy that arises from ion dehydration. Demichelis *et al*[9] studied the association of calcium carbonate in concentrated solutions leading to the identification of dynamic polymers of calcium carbonate in water known as DOLLOP. Under some conditions it was possible to observe dynamic ion cluster formation and dissociation from which equilibrium constants were extracted that suggest that ion pair association is of similar stability to ion pairing itself in accord with the multiple binding hypothesis. In contrast, simulations by Smeets *et al*[10] at lower concentrations suggest that ion pair association is less favorable. One possible explanation for this discrepancy could be that there is a change in stability after crossing the binodal beyond which a dense-liquid phase is formed, which occurs at a saturation index of approximately 2 relative to calcite.[66]

To date, the majority of simulations of ion association for Ca^{2+} and CO_3^{2-} have used either rigid ion force fields or more occasionally a shell model, though with inaccurate thermodynamics for aqueous calcium carbonate systems in the latter case. Here we introduce a new force field for this system that is based on the AMOEBA model,[11] which includes not only static multipoles, but self-consistent induced dipoles. Based on this model we determine the stability for ion pair association up to the tetramer and examine the influence of including polarization by comparing to data for the best rigid ion model. As a further means of validation, we compare to results from *ab initio* molecular dynamics for the case of the contact to solvent-shared ion pair.

Methods

As the present study involves the development and calibration of a new force field, followed by the application to the problem of the initial speciation of calcium carbonate in water, a variety of different computational techniques have been used as appropriate to the quantity of interest. Below we describe each of the methods and the computational settings employed for each stage of the development processes and subsequent application of the force field.

Force field and initial parameterization

Previous studies have shown that it is possible to generate reasonably accurate force fields for ions in aqueous solution based on simple rigid ion potentials, provided the training set includes all of the relevant thermodynamic quantities of interest, such as hydration free energies. However, the challenge when

simulating crystallization is that any model must handle all extremes from the aqueous phase, where the dielectric constant is high, through to the crystalline state where interactions are no longer screened to the same extent. While fixed charge rigid ion models can capture one extreme or the other, achieving transferability between the two states is more difficult. This is equally true for water when comparing the gas phase to the liquid where the dipole moment changes substantially. The obvious solution is to include polarizability or charge transfer within the force field, though this comes at a computational cost. Many different approaches to achieve this have been proposed over the last few decades. In the specific context of mineral systems, there has been extensive use of the shell model, though attempts to generate a reliable parameterization for water have proven unsuccessful.[12,13] Alternatively, damped point ion polarizability can be utilized and readily extended from dipoles to quadrupoles,[14] as well as higher moments if required. Charge transfer, which is capable of describing some of the same phenomenology, can be included through schemes such as electronegativity equalization, as used in the ReaxFF approach,[15] for example, and has been previously applied to aqueous calcium carbonate systems.[16]

In this work we explore the use of the AMOEBA force field, as developed and primarily utilized for a range of organic and biological problems. Full details of the AMOEBA approach can be found elsewhere;[11,17] in short, it includes static multipoles up to the quadrupole for molecular fragments, augmented by dipolar polarization with Thole damping. Here we use the standard AMOEBA approach in which the dipole-dipole interactions are solved self-consistently, as opposed to the computationally cheaper iAMOEBA approximation[18] in which the induced on-site dipoles are treated as non-interacting. For the development of a thermodynamically-based polarizable force field for calcium, carbonate and bicarbonate in water we started from the AMOEBA03 model,[19] whose parameters have been subsequently modified to correct an error in the ordering of the O–H stretch vibrational modes.[20] While initial parameters already exist for water and Ca^{2+}, it was necessary to develop models for both carbonate and bicarbonate. Here the standard protocol was followed to generate the reference geometries and charge distributions for both molecular anions. This involves optimization at the MP2/6-31G* level, followed by a population analysis with a larger basis set (MP2/6-311G**). The multipole moments were then determined by post-processing using GDMA,[21] followed by poledit. It is important to note that formally the carbonate anion is unstable in the gas phase and should eject one electron. However, the use of finite atom-centred basis sets makes the calculation possible.

Gas phase quantum mechanics

In order to provide reference data for the fitting and validation of short-range interaction parameters within the AMOEBA force field, we have generated gas phase quantum mechanical data for species involving Ca^{2+}, bicarbonate and water. In the case of bicarbonate, this is particularly important since there are multiple distinct interactions between water and the anion based on the carboxylate and hydroxyl groups. Hence quantities such as the hydration free energy alone are insufficient to determine a unique parameter set. All gas phase quantum mechanical data for this purpose were generated using the ORCA code (version 4.1.1).[22] Geometries were optimized at the ωB97X-D3/ma-def2-QZVPP level of theory,[23] which has been benchmarked and shown to be one of the best performing functionals for thermodynamics in its' class. Numerical integrations of the density were performed using Grid4.

Force field optimization and fitting

When refining the AMOEBA parameters against molecular quantum mechanical data, as well as computing the stability and properties of solid phases, all calculations were performed using an in-house version of GULP[24] that has been modified to include the AMOEBA energy and forces based on Tinker. All optimizations of solids were performed using the cell parameters as variables, rather than the strain tensor, and either Newton-Raphson methods, starting from a unit Hessian matrix with BFGS updating, or conjugate gradients. For solids, the electrostatics were determined using an Ewald summation. Second derivative properties were derived by central finite differences and in the case of molecular vibrational frequencies, Eckart purification was applied to separate rotational and translation modes. To compute the free energies of species in the solid or *in vacuo*, the quasi-harmonic values were computed at 298.15 K for the optimized structures with respect to internal energy, and so thermal expansion is assumed to have a negligible effect on the overall thermodynamics. In addition, the zero point energy contribution is excluded from the free energy, since the objective is to use the force field primarily for molecular dynamics (MD) simulations.

Hydration free energy

The hydration free energies of Ca^{2+}, CO_3^{2-} and HCO_3^- for the AMOEBA polarizable force field have been computed using the free energy perturbation[25] and Bennett acceptance ratio (BAR) methods[26] as

implemented in TinkerHP (v1.2)[27] The solute was placed in a water box of approximately 25 Å in side and equilibrated at 300 K and 1 atm using the Andersen thermostat and Berendsen barostat. The cell was then fixed to its average value and the free energy perturbation was carried out at constant volume. Eleven simulations were run for the perturbation of the electrostatic and van der Waals interactions separately, with the perturbation parameter λ set to [0.0, 0.2, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, 0.8, 0.9, 1.0] for the van der Waals interactions and to [0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0] for the electrostatic interactions. A cutoff of 9 Å was used for the van der Waals interactions and the real space part of the electrostatics; the Ewald summation was used for the reciprocal space part of the electrostatic interactions and the criterion for the convergence of the self-consistent induced dipoles was set to 10^{-5}. Each simulation was run for 1 ns with a timestep of 1 fs and the atomic coordinates used for the BAR calculations were recorded every 100 fs. The reported hydration free energies are the average of those computed for the creation and annihilation of the solute particles and the error was obtained from the BAR code. The finite size correction for the solvation of ionic species was also included.[28] For comparison, we have also computed the hydration free energy of bicarbonate based on a modified version of our earlier rigid ion force field, which was computed with LAMMPS[29] following the same procedure described in our previous work for calcium and carbonate.[30]

Hydration properties

Standard NVT MD simulations have also been run with the OpenMM[31] code to characterize the hydration structure and properties of the calcium, carbonate and bicarbonate ions in water. The equations of motions were integrated with the velocity Verlet algorithm and the temperature was maintained at 300 K by using the Nosé-Hoover thermostat. Each simulation was at least 5 ns long. The self-diffusion coefficient of species i was computed using the Einstein equation;

$$D_i^L = \lim_{t \to \infty} \frac{1}{6} \left\langle \left| r(t) - r(0) \right|^2 \right\rangle$$

in three different cubic boxes, $L = 25$ Å, $L = 50$ Å and $L = 75$ Å, and extrapolated to infinite size to remove any finite size effects using the equation;

$$D^\infty = D^L + \frac{k_B T \zeta}{6\pi\eta L}$$

where η is the system viscosity and $\zeta = 2.837297$.[32]

Ab initio molecular dynamics

To complement the gas phase benchmark quantum mechanical data, we have also used *ab initio* molecular dynamics in order to provide information on the species in solution for comparison with the force field results. Here all calculations have been performed using the program CP2K[33] using the Gaussian-planewave approach of the Quickstep module.[34] The majority of simulations have been performed using the BLYP exchange-correlation functional,[35,36] with occasionally PBE being used for comparison to determine the sensitivity to the choice of GGA. In both cases dispersion corrections were included using the D3 approach of Grimme and co-workers.[37] The norm-conserving pseudopotentials of Goedecker-Teter-Hutter[38] were used, with a small core for Ca, in conjunction with the triple-zeta double polarized basis set (TZV2P) except for the more ionic case of Ca^{2+} where DZVP was chosen instead. The auxiliary basis set planewave cutoff was set to be at least 300 Ry. Solution of the Kohn-Sham equations was achieved by using the orbital transformation algorithm with full kinetic preconditioning.[39]

All *ab initio* molecular dynamic simulations were performed in the NVT ensemble at a temperature of 330 K with a hydrogen mass of 2 amu. This set of conditions was selected based on previous observations that the use of an elevated temperature partially compensates for the over-structuring of liquid water by most GGA functionals. As will be shown later, based on the force field data, the use of a higher temperature might lead to a small over-stabilization of ion pairs, though the magnitude is likely to be of the order of thermal energy and therefore no greater than the statistical uncertainty in any thermodynamic properties computed. A time step of 0.5 fs was used, while temperature control was maintained using the canonical sampling through rescaling of velocities algorithm.[40]

To provide comparative data for the hydration structure of ions in water, unbiased molecular dynamics was performed for each ion in a cubic box of water with a side length of at least 14.3 Å. Initial coordinates were taken from equilibrated force field simulations and then run using *ab initio* molecular dynamics for

at least 50 ps. In addition to these simulations, runs were also performed to examine the formation of the calcium carbonate ion pair in water. To do this requires at least 2 collective variables; one that relates to the separation of the ions (in this case the Ca-C distance), and also the Ca-water coordination number. Unlike the force field simulations, it is not feasible to sample the equilibrium calcium hydration state over the potential energy surface without an accelerating bias, given that the timescale for water exchange is comparable to the duration of the run. Here we have used both 2-D umbrella sampling (PBE-D3) and metadynamics (BLYP-D3) based on these collective variables in order to assess the reliability of the approaches when used with the relatively meagre statistics available from *ab initio* molecular dynamics with current computational resources. For the metadynamics simulations, ten walkers were used with an initial Gaussian height of 2 kJ/mol, a width of 0.2 Å or 0.1 for distance and coordination number, respectively, and well-tempering with a ΔT of 1500 K. For the umbrella sampling, 205 umbrellas were ultimately used to span the two-dimensional space, with umbrellas being added as necessary to ensure overlapping distributions. The spring constants varied between 1-4 eV/Å^2 and 0.05-0.1 eV for the distance and coordination number collective variables, respectively. In both cases the cumulative simulation time over all walkers/umbrellas was of the order of 1 ns.

Ion binding free energy

All molecular dynamics simulations for the calculations of the binding free energies with the rigid ion force field were computed using LAMMPS,[29] while TinkerHP[27] was used for the ion pairing free energies with the AMOEBA polarizable force field. The OpenMM[31] code was instead used to compute the AMOEBA multiple binding free energies, which required larger simulation cells, due to its better performance on GPUs using mixed precision.[41,42] The open-source, community-developed PLUMED library,[43] version 2.5,[44] was used for all the ion binding free energy calculations.

The multiple walker well-tempered metadynamics[45–47] technique was used to reconstruct the binding free energy as a function of the distance between the centers of mass of the two binding species, which were either individual molecules or small clusters. In the case of small clusters, a restraining potential was added to ensure that the cluster remained bound. This was achieved by restraining the calcium distance or the calcium by carbonate coordination number within the cluster.[48] The classical molecular dynamics simulations were run following the same protocol as in our previous work.[8] The AMOEBA simulations were run with at least 20 independent walkers and a bias factor of 5. The temperature was controlled

with a Langevin thermostat, the equation of motions were integrated with a time step of 1 fs and the convergence criterion for the self-consistent induced dipoles used was again set to 10^{-5}. The biasing potential was constructed by successive additions of Gaussians laid every 0.5 ps, which had an initial height of 2.6 kJ/mol and width 0.2 Å, and at least 128,000 Gaussians were deposited. The convergence of the free energy calculation was checked by ensuring that the final height of the Gaussians was less than 0.01 kJ/mol in the relevant areas of the free energy landscape. An upper wall to the distance was also used to limit the collective variable (CV) space explored. Because of the fast water exchange rate around the calcium ion, relative to the length of the classical MD simulations, it was deemed unnecessary to use the calcium by water coordination number as a second collective variable. The calcium bicarbonate and carbonate ion pairing free energies were also computed at different temperatures in the range 290-350 K. At each temperature the water density had been previously equilibrated by using a Monte Carlo barostat[49] with a target pressure of 1 atm.

The calculation of the binding free energy of the 2 formula unit cluster has been calculated following three different pathways, by addition of one ion pair to another one, or by successive addition of Ca^{2+} or CO_3^{2-} ions. From a thermodynamic stand point, the three pathways should give the same formation free energy of the 2 formula unit cluster. However, because of the complexity and unavoidable statistical errors in the calculations a certain discrepancy between the three pathways has to be expected. For both the rigid ion force field and AMOEBA calculations the maximum discrepancy observed is about 3 kJ/mol (i.e. close to ambient thermal energy), which could be seen as an indication of the precision of the calculations. In the case of the rigid ion force field, the same procedure was carried out for the calculation of the formation free energy of a 3 formula unit cluster, while this was too demanding for the AMOEBA force field. Analogously the formation free energy of a 4 formula unit cluster was computed only via the direct addition of an ion pair to the $(CaCO_3^0)_3$ cluster.

Results and Discussion

Parameterization of AMOEBA

The objective of the current study is to develop a thermodynamically accurate model for aqueous calcium carbonate chemistry based on a polarizable force field. As the AMOEBA model already offers a good

description of water, and has a more sophisticated treatment of higher multipole moment contributions to the electrostatic energy than the majority of other force fields, we decided to use this as the starting point for further parameterization. There are several variants of the AMOEBA parameters for water, most notable the 2003 and 2014 models.[19,20] Here we have chosen to use the AMOEBA03 water as the basis for our model as it was found to more readily fit both the hydration free energy and water coordination of Ca^{2+} within the constraints imposed by the combination rules. As noted above, there is a more recent AMOEBA water model (AMOEBA14), however, because of the use of combination rules, the parameters developed in this work cannot be used with the AMOEBA14 water model without appropriate reparameterization.

Having adopted a parameter set for water, the next step is to create initial models for carbonate (CO_3^{2-}) and bicarbonate (HCO_3^-) as the two key anions of interest. Here we followed the traditional procedure of performing gas phase quantum mechanical calculations to determine the optimized geometry and charge density, followed by subsequent multipole analysis in order to define the charges, dipoles and quadrupoles which were then held fixed. Initial polarizabilities and short-range interaction parameters were then assigned based on similar functional groups already present within the AMOEBA03 library, such as carboxylate and hydroxyl groups. Similarly, existing parameters for calcium ions[50] were taken as a starting point.

In the first stage of parameter refinement we examined the properties of the three ions in aqueous solution. An important quantity for a thermodynamically-focused force field is the hydration free energy of the ions and so this was selected as a target quantity for the fitting process. Because single ion hydration free energies are not directly obtainable from experiment, there is considerable variability in some of the literature values. In particular, for Ca^{2+} the values range between -1444 and -1527 kJ/mol.[51,52] In our earlier fitting of a rigid ion model with SPC/Fw, the lower bound to the hydration free energy was reproduced. One of the reasons for this is that in the absence of a polarizable water model it is not possible to obtained hydration free energies that are more negative than this without having calcium-water distances that are too short, which then has negative consequences for the water coordination number and dynamics of water exchange. With the use of AMOEBA water this constraint is lifted and so it is possible to reappraise the choice of the most appropriate experimental hydration free energy. Quantum mechanical calculations of Ca^{2+} surrounded by an explicit shell of either six or seven water molecules embedded in a polarizable continuum model yield hydration free energies that are more exergonic than -1500 kJ/mol and so it was decided to lower the target free energy commensurately for this model. In the case of bicarbonate,

there is also a range of literature values from -335 to -368 kJ/mol.[52,53] As per Ca^{2+}, the rigid ion force field was limited to yielding a hydration free energy beyond the upper bound due to the compromise between solvent structure and thermodynamics, whereas for the AMOEBA-based model we are able to obtain a value that sits at the mid point of the literature range. Finally, for carbonate the hydration free energy must be regarded as having considerable uncertainty since it is a hypothetical reaction. Hence we take the same target value from the literature as in the case of our previous non-polarizable force field. A full comparison of the final hydration free energies is given in Table 1.

Alongside the fitting of the solvation free energy, we also considered information on the hydration structure of each species in water. This is particularly important for the case of bicarbonate where there is an ambiguity based on the hydration free energy alone since it depends on how the interaction strength is partitioned between the carboxylate and hydroxyl regions of the anion. To handle this, molecular quantum mechanical calculations were first performed to compute the binding energy and intermolecular distances for water molecules coordinating to the species *in vacuo*. This data was then used to perform an initial refinement of the short-range interaction parameters. Subsequently, as well as computing the hydration free energy in bulk water, we also determined the radial distribution functions (RDFs) for the water surrounding the species for comparison with the results of *ab initio* molecular dynamics, both from previous literature and as computed in this study with the addition of dispersion corrections. The RDFs for the final parameters for carbonate and bicarbonate are shown in Figures 1 and 2, respectively. Considering the simpler case of carbonate first, the AMOEBA water structure is in excellent agreement with that computed at the BLYP-D3 level of theory based on the peak positions, with the only slight discrepancy being the narrower peak width and therefore greater height for the inner peak. Of course it is also important to note that the statistical quality of the AMOEBA data is far higher due to the short run lengths accessible to *ab initio* methods at present. While the rigid ion force field also gives a good description of the water structure around the oxygen atoms of carbonate, there is a noticeable error in the RDF between carbon of carbonate and oxygen of water, which is distinctly bimodal. One of the likely reasons for the improvement in this regard by the AMOEBA model is the inclusion of atom-centred quadrupoles to improve the electrostatic potential of carbonate in the out of plane direction where this interaction occurs. Turning now to consider the case of bicarbonate, there is an obvious strong correspondence between the RDFs obtained from DFT and AMOEBA; the only issue is that there is a systematic shift for AMOEBA by 0.2 Å to longer distances. This discrepancy may point to the DFT

results, at least at the BLYP-D3 level, being more consistent with the exergonic extreme of the literature hydration free energy range. As per the case of carbonate, the AMOEBA model performs much better than the rigid ion force field in describing the RDFs, where the latter model tends to overestimate the strength of the hydrogen bond from H of bicarbonate to water, while underestimating the strength of the hydrogen bond accepted by the carboxylate group. Importantly, the present fit for AMOEBA shows a good balance between the two different functional groups of bicarbonate.

To complement and supplement the data from the RDFs, further properties of the ions in bulk aqueous solution are given in Table 2. Across the whole suite of data the AMOEBA model is almost uniformly in better agreement with experiment than the non-polarizable force field, or at least no worse. Our new model exhibits uniformly lower self-diffusion coefficients bringing the values closer to experimental estimates. For the calcium cation both force field models are consistent with a range of hydration states spanning 6, 7 and 8 water first neighbors, and the average is 7.2 in both cases. The carbonate ion now yields a slightly lower number of directly hydrogen bonded water molecules more in line with other literature estimates, while bicarbonate shows the reverse trend when changing to the polarizable model. The water residence times for the AMOEBA model were calculated following the same procedure used in Ref. 54 and are comparable to those for the rigid ion force field (see Table 2).

Having determined an AMOEBA parameter set that provides a good overall description of the ions in water, the next step is to be able to reproduce the relative thermodynamics of the solid state for calcium carbonate. To do this, we consider the solubility via the free energy of dissolution for calcite which is well known experimentally. As for our earlier force fields, we determine this by calculating the free energy of the solid phase (excluding zero point energy for compatibility with molecular dynamics) and then taking the difference with the sum of the hydration free energies, plus corrections for gas phase contributions. In principle, no further parameters are required to simulate calcite once the ions have been fitted in water since the calcium-carbonate short-range interactions are determined by combination rules. However, because we wish to fit the solubility some additional modification is required. One option is to introduce pair-wise specific values of epsilon and sigma. However, because this is not supported by all implementations of AMOEBA (specifically that of OpenMM on GPUs) we have avoided this approach for now. Instead, we choose to add an exponential repulsion between calcium and the oxygen of carbonate as a supplementary short-range contribution that is tapered to zero over 3 Å to a cut-off of 9 Å. Here the exponent was fixed and only the scale factor varied in order to fit the stability of calcite. The resulting free energy of dissolution

and optimized structural and mechanical properties of calcite are compared in Table 3. The optimized cell parameters for calcite using the augmented AMOEBA model closely match the experimental values to within +0.33% and -0.06 % for a and c, respectively, which is superior to the rigid ion force field. It should be noted that here we are comparing optimized structures at 0 K with experimental data at room temperature, though given the thermal expansion coefficients of calcite are of the order of 10^{-5} K^{-1} or less then this does not significantly alter the quality of the comparison. The dissolution free energy of calcite with AMOEBA is also very close to the experimental value. Note that the tabulated value is computed based on lattice dynamics (i.e. quantized phonons), but could also be computed based on equipartition theory to approximate the result of a molecular dynamics simulation, in which case a value of +49.3 kJ/mol is obtained. Hence the two theoretical estimates bracket the experimental range. Mechanically, the AMOEBA parameterization gives calcite properties that are too hard, as is generally the case for force fields that adopt formal charges, despite the ability of polarization to offset this to some extent. Given the objective here is to obtain accurate thermodynamics, we have prioritized this and structural properties over the mechanical behavior, which is less critical in the context of crystallization.

An important objective of employing a polarizable force field model is to try to achieve greater transferability between solution phase and low dielectric environments. Hence we can also compare how well the final parameters perform against the data used to determine the initial interaction parameters with water. Furthermore, in the case of the calcium bicarbonate interaction, where there is no solubility data to compare to, it also provides a means of validating the thermodynamics. The comparison of the results for the final parameters against the reference QM data for the gas phase are given in Table 4. The benefits of the inclusion of polarizability are clearly seen in the results for a calcium cation interacting with water. Here the rigid ion model underestimates the binding of a single water molecule, but overestimates the strength of water coordination for six or seven molecules, which is close to the environment in solution. In contrast, the AMOEBA model follows the trend of the QM data, though slightly underestimates the binding energy in each case. Further refinement might be possible by varying the polarizability of water, though this was not attempted here due to the risk of adversely changing some other property of the water model. It should be noted that the error in the binding energy per water is of the order of k$_B$T at ambient conditions, which is acceptable.

When comparing the gas phase quantum mechanical results versus AMOEBA for bicarbonate hydration then the force field systematically underestimates the strength of interaction, as shown by both the

13

energetics and the hydrogen bonding distances. This accords with the observation presented above for the RDFs of bicarbonate in bulk water. Collectively, these results are suggestive of the literature hydration free energy of bicarbonate being insufficiently exergonic since the fitting of this quantity largely dictates the overall strength of binding.

The final parameters of the rigid ion force field for bicarbonate and of the AMOEBA model for aqueous calcium carbonate and bicarbonate systems are given in Supplementary Information.

Benchmarking models for calcium carbonate ion pairing in water

The ion pairing of calcium carbonate has already been studied extensively in the past, both experimentally and computationally.[8,55,56] Here the objective is to benchmark how well our AMOEBA-based model performs in this regard. Shown in Figure 3 are the ion pairing free energies as a function of ion separation at 300 K for both calcium carbonate and calcium bicarbonate. In both cases the curves align well with the theoretical long-range limit for two charged particles in a dielectric medium (based on the dielectric constant of the appropriate water model), plus the entropy for the radial volume of the state. The dielectric constant for the AMOEBA 03 water was taken from Ref. 20. This demonstrates that both calculations are well-converged and allows the determination of the arbitrary shift in the free energy, as well as to extract the binding free energy at standard conditions by assuming that the bound state includes both the contact and solvent-shared minima.[30] For the case of calcium carbonate, there are several clear differences in the free energy landscape between the rigid ion force field and AMOEBA. Firstly, AMOEBA suggests that the bidentate contact ion pair is the stable form, while the monodentate configuration is effectively a point of inflection; in contrast the rigid ion force field gives the opposite order of stability with a weak minimum for bidentate coordination. Secondly, the overall depth of the well for the contact ion pair is diminished in the case of AMOEBA. Thirdly, and finally, AMOEBA suggests that there is a more significant barrier of approximately 16 kJ/mol to forming the contact ion pair, while our previous force field gave a barrier comparable to $k_B T$. Interestingly, in the case of the ion pairing free energy curve for calcium bicarbonate the first of these observations is reversed when comparing the two models, while both give a similar and more substantial barrier to dehydration of the solvent-shared state on route to the contact ion pair.

Given the differences between the rigid ion and polarizable model for the $CaCO_3^0$ ion pair it is important to seek further validation. Henzler et al[57] have previously computed the free energy of ion pairing for the system using ab $initio$ molecular dynamics at the BLYP-D2/DZ level based on umbrella sampling. In

many regards their free energy profile up to an ion separation of ≈ 4.5 Å agrees well with the present AMOEBA result, in that the bidentate contact ion pair is the more stable state, the relative depth of this to the solvent-shared state is similar, and there is an increased barrier height between these two aforementioned states. Despite this seemingly favorable comparison, there is considerable uncertainty regarding the DFT free energy since it was determined using a single distance collective variable, even though the hydration state of the ions may not reach equilibrium on the timescale of 45 ps available to each umbrella. Consequently we decided to perform further *ab initio* molecular dynamics in this case using similar levels of theory, but a larger basis set (BLYP-D3/TZ2P and PBE-D3/TZ2P). Importantly, the free energy was mapped in 2D using both the Ca-C distance and the coordination number of calcium by oxygen of water in a manner similar to the force field simulations. Recently it has been shown that water exchange at carbonate has a timescale of the order of 12.7 ps from *ab initio* molecular dynamics,[58] while estimates for calcium are far slower and typically lie in the range of 60-80 ps.[59,60] It should be noted that the water exchange rate of calcium is particularly sensitive to the coordination number; exchange within between 6, 7 and 8 waters being faster than states where less than 6 waters are present, such as at surfaces.[54] Based on this the choice is made to only bias the solvation of the cation in order to minimize the number of collective variables. The resulting 2-D free energy maps and integrated 1-D profiles are shown in Figures 4 and 5, respectively.

Analysis of the free energy landscape for ion pairing from *ab initio* molecular dynamics shows that both the contact and solvent-shared ion pair states are composed of multiple minima. For the contact ion pair there are two distinct minima each for the bidentate and monodentate states that differ in the number of waters coordinated to Ca^{2+}, with the values being 4/5 and 5/6, respectively. Arguably, there may be a further minimum for the bidentate case with 6 waters, though this currently appears as an extension of the monodentate basin with no discernible separating barrier. For both coordinations of carbonate to calcium it appears that 5 waters of hydration are preferred for the cation. The 2D free energy landscape also provides insight into the mechanism for contact ion pair formation: The most probable pathway is that the solvent-shared ion pair with 6-fold water coordination transitions to the contact ion pair first, followed by subsequent loss of a water molecule to the favored state with 5 waters of hydration for the cation. The fact that one of the activated steps along the pathway is orthogonal to the Ca-C distance demonstrates why it is important to sample the free energy with an extra collective variable for the coordination number. Even with two collective variables it is important to sound a note of caution as to the likely convergence

of the free energy landscape. During analysis of the configurations sampled from the umbrellas it was observed that the solvent-shared ion pair often exhibited a structure where two water molecules that were coordinated via oxygen to Ca^{2+} both formed a hydrogen bond to the carbonate ion. This state, with two bridging waters that are strongly polarized by the electric field between the ions, is particularly stable and exists with a timescale that is likely to exceed the length of any single umbrella. Although it is not completely orthogonal to the distance collective variable, there is a risk of incomplete sampling for the run lengths currently accessible to *ab initio* molecular dynamics. Attempts to include a third collective variable based on the combined probability of water being bonded to calcium and carbonate were unsuccessful due to the slow convergence of the 3D free energy landscape.

Considering in more detail the solvent-shared ion pair, there is competition between a calcium coordination number by water of 6 and 7, with the former being the more stable minimum. No minimum for eight-coordination was observed, which is likely to exist for the solvent-separated ion pair and beyond to the dissociated ions. Because of the small size of the simulation box (14.3 Å), this limits the distance to which the free energy can be sampled with a wall being placed just inside half the box length to ensure a dominant nearest image. An additional set of simulations were performed with a larger box size (18.11 Å) in order to explore the distance range from the solvent-shared ion pair out to the solvent-separated ion pair. In the region of overlap between the two boxes similar free energy landscapes were obtained, with the larger box showing two channels for dissociation based on either the 6- or 7-fold water coordination of calcium. This differs from the rigid ion force field results where 7- and 8-fold coordination are the preferred hydration states in this region.

In Figure 5, the free energy profiles for ion pairing as a function of just the Ca-C distance are shown after integrating out the cation-water coordination number collective variable. Here we compare the results of independent runs with both umbrella sampling and metadynamics based on two different GGA functionals, both with dispersion corrections. As is evident, in the region that spans the contact ion through to the solvent-shared ion pair both enhanced sampling methods and functionals give excellent agreement with differences that are likely to be of the order of the statistical convergence of the data. This indicates that the results are not likely to be strongly sensitive to which form of GGA is used, though a greater shift might be expected for those that are parameterized for optimal performance in the solid-state. At distances above 5 Å the two free energy curves deviate. This is almost certainly due to the rapidly expanding volume of configuration space for larger distances making it hard to converge the increasingly important entropic

contribution. Given this, and the limited timescales of *ab initio* molecular dynamics, the free energy past 5 Å should be disregarded due to the incomplete sampling.

When comparing the 1D free energy profile between the present study and the previously published curve of Henzler *et al.*[57] then there are some clear differences, even though both are nominally GGA calculations using a similar sized unit cell. In particular, the order of stability between the contact and solvent-separated ion pairs is reversed, in addition to quantitative differences in the barrier heights. The main differences in the methodology are that this study uses a larger basis set (TZ2P vs DZ as stated by Henzler *et al*) and in the present case a second collective variable was included to accelerate the sampling of the cation hydration state that leads to multiple overlapping minima at the same distance. In addition, two sets of simulations were performed here with different enhanced sampling techniques leading to essentially identical results. Based on this collective evidence, it seems likely that the present results should be a more accurate reflection of the converged GGA result. A further point of difference is that Henzler *et al* used a temperature of 300 K, whereas in this work we chose to use the elevated temperature of 330 K to correct for the overstructuring of liquid water by the GGA functionals, as proposed in the literature.[61] It should be noted that the elevated temperature will also facilitate a greater sampling of the configuration space, again aiding convergence. However, this raises the question of whether the difference in contact ion pair stability is just a consequence of the temperature change. There are two reasons to believe this is not the cause of the discrepancy: Firstly, as will be shown later, the free energy of the ion pair decreases with increasing temperature and thus the effect of temperature goes in the opposite direction to the discrepancy. Secondly, the magnitude of the difference is also quantitatively larger than the change in free energy with temperature over this range.

Having argued for the veracity of the current DFT-based free energy profiles, this creates the seemingly unfortunate situation that the GGA results disagree with those of AMOEBA in terms of the relative stability of the two inner ion pair states, though the initial barrier to go from the solvent-shared ion pair to the contact one does agree nicely. At this point it is necessary to consider the systematic errors that might be associated with the DFT results. One obvious issue is that the *ab initio* molecular dynamics is necessarily performed in a small periodic cell, whereas the force field results are examined in detail to ensure convergence with respect to cell length. Based on the force field models it is straightforward to test the likely magnitude of the error associated with using a cell of dimension 14.3 Å in comparison to a much larger unit cell. The smaller cell does indeed lead to a positive shift in the free energy of binding for the ion

pair, though this only accounts for ∼3 kJ/mol. A more difficult to quantify contribution is the intrinsic error due to the limitations of the generalized gradient approximation as one of the lower rungs of Perdew's Jacob's Ladder.[62] Certainly, assessments of possible thermochemical errors from GGA functionals would be consistent with the magnitude of the issue here, though free energies in solvent are rarely part of such benchmarks due to the lack of robust data. A specific known weakness of GGA functionals is a tendency to exaggerate delocalization of charge, which could be a problem for an anion such as carbonate where the second electron affinity is positive. To test the influence of charge transfer, free energy profiles were recomputed for calcium carbonate ion pairing based on the rigid ion force field, but with varying amounts of negative charge transferred from the oxygens of carbonate to calcium, as shown in Figure 6. It can be seen that the reduction from formal charges to partial charges indeed destabilizes the contact ion pair relative to the solvent-shared state. Transfer of only 0.24 of an electron would be sufficient to account for the difference between the force field and DFT results.

To further resolve the question as to whether the problem is that the DFT simulations fail to localize the charge on carbonate, or instead whether the force field is too ionic, we have run a limited number of *ab initio* molecular dynamics runs using a hybrid functional. Guidon *et al*[63] demonstrated that the addition of 50% Hartree-Fock exchange to PBE was sufficient to change an excess electron in liquid water from being delocalized to localized. Therefore we have used PBE0, but with twice as much Hartree-Fock exchange to try to sure that a more localized anion charge distribution is sampled. By averaging the Mulliken charges of the carbonate anion over the simulations for both PBE and PBE0(50%) it was found that indeed the carbonate was more negative in the hybrid simulations by the order of 0.25 electrons. Ultimately, the proof that GGA simulations will underestimate the stability of the contact ion pair must come from the free energy. While it is too computationally expensive to sample the full 2D landscape with hybrid functionals, we have used umbrella sampling to map the free energy difference between the contact and solvent-shared ion pair for the case where the water coordination number of Ca^{2+} remains at 5, since this only requires a single collective variable. The comparison between the 1-D free energy profiles for this transition are shown in Figure 7. This demonstrates that the inclusion of a substantial proportion of Hartree-Fock exchange not only helps to localize the negative charge on the carbonate ion, but also stabilizes the contact ion pair by ∼3-4 kJ/mol. Importantly, by equalizing the stability of the contact and solvent-shared ion pair states, this now agrees better with the result of the AMOEBA model.

As a final critical examination of the performance of our new polarizable model, the ion pairing free

18

energy for calcium carbonate can be plotted as a function of temperature for comparison to experiment (Figure 8). Corresponding enthalpy and entropy changes for formation of the $CaCO_3^0$ ion pair have been obtained by a linear fit to the data and are shown in Table 5. The key observation is that the new AMOEBA model is in good agreement with the experimental data across the range of accessible temperatures and consistently improves upon the rigid ion force field result. This improvement cannot be attributed to either a superior enthalpy or entropy of ion pairing, but instead is a consequence of cancellation of errors due to increased values for both quantities in the AMOEBA model. However, it should be noted that there are larger uncertainties on the AMOEBA data due to the greater computational cost leading to slightly shorter run lengths, though the uncertainties are not dissimilar to those on the experimental fits.

Ion pairing for calcium bicarbonate in water

Although there is a strong focus on calcium carbonate ion pairing in water as a precursor to crystallization, it is also important to account for the speciation of bicarbonate as this is often the more numerous anion at typical pH conditions, especially in the environment. Hence we have also generated an AMOEBA-based model for calcium bicarbonate ion pairing. In addition, we have modified our earlier parameterization of a rigid ion force field for bicarbonate to correct deficiencies in both the hydration structure of this anion and the overestimation of ion pair stability. The ion pairing free energy at 300 K as a function of Ca-C distance is shown in Figure 3. Both force fields give very similar profiles, except for the contact ion pair state, with the rigid ion model even reproducing the dehydration barrier from the solvent-shared state, unlike in the case of $CaCO_3^0$. As for the carbonate ion pair, the two force fields give opposing relative stabilities for bidentate versus monodentate coordination, though now in the inverse direction; AMOEBA favors monodentate coordination for the contact ion pair, while the rigid ion model essentially only has a minimum for the bidentate configuration.

Extended comparison of the overall thermodynamics for the $CaHCO_3^+$ ion pair as a function of temperature is shown in Figure 9 with the results of linear fits to determine the enthalpy and entropy components given in Table 6. The plots of free energy show that there is far greater uncertainty regarding the experimental data, especially at elevated temperatures. Here our force field models are more consistent with the slope of the Plummer and Busenberg results, while the other literature values suggest higher entropy changes. Similarly, the computational data more closely aligns with the enthalpy of these authors with the rigid ion and AMOEBA bracketing this experimental value. If we focus on the free energy at ambient

conditions, where there is least uncertainty for all experiments, then the rigid ion and AMOEBA values are less and more exergonic than experiment, respectively. However, the discrepancy is below 2 kJ/mol in both cases, which is less than $k_B T$ at ambient conditions and comparable to the scatter of the experimental values.

Multiple binding equilibria for calcium carbonate

Having validated the AMOEBA model for aqueous calcium carbonate systems it is now possible to apply this to the question of the thermodynamics of multiple binding equilibria. There has been considerable debate as to whether the association of calcium carbonate ion pairs occurs in solution and, if so, how does this vary as the size of the species grows?[57,64] Here we use metadynamics in conjunction with both the AMOEBA and rigid ion force fields to address these questions. Due to the limitations on the number of collective variables that can be effectively sampled, we have examined association processes where either individual cations or anions are added to an ion pair, or a pre-formed ion pair is used as the growth unit. During these reactions, ion pairs are constrained to remain together. Results for the thermodynamics of association for species consisting of up to 4 ion pairs are shown schematically in Figure 10. Overall, it can be seen that the rigid ion and AMOEBA force fields exhibit similar trends, though the former for the large part yields more negative free energies than the latter, in line with the intrinsic stability of a single ion pair.

Starting from the initial ion pair, addition of either Ca^{2+} or CO_3^{2-} is exergonic, though the addition of a further ion of opposite charge to produce the second ion pair is more so. The sum of these values can be compared with the free energy for direct association of two ion pairs. In the case of the rigid ion force field there is good agreement between the results for these two pathways, as would be expected according to Hess's Law, while for the AMOEBA case the poorer statistics mean that there is a greater discrepancy. Regardless of this, the standard free energy for association of two ion pairs is estimated to lie between -5.8 and -12.5 kJ/mol based on the six values obtained from the different models and pathways, though with all bar one value being clustered between -5.8 and -7.7 kJ/mol, suggesting the most exergonic value could be an outlier. For the addition of a further ion pair, the ion by ion additions to $(CaCO_3)_2$ were only computed via the rigid ion force field, leading to values that are similar to those for adding individual ions to the single ion pair. Again the associations of a third or fourth ion pair are found to have favorable standard free energies with values in the range of -4.6 to -10 kJ/mol depending on the force field and pathway. If

we focus on the values for direct addition of a pre-formed ion pair, the exergonic association standard free energy to form larger aggregates appears to be constant with increasing size since any variations lie within the statistical uncertainty of the calculations. Our findings seem to be consistent with the assumption made by Kellermeier *et al*[65] that the association free energy for larger clusters is constant (Table 7), albeit in our case it is smaller than the ion pairing free energy, as was instead assumed in their paper. One of the challenges here is that the structure of the ion association species is highly dynamic, as previously noted for DOLLOP species.[9] Therefore the use of constraints and a limited set of collective variables for sequential addition becomes increasingly problematic as the average cluster size grows. Hence the true free energy may be lower than the value computed here if the constraints were removed.

Previously, we have computed the free energies for ion pairing and ion pair association in a different way that used direct simulation.[9] Here ions were placed at high concentrations in a simulation box and the speciation monitored. Under selected conditions of low pH (i.e. high ratio of bicarbonate to carbonate ions), it was possible to obtain dynamic equilibrium leading to a converged cluster size distribution, despite being heavily supersaturated. Fitting equilibrium constants for a multiple binding model to this system it was possible to extract free energies of -20 and -22 kJ/mol for ion pairing and ion pair association, respectively. While there is reasonable agreement for the ion pair formation thermodynamics (while noting the force field is not exactly the same), our present values for ion pair association are about one third of these previous values. There are several possible reasons for this discrepancy: Firstly, the direct fitting was for very high concentrations and so the free energy relative to the dilute dissociated regime may never be sampled. Secondly, the dielectric constant of the medium will decrease as the concentration of ions increases, though the increased electrostatic screening due to higher ionic strength would outweigh this effect (the $CaCO_3^0$ pairing free energy as a function of the ionic strength is reported in the SI). Thirdly, it was subsequently shown that binodal separation of a dense-liquid phase occurs in this system and the concentration of the simulation lies well past the point where this transformation should take place.[66] As a result, it seems likely that the equilibrium constants may well be more representative of those for the dense-liquid phase, rather than dilute solution as is the case in the present study.

Conclusions

In this study we have derived a new polarizable model for aqueous calcium carbonate based on the more sophisticated electrostatic treatment incorporated within the AMOEBA force field. As a result, this has led to the creation of a means of obtaining improved thermodynamics for the aqueous speciation of this system. In particular, the standard free energy of ion pairing to form $CaCO_3^0$ is significantly improved relative to our earlier rigid ion force field while targeting similar fitting quantities, such as the ion hydration free energies and the solubility of calcite. Not only is the thermodynamics improved, but there are also enhancements in the solvation structure around the species, as well as the transferability between solution and low dielectric environments, such as the gas phase.

Finally, we have used our new model to compute the multiple binding equilibria for calcium carbonate ion pairs in water. In line with our earlier work that used unbiased MD to directly obtain equilibrium constants for association of species, we find the standard free energy for such associations is favorable and largely constant, or perhaps very slightly decreasing, with increasing cluster size. Although the two sets of results are in qualitative agreement, there are quantitative differences with the standard free energies computed here being substantially less exergonic and smaller in magnitude than that for ion pairing itself, rather than comparable to it. That said, the equilibrium constants for association of all species beyond ion pairing alone are greater than 1 and are thus favorable, though because of the typically low concentrations of calcium carbonate solutions (mM) the fraction of higher species will remain small, at least until the binodal is reached. To verify the consistency of the current thermodynamics for multiple binding with experiment we have constructed a speciation model within PHREEQC.[67] In line with experiment and the previous results of Smeets *et al*,[10] we obtain a linear-binding profile over the experimentally accessible range of concentrations.

Acknowledgements

putational Merit Allocation Scheme.

Table 1: Solvation free energies of water, calcium and carbonate in water computed using the AMOEBA force field developed in this work. Experimental values, as well as the results of our earlier rigid ion force field, are also given for comparison. All energies are in kJ/mol.

Solvation free energy	Rigid ion force field	AMOEBA	Experiment
H_2O	-27.6	-23.6	-26.4[68]
Ca^{2+}	-1444	-1493 ± 5	-1444/-1505/-1515/-1527[51–53,69]
CO_3^{2-}	-1315	-1314 ± 3	-1315/-1324[52,53]
HCO_3^-	-324	-346 ± 1	-335/-343/-368[52,53,69]

Table 2: Structural and dynamical properties of the ions in water at 300 K computed from unbiased MD simulations. The uncertainty on the AMOEBA and rigid ion force field simulations is on the last digit shown. The rigid ion force field results for the bicarbonate are for the potential developed in this work while those for calcium and carbonate are for the potential described in Ref. 30. The water potential is SPC/Fw water from Wu et al.[70] The self-diffusion coefficient, D^∞, was computed with three different box sizes (25, 50 and 75Å) and extrapolated to infinite size. The water residence time, τ, was computed by counting the water exchanges during 20 ns long simulations where the frames were recorded every ps; the solvation shell radius was defined as the position of the first minimum of the corresponding pair distribution functions. In the case of HCO_3^- with the AMOEBA forcefield τ varies slightly depending on which oxygen is considered, with the fastest exchange rate observed for the O_c–$H_w O_w$ hydrogen bond and the slowest for the $O_c H_c$–O_w hydrogen bond.

		Experiment	Rigid ion	AMOEBA
Ca^{2+}	r_{Ca-Ow} (Å)	2.33–2.44[71]	2.36	2.37
	N_{Ow}	6–10[71]	7.2	7.2
	D^∞ (10^{-5} cm^2/s)	0.79[72]	0.95	0.9
	τ (ns)	1.1 / 1.6[73]	0.23	0.1
CO_3^{2-}	r_{C-Ow} (Å)	3.35[74]	3.24	3.55
	r_{C-Hw} (Å)	2.68[74]	2.29/2.74	2.63
	N_{Ow}	9.1[74]	12.1	10.7
	D^∞ (10^{-5} cm^2/s)	0.8/0.955[75,76]	1.0	0.71
	τ (ns)	–	0.017	0.03
HCO_3^-	r_{C-Ow} (Å)		3.6	3.6
	r_{O-Ow} (Å)		2.7/2.8	2.9
	r_{O-Hw} (Å)		1.6/1.8	1.8/2.0
	N_{Ow}		10.1	11.3
	D^∞ (10^{-5} cm^2/s)	1.0[77]/1.18[75,76]	1.7	1.5
	τ (ns)		0.004/0.02	0.004/0.01
H_2O	r_{Ow-Ow} (Å)	2.8[78]	2.7	2.8
	N_{Ow}	4.4[78]	4.5	4.6
	D^∞ (10^{-5} cm^2/s)	2.35[79]	2.86	2.36
	τ (ns)	0.005[71]	0.005	0.005

Table 3: Calcite lattice parameters and dissolution free energy computed with a rigid ion force field[30] and the AMOEBA forcefield developed in this work. Experimental data is also included for comparison.

	Rigid ion[30]	AMOEBA	Experiment
a (Å)	4.9398	5.007	4.991[80]
c (Å)	17.1013	17.053	17.062[80]
K (GPa)	85	90	73.5[81]
ΔG_{dis} (kJ/mol)	45.1	46.7	48.1/48.5[82]

Table 4: Comparison of gas phase structural properties and internal energies of binding for clusters. Quantum mechanical data is computed at the ωB97X-D3/ma-def2-QZVPP level of theory. Binding energies are with respect to the sum of all molecular fragments including the calcium ion, water and bicarbonate, as appropriate.

Species	Quantity	QM	Rigid ion	AMOEBA
$Ca(H_2O)^{2+}$	r_{Ca-Ow} (Å)	2.222	2.262	2.186
	BE (kJ/mol)	-239.1	-208.2	-232.6
$Ca(H_2O)_6{}^{2+}$	r_{Ca-Ow} (Å)	2.383	2.321	2.361
	BE (kJ/mol)	-1027.4	-1060.6	-1012.8
$Ca(H_2O)_7{}^{2+}$	r_{Ca-Ow} (Å)	2.397-2.462	2.358-2.399	2.381-2.440
	BE (kJ/mol)	-1103.7	-1141.5	-1083.6
$Ca(HCO_3)^+$	r_{Ca-O5} (Å)	2.133/2.154	2.143/2.184	2.167/2.208
	BE (kJ/mol)	-1319.0	-1115.2	-1282.1
$Ca(HCO_3)(H_2O)_5{}^+$	r_{Ca-O5} (Å)	2.369/2.411	2.305/2.514	2.358/2.504
	BE (kJ/mol)	-1837.6	-1788.4	-1819.6
$(HCO_3)(H_2O)^-$	r_{O5-H2} (Å)	1.613	1.943	1.789
	r_{O2-H6} (Å)	2.231	2.546	2.296
	BE (kJ/mol)	-69.3	-52.4	-59.2
$(HCO_3)(H_2O)_2{}^-$	r_{O5-H2} (Å)	1.701/1.829	1.766/1.842	1.820/1.963
	BE (kJ/mol)	-141.0	-122.7	-133.0

Table 5: $CaCO_3^0$ pairing enthalpy and entropy obtained by a linear fit of the data reported in Figure 9. The errors are the uncertainties obtained from the linear fitting of the data points, while the error for $\Delta G(300K)$ is the 95% confidence interval

	ΔH (kJ/mol)	ΔS (J/mol K)	$\Delta G(300K)$ (kJ/mol)
Plummer & Busenberg[83]	21.0 ± 3	133 ± 9	-19.0 ± 0.9
Kellermeier[8]	8.5 ± 2	88 ± 8	-18.0 ± 0.3
Forcefield[8]	12.4 ± 2	118 ± 5	-23.0 ± 0.3
AMOEBA	28.4 ± 4	152 ± 12	-17.3 ± 0.7

Table 6: $CaHCO_3^+$ pairing enthalpy and entropy obtained by a linear fit of the data reported in Figure 9. The errors are the uncertainties obtained from the linear fitting of the data points, while the error for $\Delta G(300K)$ is the 95% confidence interval

	ΔH (kJ/mol)	ΔS (J/mol K)	$\Delta G(300K)$ (kJ/mol)
Plummer & Busenberg[83]	8.4 ± 1	49 ± 4	-6.2 $\pm$ 0.9
Martynova[84]	20.4 ± 1	93 ± 3	-7.5 $\pm$ 0.3
Jacobson & Langmuir[85]	24.6 ± 1	102 ± 5	-6.0 $\pm$ 0.3
Forcefield	6.0 ± 2	35 ± 7	-4.5 $\pm$ 0.3
AMOEBA	10.3 ± 2	61 ± 7	-8.0 $\pm$ 0.7

Table 7: Formation free energy for $CaCO_3^0$ clusters of increasing sizes compared to the free energy of isolated ion pairs using the average of all possible pathways (Δ Isolated) and per ion pair added to the first one (Δ_n Isolated)

$\Delta_f G$ (kJ/mol)	Rigid ion forcefield			AMOEBA		
	Cluster	Δ Isolated	Δ_n Isolated	Cluster	Δ Isolated	Δ_n Isolated
$CaCO_3^0$	-23.0	0	0	-17.3	0	0
$(CaCO_3^0)_2$	-51.4	-5.4	-5.4	-41.9	-8.8	-8.8
$(CaCO_3^0)_3$	-80.4	-11.4	-5.7	-76.3	-15.2	-7.6
$(CaCO_3^0)_4$	-115.8	-23.8	-7.9	-104.9	-25.2	-8.4
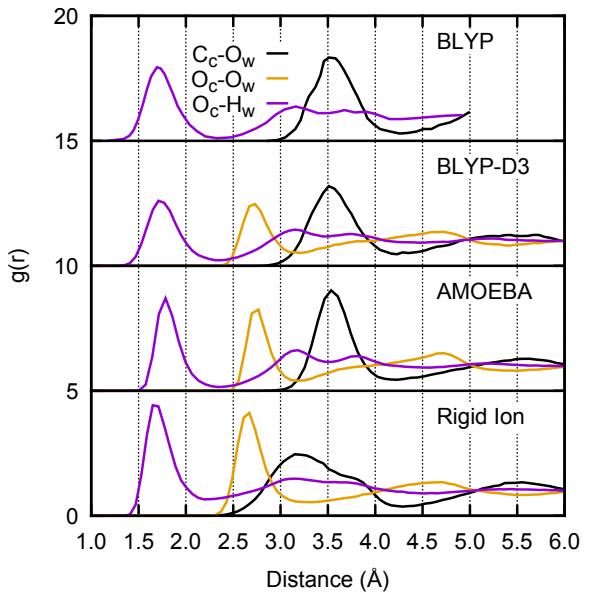
Figure 1: Radial pair distribution functions between carbonate and water computed using the rigid ion force field from Ref. 30, the polarizable AMOEBA force field developed in this work, DFT (BLYP-D3) calculations performed in this work with CP2K and literature DFT (BLYP) calculations performed with CPMD[86]
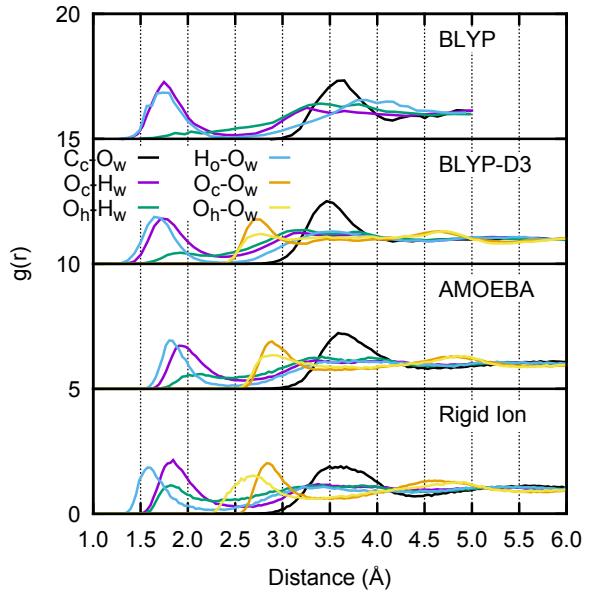
Figure 2: Radial pair distribution functions between bicarbonate and water computed using the rigid ion and the polarizable AMOEBA force field developed in this work, using DFT (BLYP-D3) calculations performed in this work with CP2K and literature DFT (BLYP) calculations performed with CPMD[86]
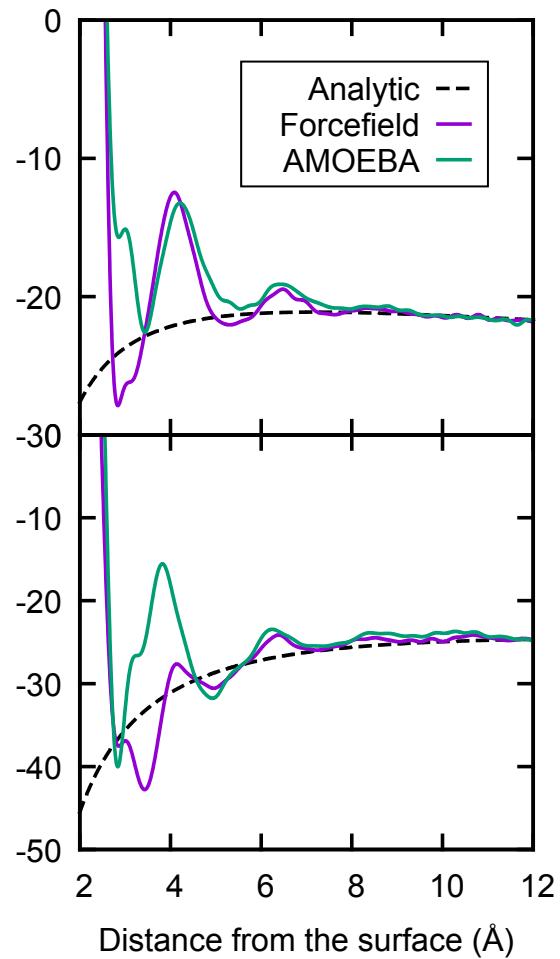
Figure 3: Pairing free energy between calcium and bicarbonate (top) and between calcium and carbonate (bottom) at 300K computed with the rigid ion force field (purple line) and the AMOEBA polarizable model (green). The dashed lines represent the analytic solution for the pairing free energy of two point charges interacting via a screened electrostatic potential.
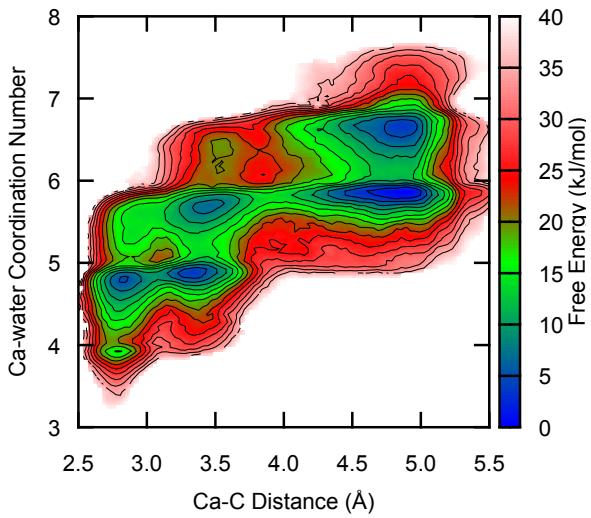
Figure 4: Free energy landscape for ion pairing of calcium carbonate computed at the PBE-D3/TZ2P level of theory. The free energy (in kJ/mol relative to the most stable state) is plotted as a function of two collective variables; the Ca-C distance (x-axis) and the Ca-O(water) coordination number (y-axis). Contour lines are shown at intervals of 2.5 kJ/mol.
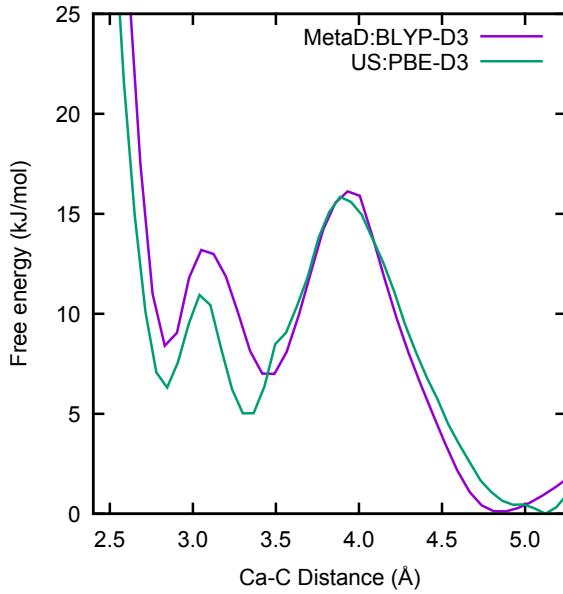
Figure 5: Free energy versus Ca-C distance for calcium carbonate in water in the region spanning the contact and solvent-shared ion pairs. Two curves are shown; one for PBE-D3 computed with 2D umbrella sampling, and a second for BLYP-D3 from multiple walker metadynamics.
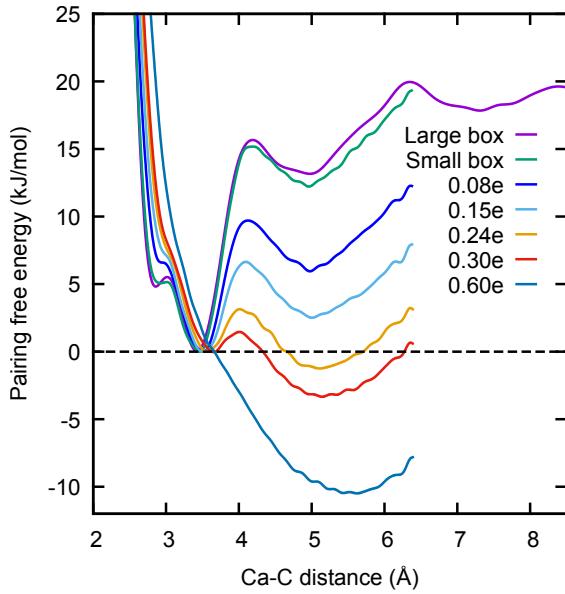
Figure 6: Calcium carbonate pairing free energy computed using the rigid ion forcefield with a partial charge transfer from the carbonate oxygen atoms to the calcium atom. In order to facilitate the comparison of the relative stability of the contact and solvent-shared ion pair states, the zero of free energy curves have been set to that of the monodentate contact ion pair state. The dashed line is the analytic solution and the horizontal black line is just a guide to the eye.
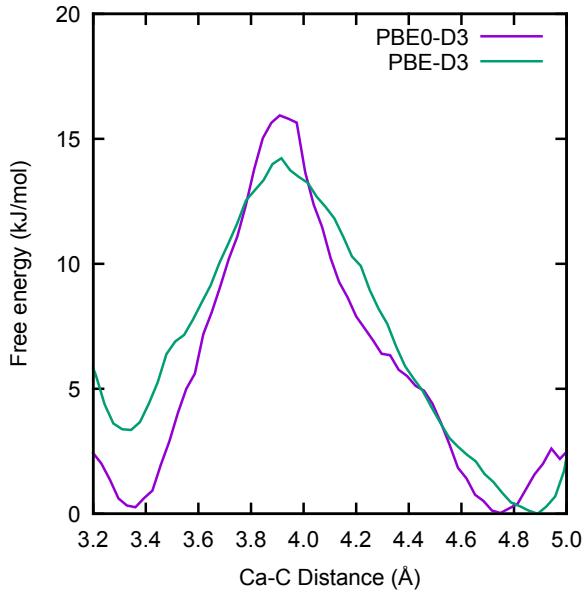
Figure 7: Free energy versus Ca-C distance for calcium carbonate ion pairing in water in the region spanning the barrier between the contact and solvent-shared states. Results at the PBE-D3 and PBE0(50%)-D3 levels of theory are shown as obtained from 1-D umbrella sampling using configurations initially taken from the 2-D sampling at the PBE-D3 level where the Ca water coordination number umbrella was set to 5.
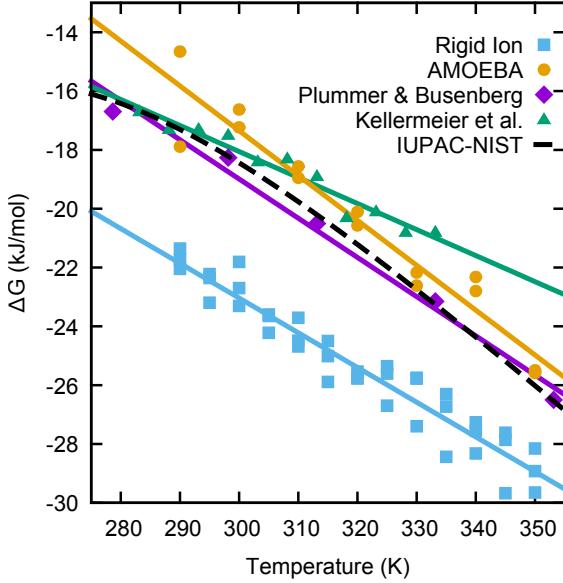
Figure 8: Calcium carbonate pairing free energy as a function of temperature computed with the AMOEBA force field developed in this work (circles). The purple diamonds are the experimental data from Plummer and Busenberg,[83] the blue squares are the simulated pairing free energy using a classical non-polarizable force field developed using the same protocol followed in this work[8,30] and the green circles are the experimental data from Ref. 8. Solid lines have been obtained with a least squares fit of the points, see Table 5. The black dashed line represents the best estimate of the temperature dependence of the pairing free energy from the IUPAC-NIST Solubility Data Series,[82] which was obtained from a critical evaluation of all the experimental data.
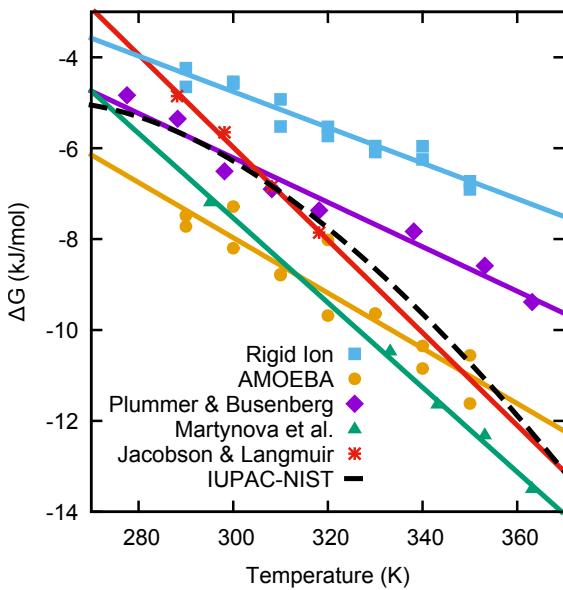
Figure 9: Calcium bicarbonate pairing free energy as a function of temperature computed with the rigid ion and polarizable AMOEBA forcefield developed in this work (squares and circles, respectively) compared with the experimental data from Plummer and Busenberg[83] (diamonds) Martynova et al.[84] (triangles) and Jacobson and Langmuir[85] (stars). The black dashed line represents best estimate of the temperature dependence of the pairing free energy from the IUPAC-NIST Solubility Data Series,[82] which was obtained from a critical evaluation of all the experimental data.
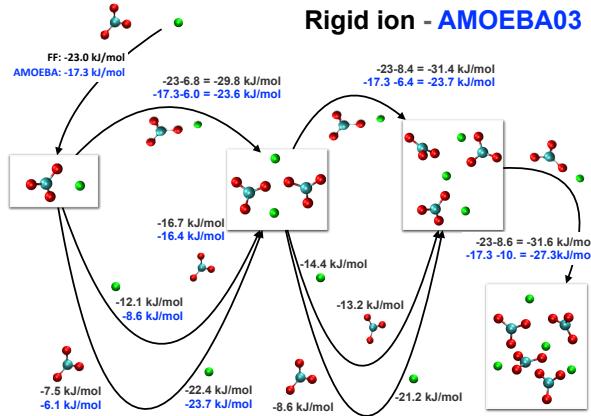
Figure 10: Simulation pathways for the calculation of the stability of calcium carbonate clusters of different sizes. The black numbers have been obtained with the rigid ion force field and the blue ones with the AMOEBA force field developed here. The arrows indicate the addition mechanisms to form the multiple ion clusters, either by direct addition of an ion pair or via a two step process.

References

(1) Lackner, K. S. Carbonate Chemistry for Sequestering Fossil Carbon. *Annual Review of Energy and the Environment* **2002**, *27*, 193–232.

(2) Gebauer, D.; Cölfen, H. Prenucleation clusters and non-classical nucleation. *Nano Today* **2011**, *6*, 564–584.

(3) Wallace, A. F.; Hedges, L. O.; Fernandez-Martinez, A.; Raiteri, P.; Gale, J. D.; Waychunas, G. A.; Whitelam, S.; Banfield, J. F.; De Yoreo, J. J. Microscopic Evidence for Liquid-liquid Separation in Supersaturated CaCO$_3$ Solutions. *Science* **2013**, *341*, 885–889.

(4) Gebauer, D. et al. Proto-calcite and Proto-vaterite in Amorphous Calcium Carbonates. *Angewandte Chemie International Edition* **2010**, *49*, 8889–8891.

(5) Gebauer, D.; Völkel, A.; Cölfen, H. Stable prenucleation calcium carbonate clusters. *Science* **2008**, *322*, 1819–1822.

(6) Tribello, G. A.; Bruneval, F.; Liew, C.; Parrinello, M. A Molecular Dynamics Study of the Early Stages of Calcium Carbonate Growth. *Journal of Physical Chemistry B* **2009**, *113*, 11680–11687.

(7) Raiteri, P.; Gale, J. D.; Quigley, D.; Rodger, P. M. Derivation of an Accurate Force-field for Simulating the Growth of Calcium Carbonate from Aqueous Solution: A New Model for the Calcite-water Interface. *Journal of Physical Chemistry C* **2010**, *114*, 5997–6010.

(8) Kellermeier, M.; Raiteri, P.; Berg, J. K.; Kempter, A.; Gale, J. D.; Gebauer, D. Entropy Drives Calcium Carbonate Ion Association. *ChemPhysChem* **2016**, *17*, 3535–3541.

(9) Demichelis, R.; Raiteri, P.; Gale, J. D.; Quigley, D.; Gebauer, D. Stable Prenucleation Mineral Clusters are Liquid-like Ionic Polymers. *Nature Communications* **2011**, *2*, 590.

(10) Smeets, P. J. M.; Finney, W. J. E. M., Aaron R ad Habraken; Nudelman, F.; Friedrich, H.; Laven, J.; De Yoreo, J. J.; Rodger, P. M.; Sommerdijk, N. A. J. M. A classical View on Nonclassical Nucleation. *Proceedings of the National Academy of Sciences* **2017**, *114*, E7882–E7890.

(11) Ponder, J. W. et al. Current Status of the AMOEBA Polarizable Force Field. *Journal of Physical Chemistry B* **2010**, *114*, 2549–2564.

(12) de Leeuw, N. H.; Parker, S. C. Molecular-dynamics Simulation of MgO Surfaces in Liquid Water Using a Shell-model Potential for Water. *Physical Review B* **1998**, *58*, 13901–13908.

(13) van Maaren, P.; van der Spoel, D. Molecular Dynamics Simulations of Water with Novel Shell-model Potentials. *Journal of Physical Chemistry B* **2001**, *105*, 2618–2626.

(14) Wilson, M.; Madden, P. A.; Costa-Cabral, B. J. Quadrupole Polarization in Simulations of Ionic Systems: Application To AgCl. *Journal of Physical Chemistry* **1996**, *100*, 1227–1237.

(15) van Duin, A. C. T.; Dasgupta, S.; Lorant, L.; Goddard III, W. A. G. ReaxFF: A Reactive Force Field for Hydrocarbons. *Journal of Physical Chemistry A* **2001**, *105*, 9396–9409.

(16) Gale, J. D.; Raiteri, R.; van Duin, A. C. T. A Reactive Force Field for Aqueous-calcium Carbonate Systems. *Physical Chemistry Chemical Physics* **2011**, *13*, 16666–16679.

(17) Ren, P.; Ponder, J. W. Constitent Treatment of Inter- and Intramolecular Polarization in Molecular Mechanics Calculations. *Journal of Computational Chemistry* **2002**, *23*, 1497–1506.

(18) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; J, C.; P, E.; J, M. T.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *Journal of Physical Chemistry B* **2013**, *117*, 9956–9972.

(19) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation . *Journal of Physical Chemistry B* **2003**, *107*, 5933–5947.

(20) Laury, M. L.; Wang, L.-P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W. Revised Parameters for the AMOEBA Polarizable Atomic Multipole Water Model. *Journal of Physical Chemistry B* **2015**, *119*, 9423–9437.

(21) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comp.* **2005**, *1*, 1128–1132.

(22) Neese, F. The ORCA Program System. *Comput. Mol. Sci.* **2012**, *2*, 73–78.

(23) Lin, Y.-S.; Li, G.-D.; S-P, M.; Chai, J.-D. Long-range Corrected Hybrid Density Functionals with Improved Dispersion Corrections. *J. Chem. Theory Comp.* **2013**, *9*, 263–272.

(24) Gale, J. D. GULP: Capabilities and Prospects. *Zeitschrift Fur Kristallographie* **2005**, *220*, 552–554.

(25) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *Journal of Chemical Physics* **1935**, *3*, 300.

(26) Bennett, C. H. Efficient Estimation of Free-Energy Differences From Monte-Carlo Data. *Journal of Computational Physics* **1976**, *22*, 245–268.

(27) Lagardère, L. et al. Tinker-HP: a Massively Parallel Molecular Dynamics Package for Multiscale Simulations of Large Complex Systems with Advanced Point Dipole Polarizable Force Fields. *Chem. Sci.* **2018**, *9*, 956–972.

(28) Hummer, G.; Pratt, L. R.; Garcia, A. Ion sizes and finite-size corrections for ionic-solvation free energies. *Journal of Chemical Physics* **1997**, *107*, 9275–9277.

(29) Plimpton, S. Fast Parallel Algorithms for Short-range Molecular Dynamics. *Journal of Computational Physics* **1995**, *117*, 1–19.

(30) Raiteri, P.; Demichelis, R.; Gale, J. D. Thermodynamically Consistent Force Field for Molecular Dynamics Simulations of Alkaline-Earth Carbonates and Their Aqueous Speciation. *Journal of Physical Chemistry C* **2015**, *119*, 24447–24458.

(31) Eastman, P. et al. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Computational Biology* **2017**, *13*, e1005659–17.

(32) Yeh, I.; Hummer, G. System-Size Dependence of Diffusion Coefficients and Viscosities From Molecular Dynamics Simulations with Periodic Boundary Conditions. *Journal of Physical Chemistry B* **2004**, *108*, 15873–15879.

(33) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: Atomistic Simulations of Condensed Matter Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.

(34) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Comm.* **2005**, *167*, 103–128.

(35) Becke, A. Density-functional Thermochemistry. I. The Effect of the Exchange-only Gradient Correction. *J. Chem. Phys.* **1992**, *96*, 2155–2160.

(36) Lee, C.; Yang, W.; G, P. R. Development of the Colle-Salvetti Correlation-energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785–789.

(37) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate ab initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(38) Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710.

(39) VandeVondele, J.; Hutter, J. An Efficient Orbital Transformation Method for Electronic Structure Calculations. *J. Chem. Phys.* **2003**, *118*, 4365–4369.

(40) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *Journal of Chemical Physics* **2007**, *126*, 014101.

(41) Friedrichs, M. S.; Eastman, P.; Vaidyanathan, V.; Houston, M.; Legrand, S.; Beberg, A. L.; Ensign, D. L.; Bruns, C. M.; Pande, V. S. Accelerating Molecular Dynamic Simulation on Graphics Processing Units. *Journal of Computational Chemistry* **2009**, *30*, 864–872.

(42) Eastman, P.; Pande, V. S. Efficient Nonbonded Interactions for Molecular Dynamics on a Graphics Processing Unit . *Journal of Computational Chemistry* **2009**, *30*, NA–NA.

(43) Bonomi, M. et al. Promoting Transparency and Reproducibility in Enhanced Molecular Simulations. *Nature Methods* **2019**, *16*, 670–673.

(44) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird . *Computer Physics Communications* **2014**, *185*, 604–613.

(45) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proceedings of The National Academy of Sciences of The United States of America* **2002**, *99*, 12562–12566.

(46) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics. *Journal of Physical Chemistry B* **2006**, *110*, 3533–3539.

(47) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Physical Review Letters* **2008**, *100*, 20603.

(48) Garcia, N. A.; Innocenti Malini, R.; Freeman, C. L.; Demichelis, R.; Raiteri, P.; Sommerdijk, N. A. J. M.; Harding, J. H.; Gale, J. D. Simulation of Calcium Phosphate Prenucleation Clusters in Aqueous Solution: Association beyond Ion Pairing. *Crystal Growth & Design* **2019**, 1–9.

(49) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. Molecular Dynamics Simulations of Water and Biomolecules with a Monte Carlo Constant Pressure Algorithm . *Chemical Physics Letters* **2004**, *384*, 288–294.

(50) Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. Simulation of Ca^{2+} and Mg^{2+} Solvation Using Polarizable Atomic Multipole Potential. *Journal of Physical Chemistry B* **2006**, *110*, 18553–18559.

(51) David, F.; Vokhmin, V.; Ionova, G. Water Characteristics Depend on the Ionic Environment. Thermodynamics and Modelisation of the Aquo Ions. *Journal of Molecular Liquids* **2001**, *90*, 45–62.

(52) Marcus, Y. *Ions in Solution and Their solvation*; John Wiley & Sons, Inc., 2016.

(53) Marcus, Y. *Ion Solvation*; Wiley, 1985.

(54) De La Pierre, M.; Raiteri, P.; Gale, J. D. Structure and Dynamics of Water at Step Edges on the Calcite 10-14 Surface. *Crystal Growth & Design* **2016**, *16*, 5907–5914.

(55) Garrels, R. A.; Thompson, M. E.; Siever, R. Control of Carbonate Solubility by Carbonate Complexes. *American Journal of Science* **1961**, *259*, 24–45.

(56) Genovese, D.; Montalti, M.; Otalora, F.; Gomez-Morales, J.; Sancho-Tomas, M.; Falini, G.; Garcia-Ruiz, J. M. Role of $CaCO_3^0$ Neutral Pair in Calcium Carbonate Crystallization. *Crystal Growth and Design* **2016**, *16*, 4173–4177.

(57) Henzler, K. et al. Supersaturated Calcium Carbonates Are Classical. *Science Advances* **2018**, *4*, eaao6283.

(58) Yadav, S.; Chandra, A. Structural and Dynamical Nature of Hydration Shells of the Carbonate Ion in Water: An ab initio Molecular Dynamics Study. *Journal of Physical Chemistry B* **2018**, *122*, 1495–1504.

(59) Schwenk, C. F.; Rode, B. M. Ab initio QM/MM MD Simulations of the Hydrated Ca^{2+} ion. *Pure Applied Chemistry* **2004**, *76*, 37–47.

(60) Koskamp, J. A.; Ruiz-Hernandez, S. E.; Di Tommaso, D.; Elena, A. M.; de Leeuw, N. H.; Wolthers, M. Reconsidering calcium Dehydration as the Rate-determining Step in Calcium Mineral Growth. *Journal of Physical Chemistry C* **2019**, *123*, 26895–26903.

(61) Bankura, A.; Karmakar, A.; Carnevale, V.; Chandra, A.; Klein, M. L. Structure, Dynamics, and Spectral Diffusion of Water from First-principles Molecular Dynamics. *Journal of Physical Chemistry C* **2014**, *118*, 29401–29411.

(62) Perdew, J. P.; Schmidt, K. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conference Proceedings* **2001**, *577*, 1–20.

(63) Guidon, M.; Hutter, J.; VandeVondele, J. Auxiliary Density Matrix Methods for Hartree-Fock Exchange Calculations. *Journal of Chemical Theory and Computation* **2010**, *6*, 2348–2364.

(64) Gebauer, D.; Raiteri, P.; Gale, J. D.; Cölfen, H. On Classical and Non-classical Views on Nucleation. *American Journal of Science* **2018**, *318*, 969–988.

(65) Kellermeier, M.; Picker, A.; Kempter, A.; Cölfen, H.; Gebauer, D. A Straightforward Treatment of Activity in Aqueous CaCO 3Solutions and the Consequences for Nucleation Theory. *Advanced Materials* **2014**, *26*, 752–757.

(66) Sebastiani, F.; Wolf, S. L. P.; Born, B.; Luong, T. Q.; Cölfen, H.; Gebauer, D.; Havenith, M. Water Dynamics from THz Spectroscopy Reveal the Locus of a Liquid-liquid Binodal Limit in Aqueous $CaCO_3$ Solutions. *Angewandte Chemie International Edition* **2017**, *56*, 490–495.

(67) Charlton, S. R.; Parkhurst, D. L. Modules Based on the Geochemical Model PHREEQC for Use in Scripting and Programming Languages. *Computers and Geosciences* **2011**, *37*, 1653–1663.

(68) Ben-Naim, A.; Marcus, Y. Solvation Thermodynamics of Nonionic Solutes. *Journal of Chemical Physics* **1984**, *81*, 2016–13.

(69) Marcus, Y. A Simple Empirical Model Describing the Thermodynamics of Hydration of Ions of Widely Varying Charges, Sizes, and Shapes. *Biophysical chemistry* **1994**, *51*, 111–127.

(70) Wu, Y.; Tepper, H.; Voth, G. A. Flexible Simple Point-charge Water Model With Improved Liquid-state Properties. *Journal of Chemical Physics* **2006**, *124*, 024503.

(71) Ohtaki, H.; Radnai, T. Structure and Dynamics of Hydrated Ions. *Chemical Reviews* **1993**, *93*, 1157–1204.

(72) Wang, J. H. Tracer-diffusion in Liquids. IV. Self-diffusion of Calcium Ion and Chloride Ion in Aqueous Calcium Chloride Solutions. *Journal of the American Chemical Society* **1953**, *75*, 1769–1770.

(73) Atkinson, G.; Emara, M. M.; Fernandez-Prini, R. Ultrasonic Absorption in Aqueous Solutions of Calcium Acetate and Other Bivalent Metal Acetates. *Journal of Physical Chemistry* **1974**, *78*, 1913–1917.

(74) Kameda, Y.; Sasaki, M.; Hino, S.; Amo, Y.; Usuki, T. Neutron Diffraction Study on the Hydration Structure of Carbonate Ion by Means of $^{12}C/^{13}C$ Isotopic Substitution Method. *Physica B: Condensed Matter* **2006**, *385-386*, 279–281.

(75) Kigoshi, K.; Hashitani, T. The Self-Diffusion Coefficients of Carbon Dioxide, Hydrogen Carbonate Ions and Carbonate Ions in Aqueous Solutions . *Bulletin of the Chemical Society of Japan* **1963**, *36*, 1372.

(76) Li, Y.; Gregory, S. Diffusion of Ions in Sea Water and in Deep-sea Sediments. *Geochimica et Cosmochimica Acta* **1974**, *38*, 703–714.

(77) Poisson, A.; Papaud, A. Diffusion-Coefficients of Major Ions in Seawater. *Marine Chemistry* **1983**, *13*, 265–280.

(78) Soper, A. K. The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure? *ISRN Physical Chemistry* **2013**, *2013*, 1–67.

(79) Kell, G. S. Precise Representation of Volume Properties of Water at One Atmosphere . *Journal of Chemical & Engineering Data* **1967**, *12*, 66–69.

(80) Maslen, E. N.; Streltsov, V. A.; Streltsova, N. R. X-ray Study of the Electron Density in Calcite, $CaCO_3$. *Acta Crystallographica Section B Structural Science* **1993**, *B49*, 636–641.

(81) Redfern, S. A. T.; Angel, R. J. High-pressure Behaviour and Equation of State of Calcite, $CaCO_3$. *Contributions in Mineralogy and Petrology* **1999**, *134*, 102–106.

(82) De Visscher, A.; Vanderdeelen, J.; Königsberger, E.; Churagulov, B. R.; Ichikuni, M.; Tsurumi, M. IUPAC-NIST Solubility Data Series. 95. Alkaline Earth Carbonates in Aqueous Systems. Part 1. Introduction, Be and Mg. *Journal of Physical and Chemical Reference Data* **2012**, *41*, 013105.

(83) Plummer, L. N.; Busenberg, E. The Solubilities of Calcite, Aragonite and Vaterite in H_2O-Co_2 Solutions Between 0 and 90C, and an Evaluation of the Aqueous Model for the System $CaCO_3$-Co_2-H_2O. *Geochimica et Cosmochimica Acta* **1982**, *46*, 1011–1040.

(84) Martynova, O. I.; Vasina, L. G.; Pozdniakova, S. A. Dissociation Constants of Certain Scale Forming Salts. *Desalination* **1974**, *15*, 259–265.

(85) Jacobson, R. L.; Langmuir, D. Dissociation-Constants of Calcite and $CaHCO_3^+$ From 0 to 50 Degrees C. *Geochimica et Cosmochimica Acta* **1974**, *38*, 301–318.

(86) Kumar, P. P.; Kalinichev, A. G.; Kirkpatrick, R. J. Hydrogen-Bonding Structure and Dynamics of Aqueous Carbonate Species from Car-Parrinello Molecular Dynamics Simulations. *Journal of Physical Chemistry B* **2009**, *113*, 794–802.