This is a repository copy of *Amino acid and oligopeptide effects on calcium carbonate solutions*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/160476/

Version: Published Version

**Article:**

Article

# Amino Acid and Oligopeptide Effects on Calcium Carbonate Solutions

Aaron R. Finney,* Riccardo Innocenti Malini, Colin L. Freeman, and John H. Harding

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Biological organisms display sophisticated control of nucleation and crystallization of minerals. In order to mimic living systems, deciphering the mechanisms by which organic molecules control the formation of mineral phases from solution is a key step. We have used computer simulations to investigate the effects of the amino acids arginine, aspartic acid, and glycine on species that form in solutions of calcium carbonate ($CaCO_3$) at lower and higher levels of supersaturation. This provides net positive, negative, and neutral additives. In addition, we have prepared simulations containing hexapeptides of the amino acids to consider the effect of additive size on the solution species. We find that additives have limited impact on the formation of extended, liquid-like $CaCO_3$ networks in supersaturated solutions. Additives control the amount of (bi)carbonate in solution, but more importantly, they are able to stabilize these networks on the time scales of the simulations. This is achieved by coordinating the networks and assembled additive clusters in solutions. The association leads to subtle changes in the coordination of $CaCO_3$ and reduced mobility of the cations. We find that the number of solute association sites and the size and topology of the additives are more important than their net charge. Our results help to understand why polymer additives are so effective at stabilizing dense liquid $CaCO_3$ phases.
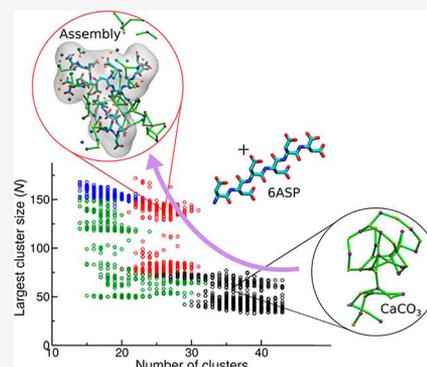
## INTRODUCTION

While there have been major advances in the production of synthetic, bioinspired materials in recent years,[1−3] the degree of control that biological organisms can exercise in engineering hard, mineralized tissues with distinct functionalities and morphologies far surpasses anything available using current technologies. Three crystalline polymorphs of calcium carbonate are found in biology: calcite, aragonite and vaterite, in order of decreasing thermodynamic stability.[4−6] These ordered materials are often formed after the initial deposition of amorphous calcium carbonate (ACC)[7,8] (which was suggested in ref 9 to exhibit short-range coordination geometries reminiscent of aragonite or calcite). Experiments examining the growth of sea urchin spicules[10,11] and corals[12] have shown that calcite and aragonite formation is preceded by hydrated ACC phases, which dehydrate before crystallization.

Organic molecules are found in mineralized tissues throughout biology, usually in the form of an organic matrix or as inclusions.[4] For example, ∼5 wt % of nacre is made from proteins and polysaccharides that glue hexagonal aragonite tablets together and give the material increased toughness and impact resistance.[13] In calcitic materials, the organic matter can inhibit crack propagation.[14] However, the organic material is not present simply to improve the mechanical function of biominerals. Proteins and peptides can control nucleation and crystal growth processes.[9,15] Certain proteins, such as the sea urchin matrix protein LSM34, are essential for the formation of calcite spicules.[15−17] Simulations have shown that the avian egg-

shell protein, Ovocleidin-17, hastens the formation of calcite from initially amorphous nanoparticles.[18] Proteins and peptides adsorbed onto a β-chitin matrix are thought to selectively nucleate nacreous aragonite in mollusk shells.[19,20] Intrinsically disordered proteins (IDPs), which are present in nacre, self-assemble during mineralization and selectively precipitate vaterite or aragonite.[21,22] The inherent conformational freedom associated with IDPs appears to modulate the mechanisms associated with both mineralization and the resulting mineral phases.[23−26] Protein control of the amount of impurities, such as Mg(II), in amorphous phases may be an important step in aragonite polymorph selection.[27] Inspired by these examples in nature, synthetic additives have been designed and applied to control the morphology and rate of crystal growth.[28,29] Other synthetic additives appear to inhibit crystallization. Polyaspartic acid (Poly(ASP)) in solution, for example, retarded the transformation of amorphous liquid and solid phases of calcium carbonate into crystalline ones.[30−32] Although a wealth of research has been dedicated to understanding how organic

additives control the nucleation and growth of calcium carbonate, these mechanisms remain surprisingly unclear.

De Yoreo et al.[33] presented possible pathways to crystals from initially free ions in solution that may involve stable or metastable particles. In the case of calcium carbonate, nanometer-sized clusters,[34,35] liquids,[30,36−38] solid amorphous[10−12,39] and crystalline[40,41] particles have all been presented in pathways to the formation of macroscopic crystalline materials. The initial formation of a $CaCO_3$ phase from ions in solution has been intensely investigated. Several studies[37,42,43] support the notion that, at mild supersaturation, the solution speciation is consistent with a classical picture of nucleation. The solution contains free $Ca^{2+}$ and $CO_3^{2-}$ ions, ion pairs, and occasionally larger ion clusters, which form stochastically but are ultimately transient, unstable species. Alternatively, prenucleation clusters (PNCs) of ions were suggested[30,34,44] as the thermodynamically stable solute species. Simulations revealed[35,45,46] that these clusters have a dynamic size distribution and that their constituent ions have a high degree of translational and rotational freedom. The prenucleation cluster pathway proposes that, on increasing supersaturation, the internal mobility of ions within PNCs decreases and an interface is formed with the solution. The resulting species aggregate to minimize surface tension and form a macroscopic amorphous phase.[32]

We can study the control exerted by proteins and peptides on nucleation and crystallization by investigating the activities of their fundamental building blocks: amino acids. Picker et al.[47] performed a complete survey of the effect of the 20 protein-forming amino acids on various aspects of calcium carbonate nucleation in titration experiments at pH 9.75. Aspartic acid (ASP) influences all stages of nucleation by increasing the amount of bound calcium before nucleation, increasing the induction time for nucleation, and favoring the growth of calcite by inducing a less-stable intermediate phase. Glutamic acid affected the amount of free calcium in solution before nucleation the most, while serine and ASP were best at delaying nucleation. Glycine (GLY) and ASP were the only amino acids that significantly affected the solubility of nucleated phases.[47] Simulations performed on ASP suggest that this amino acid stabilizes amorphous $CaCO_3$ nanoparticles.[45,46] Both citrate and ASP delay nucleation, but not necessarily by complexation to cations.[46] Simulations suggest that the binding of ASP (in the protonation state adopted here) to ACC is through the amine group rather than the carboxylate group as might be expected.[48] Indeed, other advanced sampling simulations suggest that the binding energies of calcium carbonate ion pairs to guanidinium and carboxylate groups are approximately equal.[49] Experiments confirmed that both ASP and GLY extend the lifetime of amorphous precursors.[50] Adsorption of amino acids to growing crystalline surfaces can influence both the morphology and phase of the crystalline calcium carbonate produced.[51,52]

The chemical activity of additives may be key to controlling the nucleation and growth of calcium carbonate. However, particularly for templated growth, we might also expect the size and topology of any organic matrix to be as important for nucleation and growth control. This has been shown for calcium oxalate growth, where oligopeptide chain lengths from 5 to 20 monomers of glutamic acid clearly affected the stabilities and morphologies of the resulting crystalline phases.[53] Simulations showed that oligo-glutamates organized cations in solution before nucleation. This may be the cause of the dependence of crystal morphologies on oligomer chain length.[54] Experimental

studies of artificial peptide additives suggest that the peptide's charge and folding properties in solution control crystallization.[55,56] Molecular simulations revealed that the strength of adsorption of the peptide/polymer to calcite is influenced by the flexibility of the molecule backbone. This may have consequences for crystal templating.[57]

Computer simulations have provided much insight into the early stages of phase separation in a range of systems.[58−61] Here, we report molecular dynamics (MD) simulations that investigate the effect of three amino acids (AAs) on calcium carbonate aqueous solutions. These were arginine (ARG), ASP, and GLY, the chemical structures for which are provided in Figure 1. The protonation state for each molecule was chosen for
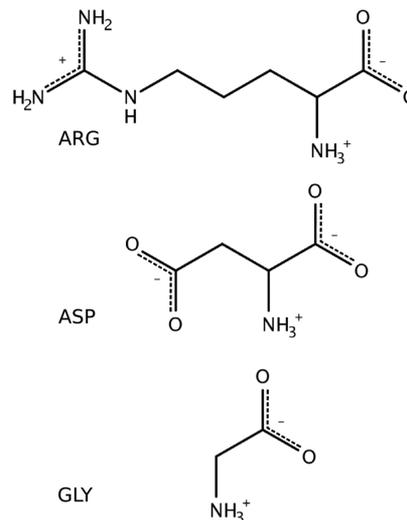


**Figure 1.** Amino acid molecules in the speciation state adopted in this work.

a solution where the pH was ∼9. This provided molecules that were positively charged (ARG), negatively charged (ASP), and neutral (GLY) overall. In addition, we explored the effect of additive size by simulating oligomers of the three AAs containing six monomers bonded through peptide links, i.e., hexapeptides of ARG, ASP, and GLY which we label 6ARG, 6ASP, and 6GLY, respectively.

## ■ COMPUTATIONAL METHODS

Typically, experiments studying $CaCO_3$ nucleation are performed under basic conditions and at millimolar ion concentrations. Limited simulation times mean that we are unlikely to observe the full effect of additives on $CaCO_3$ species during nucleation under these conditions. Therefore, we have simulated ion concentrations that approach the experimental supersaturations and also higher supersaturations where earlier simulations[35,36] showed that liquid-like ionic networks are maintained in the steady state. Simulations will be referred to as low salinity or high salinity and indicated by a subscript "LS" and "HS", respectively, in system labels.

Table 1 provides the initial composition for all systems. The number of carbonate and bicarbonate ions were chosen to neutralize the total calcium charge, setting $c(CO_3^{2-})/c(HCO_3^-)$ to give a pH of 10.3, according to the Henderson−Hasselbalch equation. At this pH, the $NH_3^+$ groups of AAs become susceptible to deprotonation. Our earlier simulations initiated with a fixed number of ions showed consistent reductions in pH (cf. the initial configurations) because of preferential binding of carbonate to calcium. The simulations reported here indicate that the steady-state pH of solutions at low salinity was 9.4−9.8; hence, the adopted speciation states for the AAs are sensible. Oligopeptides

**Table 1. System Composition for Low and High Salinity Simulations**[a]

| system | $n(Na^+)$ | $n(Cl^-)$ | $n(add)$ | $n(Ca^{2+})$ | $n(CO_3^{2-})$ | $n(HCO_3^-)$ | $n(H_2O)$ | $c(Ca^{2+})$ [mM] | $c(C)$ [mM] | $c(add)$ [mM] |
|---|---|---|---|---|---|---|---|---|---|---|
| $CON_{LS}$ | 23 | 23 | 0 | 47 | 30 | 34 | 100 000 | 26 | 36 | 0 |
| $ARG_{LS}$ | 5 | 23 | 18 | 47 | 30 | 34 | 100 000 | 26 | 36 | 10 |
| $ASP_{LS}$ | 23 | 5 | 18 | 47 | 30 | 34 | 100 000 | 26 | 36 | 10 |
| $GLY_{LS}$ | 23 | 23 | 18 | 47 | 30 | 34 | 100 000 | 26 | 36 | 10 |
| $CON_{HS}$ | 23 | 23 | 0 | 100 | 64 | 72 | 10 000 | 560 | 760 | 0 |
| $ARG_{HS}$ | 5 | 23 | 18 | 100 | 64 | 72 | 10 000 | 560 | 760 | 100 |
| $ASP_{HS}$ | 23 | 5 | 18 | 100 | 64 | 72 | 10 000 | 560 | 760 | 100 |
| $GLY_{HS}$ | 23 | 23 | 18 | 100 | 64 | 72 | 10 000 | 560 | 760 | 100 |
| $6ARG_{HS}$ | 5 | 23 | 3 | 100 | 64 | 72 | 10 000 | 560 | 760 | 17 |
| $6ASP_{HS}$ | 23 | 5 | 3 | 100 | 64 | 72 | 10 000 | 560 | 760 | 17 |
| $6GLY_{HS}$ | 23 | 23 | 3 | 100 | 64 | 72 | 10 000 | 560 | 760 | 17 |

[a]$n(X)$ and $c(X)$ indicate the number and concentration, respectively, of ion X.

were formed as a peptide chain of six AAs and the concentration of oligopeptides was a fraction of the AA concentration to ensure that we retained the same numbers of end-chain organic functional groups. Na and Cl ion concentrations were chosen to ensure that the ionic strength ($I$) in all simulations was 0.11 M (low salinity) and 2.2 M (high salinity) at $t = 0$, and that the total charge in simulation cells was zero. Ions were initially dispersed in solution by randomly populating cubic cells before adding water.

The force field of Demichelis et al.[35] was employed in the current work. This accurately captures the energetics of ion solvation and uses flexible SPC/Fw water.[62] The intramolecular interactions for L-amino acids and L-oligopeptides were modeled using the (ff14SB) AMBER force field.[63] Accurately capturing the physical interactions between atoms at the organic/inorganic interface when combining existing force fields is not straightforward. Simple mixing rules (e.g., Lorentz–Berthelot mixing) offer an inexpensive route to generating intermolecular potentials but often lead to incorrect forces between atoms.[64,65] These rules rely on the balance of short-range van der Waals/Pauli repulsion and long-range electrostatic force contributions being approximately the same in both the adopted inorganic and organic force fields. This is rarely the case for mineral-organic systems. Fitting classical potentials to data from ab initio calculations is preferable but is computationally demanding enough to render a study of multi-component systems intractable over reasonable time scales. Here, we used a generic procedure to fit the intermolecular terms between the Demichelis and AMBER force fields that has been shown to reproduce, for example, the correct adsorption energies for the binding of organic molecules to mineral surfaces, when compared to ab initio calculations. Intermolecular potentials between the organic atoms and all other solute atoms were fitted using this "Schröder method", as described by Freeman et al.,[66] while the organic–water interaction parameters were calculated using the standard Lorentz–Berthelot rules. Example force field input files are included in the Supporting Information (SI).

The DL_POLY 4.03 MD simulation package[67,68] was used throughout. Simulations were run with the velocity Verlet algorithm and a time step of 1 fs (configurations were stored every 1 ps). Simulations were performed at 298 K and 1 atm using a Nosé–Hoover[69] thermostat and barostat with 0.1 and 1.0 ps relaxation times, respectively. The number of atoms and chemical bonds were fixed in the simulations. Electrostatic energies were calculated using the smooth particle mesh Ewald method,[70] and three-dimensional periodic boundaries were implemented. A 200 ps MD simulation was performed with all solute atoms frozen to equilibrate the box volume. The final configuration was used to initiate MD simulations where all atoms were mobile. The solute cluster sizes, free ion concentrations, and coordination numbers were monitored in order to identify when a steady state had been achieved. A total of 20 and 60 ns trajectories—with the final 2 or 10 ns used to calculate equilibrium averages—were generated for low- and high-salinity systems, respectively.

The potential of mean force (PMF) for monomer–monomer association in solution was measured using umbrella sampling[71] (US) with the PLUMED plugin.[72] Approximately 50 US windows were used

to sample a monomer separation reaction coordinate from 1.5−24 Å in 0.5 Å steps, and extra windows were added to sample high-energy regions of the reaction coordinate accurately. A simple harmonic potential, with a force constant of at least 12 $k_B N_A T$ (where $k_B$ is the Boltzmann constant, $N_A$ is Avogadro's constant, and $T$ is temperature), was applied to restrain the distance between monomers. Simulations in each US window were performed for at least 1 ns and the WHAM program[73] was used to calculate free-energy changes in the reaction coordinate space. Probability distributions were monitored to ensure proper sampling of the reaction coordinate between adjacent US windows. These simulations were performed at 298 K and 1 atm and monomer concentrations were 10 mM. Na and Cl ions (also at 10 mM) were added to maintain a background electrolyte concentration. Coordination numbers between specific atom types were monitored using the function defined in PLUMED, which goes smoothly from one to zero over a range of atom separation distances. This was set using the region of the first and second minima in radial pair distribution functions to define the switching function parameters.

Clusters were defined as collections of solute ions or molecules where interatomic distances were within 3.5 Å. The position of the first peak in the Ca−O(H₂) radial distribution functions (RDFs) for free ions and ions in liquidlike amorphous calcium carbonate is less than this distance (at ∼2.5 Å). The cutoff adopted ensures that fluctuations in the number and size of clusters are minimized. Clusters were defined in terms of the number of ions or molecules ($N$) and cluster size probability distributions ($f(N)$) were calculated. In addition, the average number of clusters where $N = 1$ allowed the calculation of free ion/molecule concentrations, ionic strengths, and pH at equilibrium. This approach provides a liberal estimate for the number of "free" ions or molecules in solution, given that solvent-separated and solvent-shared coordinated ions are considered to be free. Self-diffusion coefficients ($D$) were calculated using Einstein's relation:

$$D = \frac{1}{6} \lim_{t \to \infty} \frac{1}{N} \sum_1^N [r(t) - r(0)]^2 \tag{1}$$

where $N$ is the total number of atoms of interest in a cluster and $r(t)$ are the Cartesian coordinates as a function of time $t$. We calculated the value of $D$ for the center of mass of clusters and subtracted this from the diffusion rate of ions to measure the diffusion of ions and molecules within clusters. The free-energy changes for ion pairing in low-salinity simulations were estimated using the activity of the ions once a steady state was reached. The activity coefficient ($\gamma$) values were obtained using the extended Debye−Hückel equation:

$$-\log(\gamma) = \frac{Az^2\sqrt{I}}{1 + Ba\sqrt{I}} \tag{2}$$

This model for ion activities was chosen because the $I$ values at the steady state were of the order of tens of millimoles. Here, $A = 1.82 \times 10^6(\varepsilon T)^{-3/2}$ and $B = 50.3(\varepsilon T)^{-1/2}$, where $\varepsilon$ is the relative permittivity of SPC/Fw water (79.63) and $T$ is the temperature (298 K) of the
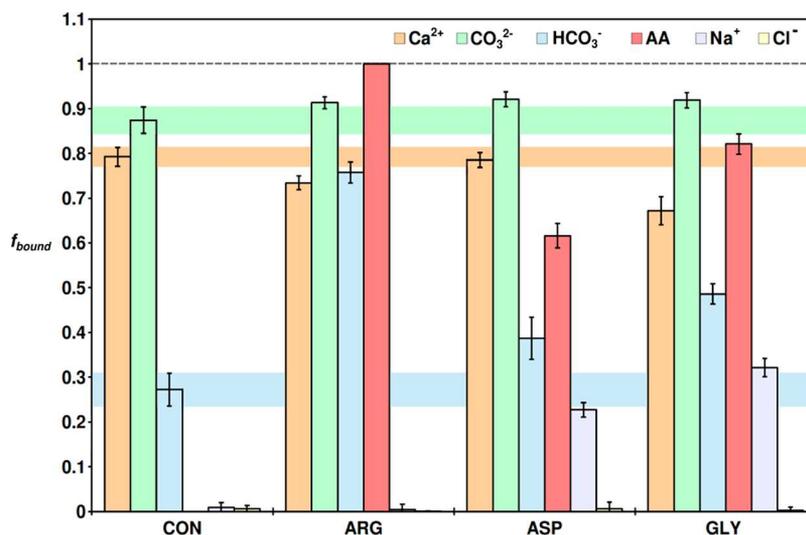
**Figure 2.** Average fraction of bound ions, $f_{bound}$, in the steady state taken from the final 2 ns of total simulation time. Error bars show uncertainties in the data of one standard deviation in the trajectory window sampled. The shaded regions indicate the fraction of bound $Ca^{2+}$, $CO_3^{2-}$ and $HCO_3^-$ in $CON_{LS}$ for ease of comparison to systems containing amino acids.

simulations. $z$ is the magnitude of the ionic charge, and $a$ is a measure of the radii of solvated ions.

## RESULTS AND DISCUSSION

**Low Salinity.** Approximately 10 ns of simulation time were required to reach a steady state in systems at low salinity. The concentrations of free ions at the steady state (see Table S1 in the SI) decreased in all systems, because of ion binding in solution. This was reflected in a change of ionic strength, which decreased to 25%−30% of the initial value. The concentration of calcium ions, $c(Ca^{2+})$, in $CON_{LS}$ (i.e., the control system; see Table S1) was slightly higher (5−9 mM) than recent estimates for the limit of solution stability (3−5 mM) in pure calcium carbonate solutions,[39] where the free-energy barrier to phase separation goes to zero. However, the free-energy landscape for this process in microscopic, thermodynamically closed systems is a function of both the size and composition of the initial state;[74] therefore, it is not surprising that concentrations beyond the limit of solution stability in the macroscopic case are found. Approximately 50% of the total (bi)carbonate was bound in $CON_{LS}$ at the steady state, but this was higher in the AA systems, most notably in $ARG_{LS}$, where only ∼17% of the total (bi)carbonate was free in solution by the end of the simulation. [Note: both here and hereafter, the term "(bi)carbonate" refers to both carbonate and bicarbonate ions.]

The fraction of bound species at low salinity are presented in Figure 2. By the end of the simulation, 60%−80% of the calcium was bound, which is consistent with typical experiments[35,47] and simulation studies.[35] Carbonate ions were preferentially bound to calcium over bicarbonate anions, and Na and Cl ions remained dispersed (using the definition of free ions adopted here) in solution. The preferential binding of carbonate over bicarbonate led to a reduction of the pH to ∼9.5 in all systems except for $ARG_{LS}$, where the pH was slightly higher (9.8 ± 0.1). ARG and GLY did lead to some reduction in the fraction of bound calcium, compared to the reference $CON_{LS}$ state. However, analysis of cluster species indicated that this was not due to calcium binding directly to AAs. Indeed, given the system size limitations, it could be argued that, in the model used here, the presence of AAs has a limited effect on the amount of

calcium carbonate binding before nucleation. Using another force field, Raiteri et al. measured the free energies of binding $Ca^{2+}$ and $CaCO_3^0$ to ASP as −12 kJ mol$^{-1}$ and approximately −2 kJ mol$^{-1}$, respectively, with a larger barrier for ASP coordination to the ion pair than to the free cation.[42] Our simulations indicated that the association of carbonate ions to calcium to form ion pairs was relatively fast (the thermodynamic driving force to forming $CaCO_3$ ion pairs is large,[75] ∼20 kJ mol$^{-1}$). Since most $Ca^{2+}$ ions are bound in these species (and occasionally in larger clusters), it is not surprising that even the additive with no net charge, GLY, had only a small effect on the number of dispersed $Ca^{2+}$ ions in solution. However, the general trend was that the amount of calcium binding was greater for ASP than for positively charged (ARG) or neutral (GLY) AAs in solution. The trend in our simulations is less obvious than in the experiments of Picker et al.,[47] where ASP had a stronger effect on the amount of free calcium in solution than ARG and GLY. This could be an effect of limited system sizes in the simulations and/or the model adopted here.

It is clear from Figure 2 that, while the binding of carbonate was not affected by the presence of AAs, there were significant differences between the systems simulated in the fractions of bicarbonate binding. These followed the sequence $ARG_{LS}$ > $GLY_{LS}$ > $ASP_{LS}$ > $CON_{LS}$. The binding trend is consistent with the opposing charges of amino acids and bicarbonate, with the most positively charged AA binding the largest fraction of the negatively charged bicarbonate anion. Indeed, upon close inspection of cluster species, it was clear that a large number of clusters that contained AAs also contained bicarbonate (see Figure S1 in the SI), where the anion was bound to −$NH_x$ functional groups. However, negatively charged additives (ASP) also showed increased bicarbonate binding, compared to the control state. The preference for calcium to bind to carbonates means that the solution is enriched with bicarbonate at early times, and bicarbonate can hydrogen-bond to AAs in solution.

The fractions of AAs bound at the steady state was high in all systems, and in the case of $ARG_{LS}$, no free ARG was found in solution by the end of the simulation. The fraction of bound AAs followed the sequence $ARG_{LS}$ > $GLY_{LS}$ > $ASP_{LS}$, following the fraction of bicarbonate binding. In addition, in $GLY_{LS}$ and
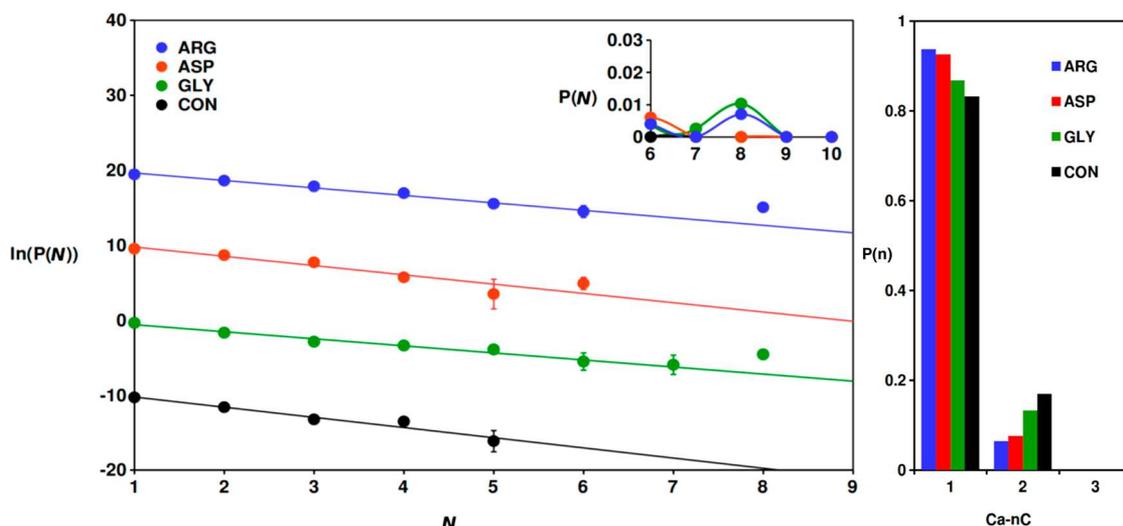
**Figure 3.** (Left) Natural logarithm of cluster size ($N$) probability distributions measured in the final 2 ns of simulations at low salinity. $N$ includes all solute ions and molecules within the clusters. The distributions were shifted as $\ln(P(N)) - 10$, $\ln(P(N))$, $\ln(P(N)) + 10$, and $\ln(P(N)) + 20$ for $CON_{LS}$, $GLYS_{LS}$, $ASP_{LS}$, and $ARG_{LS}$, respectively. The cluster size distribution for $N = 6-10$ is given in the inset. Lines are included as a visual reference, and error bars indicate one standard deviation in the data. (Right) Probabilities for calcium to coordinate to $n$ (bi)carbonate ions in the first coordination sphere.

$ASP_{LS}$, a significant number of free Na ions were removed from solution, and the trend in sodium binding matched that for AA binding. Binding of sodium was through the negatively charged carboxylate AA groups (see Figure S1). While chloride remained free in solution, up to 30% of the sodium was bound at the steady state. The presence of spectator ions in experiments maintains a background electrolyte charge density. However, it is important to consider their potential role in nucleation mechanisms, since they can bind to AA residues. In some cases, this binding enables the transport of certain AA residues in the cells of living organisms.[76,77] Experimental data on the equilibrium binding of spectator ions to amino acids is limited, but data do suggest that Na$^+$ binding to ASP, at least, is favorable at 298 K ($\log_{10}(K) = 0.42$, where $K$ is the equilibrium constant for association).[78]

Figure 3 shows that the steady-state cluster size probability distribution, $P(N)$, for clusters containing $N$ molecules/ions (using the first-shell coordination criterion to determine clustering), was dominated by free ions and ion pairs for all systems in the low-salinity regime; the probability of larger species rapidly decayed. In $CON_{LS}$, the entire probability distribution could be fitted to an exponential function. In contrast, larger clusters, containing up to eight ions or molecules, were observed with a higher probability in $GLY_{LS}$ and $ARG_{LS}$. In addition, the gradients of the linear fits of $\ln(P(N))$ in Figure 3 for systems containing AAs were slightly perturbed, compared to the control system, particularly for $GLY_{LS}$ and $ARG_{LS}$. These changes were caused by relatively large species that formed predominantly via the association of bicarbonate anions, Na$^+$ cations, and AAs (see examples in Figure S1), e.g., [ARG-$(HCO_3)_2$]$^-$ and [ASP(Na)$(HCO_3)_2$]$^{2-}$. In $ARG_{LS}$, the largest complex that we observed did contain calcium carbonate with the formula, [ARG(Ca)$_3$(CO$_3$)$_2$(HCO$_3$)$_2$]$^+$. Two bicarbonate ions were coordinated to the nitrogen-containing groups of the AA: one was bound to both the guanidinium and amine groups, while the other was more loosely bound to the amine group of the $\alpha$-carbon. The calcium carbonate ions formed a cluster around the AA guanidinium group, with carbonates interacting with $-NH_x$ groups and Ca ions coordinating to adjacent carbonate and bicarbonate O atoms. While there is obviously an

electrostatic attraction between the oppositely charged ions, the geometries of the clusters indicated that hydrogen bonding was also present. We also observed a [GLY(Ca)$_2$(CO$_3$)$_2$]$^0$ species where the binding of a (Ca)$_2$(CO$_3$)$_2$ ionic network was through carbonate oxygen groups binding to $-NH_x$ groups and Ca ions interacting with carboxylate O atoms. While these associated species appear to be mechanically stable for reasonably long times in the simulations, their relative thermodynamic stability has not been established.

Where Ca and carbonate ions were present in the associated AA species, they had a tendency to exist as small ionic networks interacting with the adjacent AA stabilized by other ions, such as bicarbonate or sodium. The calcium carbonate networks were similar to the small clusters which occurred with low probability in $CON_{LS}$. AAs seemed to provide an environment within the solution onto which the liquidlike ionic networks could associate. Figure 3 also provides the coordination probabilities for Ca ions to bind to $n$ (bi)carbonate ions in all solutions at low salinity. Whereas larger clusters of all molecule types were found in AA systems, the degree to which ionic calcium carbonate networks formed actually decreased, as indicated by the slight increase in the probability for calcium to bind to a single (bi)carbonate anion. The size of the calcium carbonate networks (in terms of the numbers of ions) in systems containing AAs was approximately the same as those in $CON_{LS}$, so the presence of AAs may produce ionic networks with reduced coordination, because of competing chemical interactions with the organic functional groups present in the additives.

The liquidlike calcium carbonate networks first identified in simulations[35] were initially thought to be the thermodynamically stable solute species which directly participated in the formation of post-nucleation phases (i.e., PNCs).[35] Wallace et al.[36] suggested that a population of small, dynamic clusters were stable in a statistical sense upon spontaneous liquid–liquid demixing. More recent experimental work suggests that these postulated PNCs transform to dense liquid calcium carbonate phases before the limit of solution stability is reached.[32] However, the proposed thermodynamic stability of liquidlike clusters in solution before nucleation has been ques-

tioned.[37,43,79] Smeets et al.[37] showed that a dense liquid phase of calcium carbonate is the first product of nucleation from solution, which is consistent with other experiments.[30,32] Experimental analyses of the species in solution before nucleation did not detect the presence of PNCs. Furthermore, Henzler et al.[43] used Ca K-edge X-ray absorption near-edge structure spectroscopy to examine the speciation of $Ca^{2+}$ in solution under a range of conditions and found this to be independent of concentration in the prenucleation regime. However, the interpretation of the data from the analyses in these works has also been questioned.[80]

Here, we will limit our discussion of PNCs to computational work that examines the stability of liquidlike clusters in solution before nucleation. Demichelis et al.[35] used a speciation model to determine the free-energy change for ion association from cluster size probability distributions in MD simulations over a range of calcium carbonate concentrations and pH levels. They adopted the assumption—following, e.g., the work of Gebauer et al.[34] and Kellermeier et al.[81]—that the equilibrium constant for direct cation−anion coordination in clusters is independent of cluster size in the so-called "multiple binding model". Fitting to all pH levels, the free-energy change for the reaction $CaCO_3^0$ + $(CaCO_3^0)_n \rightarrow (CaCO_3^0)_{n+1}$ is −21.7 kJ mol$^{-1}$ (and the equilibrium constant ($K$) is equal to 6355) for all $n$. At concentrations much closer to typical experiments, cluster size distributions over a range of pH levels show a rapid decay in cluster sizes with ion pairs dominating cluster populations (see Figure 3 and the work of Smeets et al.[37]) These results appear to conflict with the predictions of the multiple binding model. Basing their estimate on species populations in the steady state, Smeets et al. suggested that the free-energy change for the above reaction where $n = 1$ is much less negative than the equivalent change for the reaction $Ca^{2+}$ + $CO_3^{2-} \rightarrow CaCO_3^0$. The limitations of MD time scales were suggested as a possible reason for the smaller predicted equilibrium constants.[80]

To address the apparent discrepancy between the simulation results, we have performed US simulations to calculate the potentials of mean force (PMF) and establish the equilibrium constants for the following equilibria:

$$Ca^{2+} + CO_3^{2-} \rightleftharpoons CaCO_3^0 \qquad \Delta G_1 = -k_B N_A T \ln(K_1) \qquad (3)$$

$$2(CaCO_3^0) \rightleftharpoons (CaCO_3^0)_2 \qquad \Delta G_2 = -k_B N_A T \ln(K_2) \qquad (4)$$

PMFs were measured as a function of the separation distance between monomer units on the left-hand-side of the above equations. (Figure S2 in the SI shows the sampled probability densities for the reaction coordinate.) These can be used to obtain Gibbs free-energy changes ($\Delta G$) for association, provided that we add an analytical correction to the resulting PMFs to account for the increasing phase space volume as a function of the monomer−monomer separation distance, $r_{AB}$:

$$2k_B N_A T[\ln(r_{AB}) - \ln(r_c)] \qquad (5)$$

In the above equation, $r_c$ was the maximum distance sampled in the calculations. Therefore, we obtain one-dimensional free-energy profiles (see Figure 4) that asymptotically approach a limiting value of $\Delta G$ in the reaction coordinate sampled. The equilibrium constants for ion binding can then be calculated using the pair correlation function treatment that has been well-documented elsewhere:[54,75,80,82]
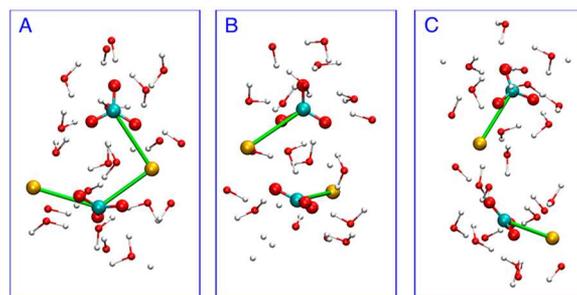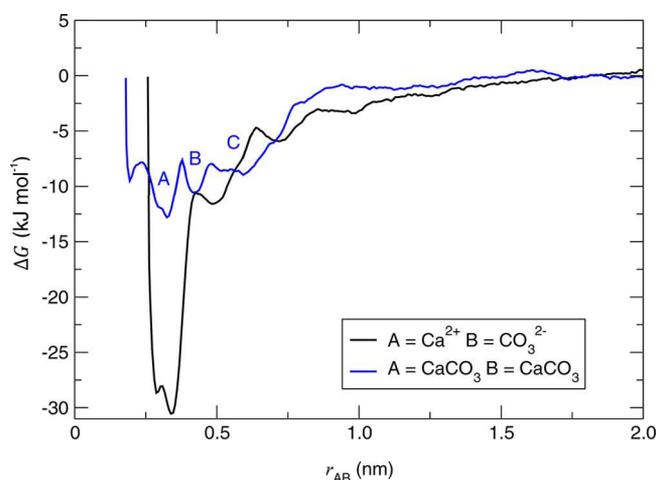


**Figure 4.** (Top) Potential of mean force energies ($\Delta G$) for the association of AB monomers in water, as a function of monomer−monomer distance $r_{AB}$, calculated from US simulations. (Bottom) Representative atomic configurations for minima A, B, and C identified for the association of two ion pairs in water. Ca, C, O, and H atoms are shown as gold, cyan, red, and white spheres with lines showing molecular bonds and green lines highlighting Ca−C connections within 3.5 Å. Only atoms of water molecule atoms within 4.1 Å of ions are shown for the sake of clarity.

$$K = 4\pi \int_{r_0}^{r_c} dr_{AB} \exp\left(\frac{-\Delta G(r_{AB})}{k_B N_A T}\right) r_{AB}^2 \qquad (6)$$

In the above equation, $r_0$ is the minimum distance in the calculated free-energy profiles. Two corrections are made to the calculated equilibrium constant. The first accounts for the fact that $K$ should be defined at standard concentrations, while the second makes use of activity coefficients, calculated using eq 2, to account for the effect of the background electrolyte in the simulations. Note that the measured $K$ value in eq 6 will be affected by the choice of integral limits, which are not predetermined. The limits chosen here account for the formation of even weakly associated species with multiple solvation shells between monomer units.

Figure 4 shows the resulting free energies. It is obvious that the free energies of binding a cation and anion in solution are much more negative than those calculated for bringing two ion pairs into direct contact. We calculate $\log(K_1) = 4.0$ and $\log(K_2) = 1.9$, giving $\Delta G_1 = -22.9$ kJ mol$^{-1}$ and $\Delta G_2 = -10.8$ kJ mol$^{-1}$. An uncertainty of $k_B N_A T = 2.5$ kJ mol$^{-1}$ applies to the data. $\Delta G_1$ is equal to earlier results from advanced sampling calculations.[75] These measurements support the earlier suggestion that the equilibrium constant for binding two ion pairs in a single step is much smaller than that for the binding of a cation with an anion.[37] Consistent with earlier calculations using the same model,[75] solvent-separated ($r_{AB} \approx 0.7$ nm) and solvent-shared

ion pair states ($r_{AB}$ = 0.5 nm) are visited as ions are brought close to each other before a $CaCO_3^0$ contact ion pair is formed. The global minimum ($r_{AB}$ = 0.34 nm) indicates a preference for monodentate carbonate binding. In the case of ion-pair–ion-pair association, only a bias potential applied to the distance between the centers of mass of the ion pairs was enforced. Therefore, it was important to ensure that, before and during association, the individual ion pairs remained intact (see Section S1 and Figure S3 in the SI). The representative atomic configurations in Figure 4 highlight that the first, relatively broad minimum in the energy landscape seen upon decreasing $r_{AB}$ (~0.6 nm) represents a solvent-separated state, where a cation and anion from adjacent ion pair monomers are coordinated via water molecules, which are oriented to maximize favorable interactions between the $Ca^{2+}$ ion and the O of the carbonate. The remaining ions in each ion pair are surrounded by solvation shells, aside from the direct ion-pair coordination. A small energy barrier must be overcome to displace water molecules from these ions in order to form a state where all ions are associated within their first two coordination spheres when $r_{AB}$ = 0.44 nm. This state is still a solvent-separated one and, while it appears as a distinct minimum, we are not biasing the sampling of different water coordination numbers surrounding each ion which could affect the topological features and relative free-energy barriers of the resulting energy landscapes.[75] The most stable state is observed when $r_{AB}$ = 0.33 nm and, here, all of the ions are directly coordinated to adjacent ions in a liquidlike chain. Similar calculations have recently been conducted by Raiteri et al.,[83] using a rigid molecule version of the force field adopted here and a newly developed polarizable force field. The rigid force field model suggests that $\Delta G_2$ = −6.0 kJ mol$^{-1}$, which was slightly less negative than the result from the polarizable force field (−6.8 kJ mol$^{-1}$). Our result suggests a more favorable binding energy; this could be due to the constraints imposed to maintain the ion pair monomers in the other work, and differences in the chosen integral limits for the calculation of equilibrium constants. Raiteri et al. extended their study to measure the free-energy changes for adding two more ion pairs into growing clusters. The binding energies were between −4.6 kJ mol$^{-1}$ and −10 kJ mol$^{-1}$, depending on the reaction step and force field used to add ion pairs into clusters. The authors highlight that the negative free energies indicate increased thermodynamic stability for the larger clusters in solution. We agree that equilibrium constants larger than unity should result in a population of larger clusters (than ion pairs) in solution, albeit the concentration of these species will be low at typical free-ion concentrations in the millimolar regime. However, it is also important to consider the different types of ion associates that will form in solution. As highlighted in this work, both ion pairs in direct contact and solvent-shared and solvent-separated clusters of ion pairs were observed in simulations. We can measure the equilibrium constant, $K_{CIPs}$, for the formation of $(CaCO_3^0)_2$, where ions are directly coordinated or where ion pair monomers are separated by water molecules. This is done by calculating the ratio of line integrals for different regions in the free-energy profile that represent these states; i.e.,

$$K_{CIPs} = \frac{\int_{0.15}^{3.7} dr_{AB} \exp\left(\frac{-\Delta G(r_{AB})}{k_B N_A T}\right) r_{AB}^2}{\int_{3.7}^{7.8} dr_{AB} \exp\left(\frac{-\Delta G(r_{AB})}{k_B N_A T}\right) r_{AB}^2} \quad (7)$$

Using the above equation, $\log(K_{CIPs})$ = −0.16 and the free-energy change for the removal of water to form an ion cluster with ions in direct contact is 0.9 kJ mol$^{-1}$. Hence, while the association of two ion pairs is thermodynamically favorable, $(CaCO_3^0)_2$ populations will be dominated by highly hydrated, weakly bound associates in the standard state. In future work, it will be necessary to explore the free energies for different solvation states in clusters containing more than a single ion pair to reduce the uncertainty in this estimate for the equilibrium between directly coordinated and solvent-separated ions in $(CaCO_3^0)_n$ states.

Both our results and those of Raiteri et al.[83] indicate that the equilibrium constant for Ca−CO$_3$ binding in clusters is larger than unity (and, therefore, association is favorable) but much smaller than for the formation of the ion pair from dispersed monomers, and Raiteri et al. also find different equilibrium constants for different steps in stepwise multiple binding, contrary to the underlying assumption in the multiple binding model. In light of these new results and interpretation, a re-evaluation of the cluster size probability distributions and the role of water in clusters that would emerge in pure calcium carbonate solutions before nucleation should be considered, and experimental efforts in this endeavor will be invaluable. A smaller population of prenucleation clusters does not have significant implications for the prenucleation cluster pathway. However, we cannot confirm the role of these clusters in nucleation from these calculations alone. We should also note that (i) the accuracy of the adopted force field has been questioned,[43] and (ii) a thermodynamic-kinetic growth model of calcium carbonate growth suggests that clusters may be present are merely spectators,[79] although the analyses in both these papers have also been criticized.[80]

Our MD calculations suggest that the presence of AAs reduces the density of calcium carbonate ionic networks in solution. Not only did we see a reduction in the average coordination by calcium, we also found that slightly more free calcium was available in solution. However, in our simulations, there is no replenishment of the (bi)carbonate which associates with AAs and therefore a constant background concentration of the anion cannot be maintained. New methods, which allow for simulations at constant supersaturation may prove useful to circumvent these issues in future studies.[84−86] If AAs sequestered bicarbonate in experiments, then the equilibrium distribution of carbonate/bicarbonate in the bulk solution would adjust to counter this change in bicarbonate concentrations. To further analyze the chemical interactions described above, approximations for the free-energy changes (provided in Table S2 in the SI), based on species populations, for the forward reactions in the following equilibria were calculated:

$$Ca^{2+} + HCO_3^- \rightleftharpoons CaHCO_3^+ \quad (8)$$

$$AA + HCO_3^- \rightleftharpoons AAHCO_3 \quad (9)$$
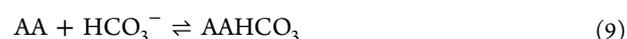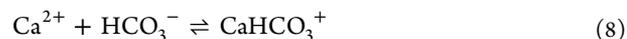
For the control system, the free-energy changes for the formation of the ion pairs in eqs 3 and 8 were −21 kJ mol$^{-1}$ and −12 kJ mol$^{-1}$, respectively, in good agreement with earlier simulations[35,75] and the free-energy calculations here. The frequencies for ion attachment/detachment events was ~0.03 ps$^{-1}$, with 60% of events occurring in ion/molecule clusters containing at least two of the monomers. This is a good indication that the systems reached a steady state. Given the large changes in the amount of free bicarbonate in solution in the steady state, due to association with AA, we calculated the
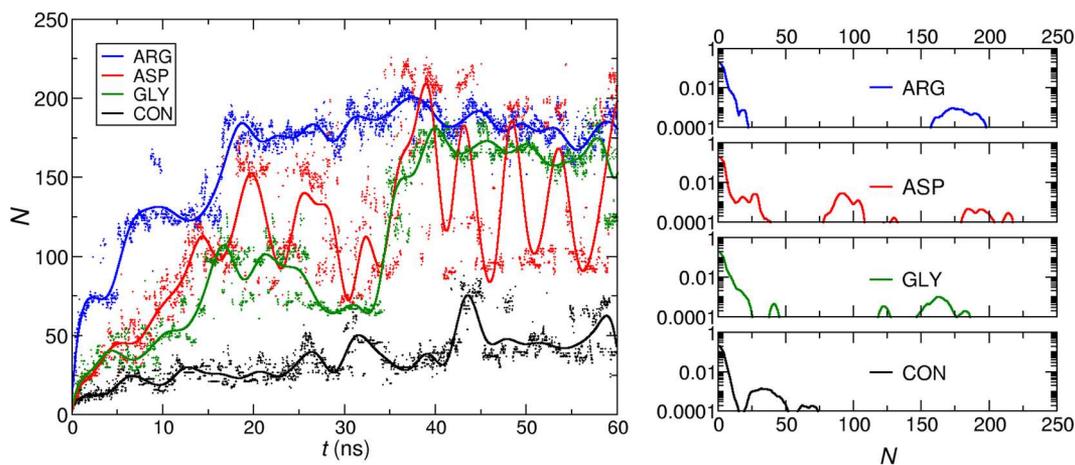
**Figure 5.** (Left) Largest cluster size in simulations as a function of time $t$ at high salinity, according to the total number of ions/molecules $N$. Points show measurements from trajectory data and lines show a running average over 50 ps. (Right) cluster size probability distributions averaged over the final 10 ns of simulation; data are shown on a logarithmic scale of probability.

approximate free-energy changes for bicarbonate association reactions (i.e., eq 9). The forward reaction was slightly more favorable in $GLY_{LS}$ than in $ASP_{LS}$ ($\Delta G = -11$ kJ mol$^{-1}$ ($GLY_{LS}$) and $-10$ kJ mol$^{-1}$ ($ASP_{LS}$)). This means that the driving force for bicarbonate association to calcium is comparable to binding with AAs. Because all of the arginine is involved in associated species, it is not possible to approximate the free energy of binding using the above approach. Instead, we performed US calculations of ARG binding to bicarbonate. The resulting $\Delta G$ profile showed a plateau at ~10 Å, as shown in Figure S4 in the Supporting Information. No energy barrier inhibits the bicarbonate from binding to ARG through the guanidinium functional group to bicarbonate oxygens ($\Delta G \approx -17$ kJ mol$^{-1}$), where the (ARG)N−O(bicarb) coordination number was slightly greater than two. Guanidinium groups are known to bind strongly to carbonate from previous simulations.[18] An energy barrier of ~4 kJ mol$^{-1}$ separates this minimum from the most stable coordination state (with a negligible kink in the free-energy profile at ~5 Å), where the conformational freedom available to ARG allows multiple $NH_x$ groups to associate with bicarbonate ($\Delta G \approx -19$ kJ mol$^{-1}$ for additional binding). Here, the mean coordination number between (ARG)N and O(bicarb) was between three and four, with much higher coordination levels through both first and second coordination shell binding (see Figure S4). Conformational freedom and multiple hydrogen bonds to bicarbonate help explain why ARG binds most strongly to anions in solution.

Since bicarbonate binds so effectively to ARG, we might expect carbonate also to bind strongly, given that the charge density of this anion is larger. The free-energy profile for ARG associated with carbonate is provided in Figure S5 in the Supporting Information. $\Delta G$ increases as a function of $r$ from the most stable state up to a maximum of ~10 Å. An energy barrier of several kJ mol$^{-1}$ separates dispersed and coordinated states. Coordination geometries for ARG−carbonate are quite similar to the ones for the bicarbonate case with minima at similar positions in the reaction coordinate. However, the kink that appeared in Figure S4 at ~4.8 Å is a defined minimum in the case of carbonate. The energy to remove carbonate from ARG is ~7 kJ mol$^{-1}$ less positive than that for bicarbonate. Combining this with the observation of a barrier to ARG−carbonate binding and the preference for calcium to bind to carbonate over bicarbonate (where the binding is around twice as strong and with minimal

energy barriers to association), it is not surprising that we see significant association of ARG with bicarbonate in solution.

**High Salinity.** At increasing levels of supersaturation, solution stability decreases and phase separation becomes increasingly likely until the limit of solution stability is reached, when phase separation is instantaneous. Demichelis et al.[35] observed the formation of extended clusters, which consumed the majority of ions in solution at high calcium carbonate concentrations and at high pH. Smeets et al.[37] presented these extended ionic networks as the dense liquid phases that were found in experiments[30,38] and predicted in simulations elsewhere.[36] Here, we examine the effect of additives during the formation of extended calcium carbonate networks in solution.

In the high salinity simulations, ion association was observed immediately. Figure 5 shows how the size of the largest clusters increased during the simulations. In the case of $CON_{HS}$, the largest clusters grew steadily up to around 40 ns after which a plateau in cluster size ($N$) against time is superimposed on significant noise. The large fluctuations in the cluster sizes indicate that the growth was not only by addition of dispersed ions and small clusters but also by the association of already large clusters of solute monomers, and that clusters occasionally dissociated into smaller species but were coordinated within a few solvation layers. It is possible that the largest clusters would further associate. However, because the concentration of these species decreases over time and because their mobility in solution decreases due to increased mass, long simulation times may be required. Over the final 10 ns of simulation, the average size of the largest cluster was $47 \pm 12$.

Compared to $CON_{HS}$, there was an increase in the growth rate of clusters when AAs were included in the solution. This was most apparent in $ARG_{HS}$, where there was a rapid increase in the largest cluster up to $N \approx 50$ in the first nanosecond, followed by a steadier rate of growth that was interspersed several times with jumps in the value of $N$ as large associates combined up to $N = 177 \pm 10$ at the steady state. A movie of the formation of clusters can be found in the SI. The growth of the largest assemblies in $GLY_{HS}$ was not as rapid, but at the steady state beyond 40 ns, a large species emerged that was slightly smaller ($N = 160 \pm 17$) than the assembly in $ARG_{HS}$. The average size of the largest cluster in $ASP_{HS}$ fluctuates significantly after an initial rapid increase in $N$ over the first 20 ns. This was due to dynamic first-shell coordination between two large regions of the cluster,

**Table 2. Solute Concentrations, *I*, and pH over the Final 10 ns of Simulation of Initially High Salinity Simulations Containing Amino Acids**[a]

| system | $c(Na^+)$ [mM] | $c(Cl^-)$ [mM] | $c(AA)$ [mM] | $c(Ca^{2+})$ [mM] | $c(CO_3^{2-})$ [mM] | $c(HCO_3^-)$ [mM] | $I$ [M] | pH |
|---|---|---|---|---|---|---|---|---|
| $CON_{HS}$ | 119 ± 1 | 122 ± 1 | 0 | 13 ± 7 | 3 ± 1 | 150 ± 9 | 0.23 ± 0.02 | 8.7 ± 0.1 |
| $ARG_{HS}$ | 0 | 127 ± 1 | 0 | 11 ± 5 | 1 ± 1 | 50 ± 8 | 0.11 ± 0.01 | 8.8 ± 0.1 |
| $ASP_{HS}$ | 0 | 28 ± 0.2 | 0 | 15 ± 5 | 4 ± 2 | 69 ± 10 | 0.09 ± 0.01 | 9.1 ± 0.1 |
| $GLY_{HS}$ | 6 ± 1 | 127 ± 0.1 | 0 | 8 ± 4 | 4 ± 3 | 62 ± 10 | 0.12 ± 0.01 | 9.2 ± 0.1 |

[a] $c(X)$ indicate the concentration of ion X.

which contained ∼100 and 200 solute monomers, respectively (the largest mean cluster size over the final 10 ns was 133 ± 46). Increasing the adopted truncation distance in the definition of associated monomers would attenuate these fluctuations, and we note that the dynamic first-shell coordination displayed here is not indicative of clusters dissociating (i.e., seemingly independent clusters were associated by ions separated by solvation shells). The size of the largest sampled clusters followed $ASP_{HS} > ARG_{HS} > GLY_{HS} > CON_{HS}$. Both the net charge and the size of the molecular additives were important for the formation of large assemblies. The simulations were repeated an additional two times from different starting configurations and the results were qualitatively the same.

Cluster size probability distributions over the final 10 ns of simulations are provided in Figure 5. In $CON_{HS}$, there was a rapid decrease in the cluster size probabilities. The probability of finding species where $N = 1, 2$, and 3, that is for free ions, ion pairs, and ion trimers, was 0.66, 0.16, and 0.08, respectively. Two additional broad peaks with low probabilities were observed centered around $N = 35$ and 65. Visual inspection (see Figure S6 in the Supporting Information) showed a handful of large, dynamically coordinated assemblies of ions that were mainly composed of calcium and carbonate. A wide distribution in Ca–C coordination probabilities was found with a maximum for calcium binding with two (bi)carbonates (see Figure S7A in the Supporting Information). Decomposing these into contributions from carbonate and bicarbonate highlight that the network is dominated by the divalent anion, with a maximum of just two bicarbonates binding to a single calcium (see Figures S7B and S7C in the Supporting Information). The average coordination number of $Ca^{2+}$ to $CO_3^{2-}$ and $HCO_3^-$ was 1.5 ± 0.1 and 0.4 ± 0.1, respectively. The high mobility of ions in clusters is evident in the measured diffusion coefficients $(D_{Ca})$ for Ca ions. The mean $D_{Ca}$ in $CON_{HS}$ was $1.61 \times 10^{-6}$ cm$^2$ s$^{-1}$—an order of magnitude lower than the diffusion coefficient, $D_{H_2O}$, for water $(1.66 \times 10^{-5}$ cm$^2$ s$^{-1}$; $D_{H_2O}$ for pure SPC/Fw water was $2.34 \times 10^{-5}$ cm$^2$ s$^{-1})$. We measured $D_{Ca}$ as a function of distance from the center of mass of the largest cluster (see Figure S8 in the Supporting Information), and for Ca ion closest to the core, $D_{Ca} = 1.02 \times 10^{-7}$ cm$^2$ s$^{-1}$, while for a Ca ion at the periphery of the cluster, $D_{Ca} = 1.27 \times 10^{-6}$ cm$^2$ s$^{-1}$ (when the diffusion of the cluster was subtracted). Even at the center of the cluster, ions diffuse faster than in solid ACC,[87] by ∼3 orders of magnitude. An order-of-magnitude change in the self-diffusion coefficient of ions in dense liquid carbonate phases was measured by Bewernitz et al., using NMR techniques.[30] The remaining solution contains the smaller associates and free ions, which were largely bicarbonate. This is reflected in the concentrations provided in Table 2. Approximately 37% of bicarbonate remained completely solvated in their first shells by the end of the simulation. In contrast, 98%−99% of calcium and carbonate ions were directly coordinated to other ions. The concentrations

of these species are approximately the same as those obtained in simulations at low salinity. In addition, both Na and Cl ions remained dispersed in solution. This type of behavior can be expected from earlier predictions.[35] At higher pH values, Smeets et al.[37] observed the complete separation of large, single calcium carbonate liquidlike associates from a lean solution. However, simulations in this work, at pH ∼9 (see Table 2), contain bicarbonate ions that cannot convert to carbonate and are less likely to associate with calcium.

All of the simulations that included AAs showed multimodal probability distributions with a much clearer separation between dispersed ions, small associates, and large associates (see Figure 5), compared to $CON_{HS}$. While there was always a high probability to find free ions and small ion associates, peaks at large N were centered around the sizes of the largest clusters, as discussed above. We do find, generally, a separation of large, single solute-associated species from a population of much smaller species that are dominated by free ions and ion pairs. In contrast to $CON_{HS}$, the probabilities for finding clusters containing $N = 1, 2$, and 3 were 0.62, 0.14, and 0.12 in $ARG_{HS}$. However, in $ASP_{HS}$, as noted above, we observed a dynamic clustering (using the adopted truncation distance for association of 3.5 Å) between two large ASP-rich regions in this system; therefore, two peaks are observed in the cluster size probability distribution at $N \approx 100$ and 200, with the smaller cluster being ∼4 times more probable (on average). All of the largest assemblies displayed some degree of dynamic ordering. The broad peak centered around $N = 180$ in $ARG_{HS}$ highlights this (also see individual data points in the left of Figure 5). The size of this cluster, measured in terms of radius of gyration, was ∼1.7 nm.

As solute monomers associated rapidly into a small number of large assemblies, which appear to be in dynamic equilibrium with smaller associates in a solution where the ionic strength is ∼10% of the initial conditions (see Table 2), it is helpful to consider mechanisms and thermodynamic driving forces for the association. It is reasonable to assume that the solution is beyond its stability limit at $t = 0$, given the level of the initial supersaturation here. Indeed, this limit was predicted to be on the order of $c(Ca)$ that we found at the steady state in low-salinity simulations.[39] Because the association of calcium with the divalent carbonate anion is thermodynamically more favorable, the initially homogeneous solution decomposes by forming $CaCO_3$ clusters, which are then in a dynamic equilibrium with the remaining ions in the surrounding solution. These dense liquids could subsequently dehydrate to form solid ACC.[36] Figure S9 in the Supporting Information provides the RDFs for Ca–C, as a function of time. These do not indicate any increasing density of the ionic network over the final 10 ns, but there are fluctuations in the ratios of monodentate versus bidentate carbonate binding to calcium. Advanced sampling simulations have indicated that the transition from liquidlike calcium carbonate nanoparticles to solidlike ACC with

stoichiometry $CaCO_3 \cdot H_2O$ occurs in clusters containing 30−40 $CaCO_3$ ion pairs.[88] However, kinetic barriers to dehydration will limit the amount of water that can be removed from the liquidlike clusters in these simulations, and calculations sampling different hydration levels should be considered.

All of the AAs were associated in clusters by the end of the simulations. In $ARG_{HS}$ and $GLY_{HS}$, the AAs were found in the single, large clusters that emerged, whereas, in $ASP_{HS}$, four of the AAs were in a cluster that was dynamically coordinated to another cluster containing the remaining 14 AAs. Table S3 in the Supporting Information shows that, in systems containing additives, the level of calcium and carbonate association was similar to that observed in $CON_{HS}$. However, there was greater association of bicarbonate, with 83%−88% bound in solution over the final 10 ns of simulation (at $t = 0$, $c(HCO_3^-) \approx 400$ mM). The amount of bicarbonate binding ranked as follows: $ARG_{HS} > GLY_{HS} > ASP_{HS} > CON_{HS}$, with the trend being consistent with that observed at low salinities. The removal of these additional ions from the solution where additives were included led to a reduction in ionic strength to approximately half the levels of $CON_{HS}$ and increases in the final pH (see Table 2). Cl ions remained dispersed in solution surrounded by water, yet Na ions were bound in associates in the final 10 ns of simulation, in stark contrast to the observations in $CON_{HS}$. While there was no correlation in the removal of sodium with additive charge, we note that the amount of sodium added was to control the initial ionic strength; therefore, where additives were positive, fewer Na ions were simulated.

Snapshots of the configurations for the AA simulations at high salinity at the end of 60 ns simulation are provided in Figure 6. The large cluster containing all of the AAs can be seen surrounded by a $CaCO_3$ ionic network, which adsorbs to the AA assembly. Other $CaCO_3$ ionic networks can be seen in the surrounding solution, along with free ions, ion pairs, and other small associates. These depictions are similar to what was seen in $CON_{HS}$. It appears that the large cluster assembly is facilitated by bicarbonate ions (see Figure 6 and the movie in the Supporting Information). A large number of bicarbonate ions is observed throughout the highlighted region and is stabilized through hydrogen bonding to ARG molecules, as found in US calculations at low salinity. Some carbonate ions are also observed at the periphery of the assemblies. Because the guanidinium functional groups of ARG have a tendency to be directed toward bicarbonates in the center of the assembly, carboxylate groups tend toward the periphery of the cluster and we observed coordination of these groups to sodium. The average coordination probabilities are $1.3 \pm 0.03$ and $0.5 \pm 0.03$ for Ca−carbonate and Ca−bicarbonate in clusters and the distribution of Ca−C binding shifts to smaller values (see Figure S7 in the Supporting Information). The decrease in carbonate coordination is due to the topology of the AA assembly limiting the mass density of the calcium carbonate network. A small amount of calcium is coordinated to the AAs through carboxylate groups (see Figure S7). From these data, we see that the ARG assembly provides an environment where calcium carbonate networks can adsorb. In doing so, the ionic network takes on coordination levels that are smaller than networks in the remaining solution, which indicates a decreasing ability to increase ionic density. In addition, the mobility of calcium decreases when adsorbed to the ARG assembly. Figure S8 shows that the diffusion coefficients for Ca ions in the largest cluster were $6.1 \times 10^{-8}$ cm$^2$ s$^{-1}$ and $3.1 \times 10^{-7}$ cm$^2$ s$^{-1}$ (when we subtract the diffusion of the cluster) for ions closest to the core
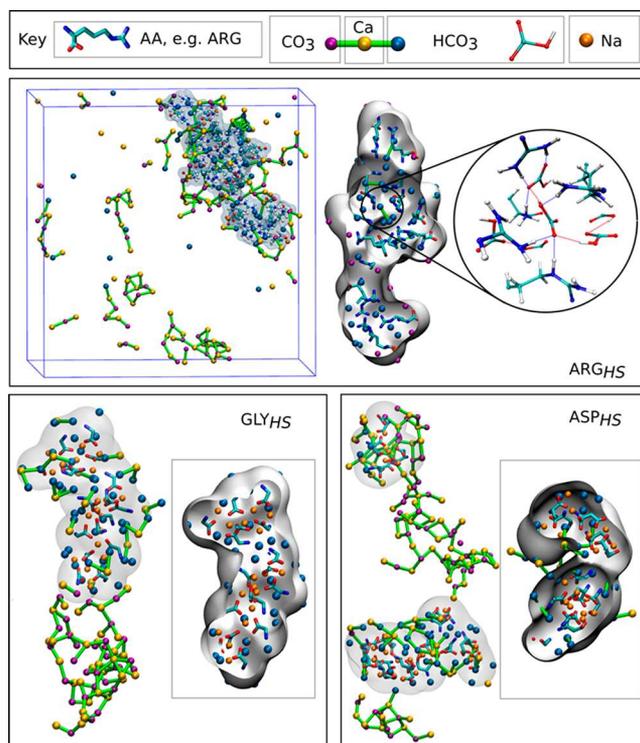


**Figure 6.** Snapshots from the end of 60 ns of simulation of solutions containing calcium (bi)carbonate and AA additives at high salinity. A key to atom types is provided; water and chloride are not shown for the sake of clarity. Green lines show calcium−(bi)carbonate inner-sphere coordination. In $ARG_{HS}$, the entire simulation cell is shown. The AA assemblies are shown with smoothed surfaces created according to approximately twice the van der Waals radii of solute atoms. The AA assemblies are enlarged on the right with a solid surface, indicating the approximate boundary with surrounding ions and water. In $ARG_{HS}$, we also highlight the hydrogen bonding (shown by the dashed lines) between ARG molecules and bicarbonate in the core of the assembly.

and those at the interface between AAs and solution, respectively. There is approximately an order of magnitude difference in the diffusion of Ca ions in the largest clusters when the AAs are present, compared to $CON_{HS}$.

The large AA assemblies that form in $GLY_{HS}$ and $ASP_{HS}$ were again facilitated by the coordination of AAs to bicarbonates through $-NH_3$ groups. However, here, we also observe a significant amount of sodium within the assemblies coordinated to the AAs through carboxylate functional groups. These Na ions had a tendency to be in the core of clusters, which is where maximum coordination can be achieved. This provides anions coordinated to AAs that can also bind to calcium. Again, we saw calcium carbonate networks adsorbing to the surfaces of assemblies and the networks themselves were quite substantial in size. The largest cluster contained 50% of the total calcium and carbonate available in the simulation, while 70% and 85% of the total bicarbonate and sodium were found in the cluster, respectively. For $GLY_{HS}$, the coordination probabilities (see Figure S7) were shifted to smaller values. In $ASP_{HS}$, the shape of the coordination distributions was closest to $CON_{HS}$. This can be explained by the composition of the clusters. As shown in Figure 6, two regions rich in ASP are connected by a $CaCO_3$ network. This liquidlike network undergoes dynamic ordering, producing the large fluctuations in the largest cluster size, as a function of time in Figure 5. The cluster comprises 65%, 70%, and 63% of the total calcium, carbonate, and bicarbonate in the
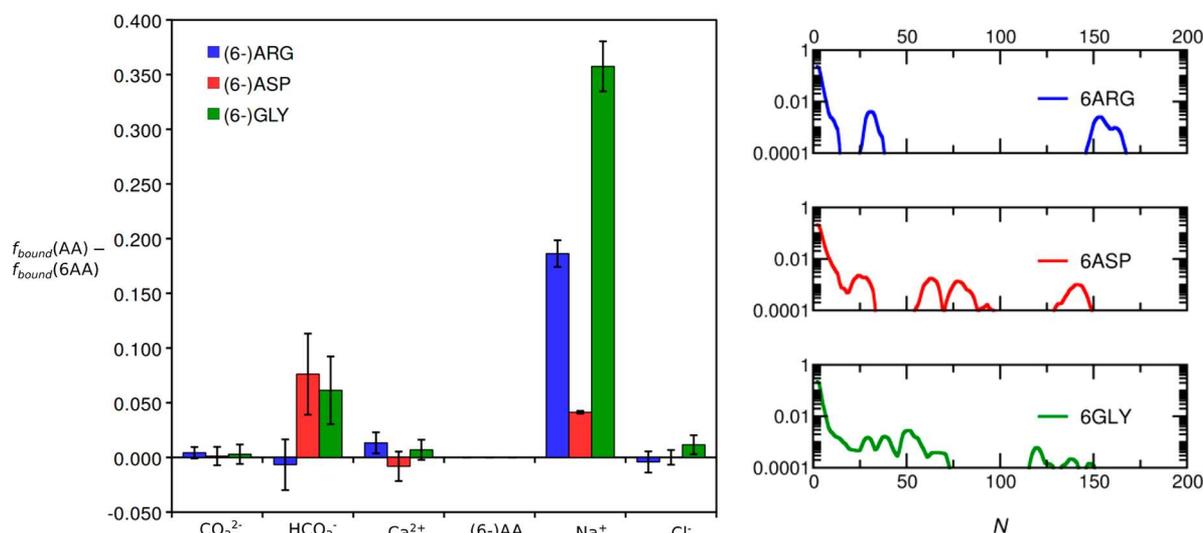
**Figure 7.** (Left) Differences between the fractions of bound solute (using a coordination distance cut-off criterion) between simulations containing AAs and 6AAs. (Right) Cluster size probability distributions in solutions containing 6AAs at high salinity, according to the total number, $N$, of ions/ molecules in clusters. Data were averaged over the final 6 ns of simulation, and error bars show uncertainties of one standard deviation.
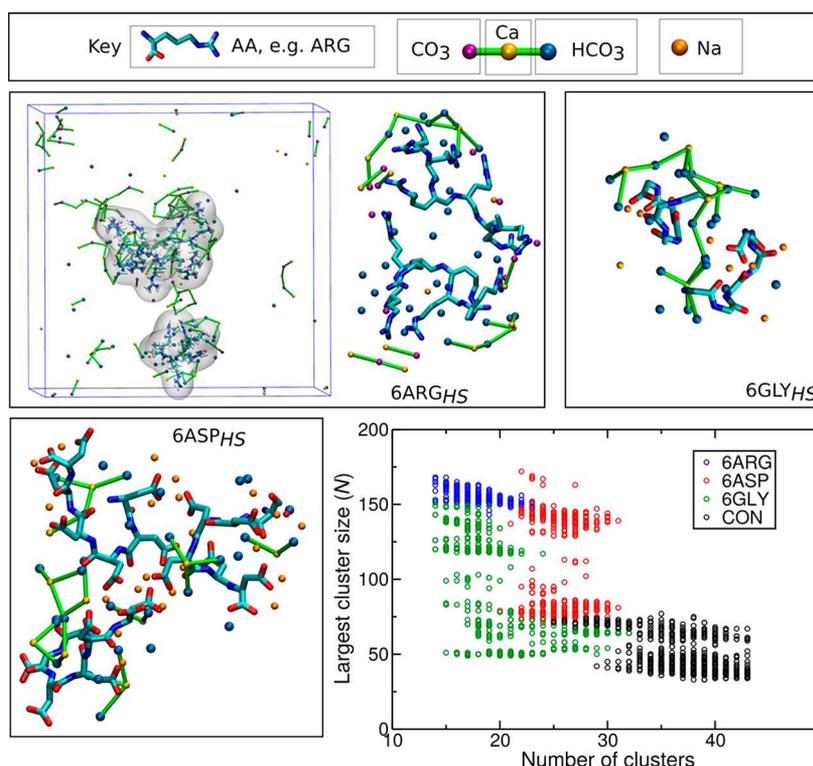


**Figure 8.** Snapshots from the end of 60 ns of simulation of solutions containing calcium (bi)carbonate and 6AA additives at high salinity. A key to atom types is provided; water and chloride are not shown for the sake of clarity. Green lines highlight calcium−(bi)carbonate inner-sphere coordination. The largest clusters containing 6AAs are shown with all atoms within 3 Å also shown. The plot shows the distributions of the total number of clusters (containing at least three solute monomers) versus the number of solute monomers ($N$) in the largest cluster from the final nanosecond of simulation. Data for $CON_{HS}$ at the steady state are plotted for comparison.

system and a significant amount of these ions is found away from the surface of the AA assemblies. In both $GLY_{HS}$ and $ASP_{HS}$, the smaller size of additives leads to different geometries for the adsorbed calcium carbonate networks, cf. $ARG_{HS}$. This affects the ionic coordination in the networks. In $ASP_{HS}$ in particular, coordination levels larger than those observed in $CON_{HS}$ are evident. In both cases, we again found a small amount of Ca binding directly through the AA carboxylate groups.

When supersaturations are high, the additives interact with the calcium carbonate liquidlike networks that are observed in control simulations. Only subtle changes in the coordination of the networks are seen. It is interesting to speculate about how additives would control nucleation if solute concentrations were lowered and system size limitations were not evident. In this case, it is reasonable to assume that several smaller additive assemblies could be stabilized by the presence of bicarbonate

and so-called spectator ions. The $CaCO_3$ networks that form could adsorb to these assemblies and connect them in solution. In the presence of poly(ASP), experiments[30,89] have shown that a dense liquid phase of calcium carbonate is stabilized in solution. It was assumed that the mechanism of action was to remove Ca ions in solution; subsequent diffusion of carbonate eventually led to crystallization. Our results suggest that Ca depletion from the mother solution is not necessary to stabilize dense liquid calcium carbonate. However, the action of extended polymers may be different from assemblies of individual AA monomers.

**Oligopeptides.** We have also considered the effect of additive size on the formation of clusters in solution, again at high salinity. Here, we simulated oligopeptide chains of ARG, ASP, and GLY, where six monomer AAs were connected by peptide bonds. In forming the peptide bonds, $NH_3^+$ and $CO_2^-$ groups are replaced and so fewer functional groups are available for coordination to ions in solution. Nonetheless, the net charges of 6ARG, 6ASP, and 6GLY were $+6e$, $-6e$ and zero in the simulation cells. Figure S10 in the Supporting Information provides the structures that were adopted in this work. Three oligopeptides were included in simulations, in order to retain the same total charge.

Rapid growth of large assemblies emerged in simulations of calcium (bi)carbonate in the presence of 6AA oligopeptides, just as when AAs were included. Large clusters were observed in equilibrium with a surrounding solution where the free ion concentrations at the steady state (see Table S4 in the Supporting Information) were generally consistent with simulations containing AAs. Again, all of the 6AAs formed clusters with other solute monomers. There were some subtle changes in solution compositions at the steady state. Most notably, there were slight increases in the pH of $6ARG_{HS}$ ($9.1 \pm 0.1$), compared with $ARG_{HS}$ ($8.8 \pm 0.1$). In addition, the ionic strength of $6GLY_{HS}$ ($0.17 \pm 0.01$ M) increased compared with $GLY_{HS}$ ($0.12 \pm 0.01$ M). Table S5 in the Supporting Information provides the fraction of bound solute species at the steady state. For ease of comparison, Figure 7 shows the differences between the binding fractions in these simulations and those containing AAs at high salinity. The formation of peptide chains from six AAs clearly does not have any impact on the level of $Ca^{2+}$ and $CO_3^{2-}$ association in solution. Both $6ASP_{HS}$ and $6GLY_{HS}$ show a reduction in the fraction of bicarbonate binding. While all 6AA simulations show a reduction in the fraction of sodium binding, this is most notable in $6ARG_{HS}$ and $6GLY_{HS}$. These changes can be reconciled by considering the difference in chemistries of the AA additives on forming peptide bonds. In all cases, the ability for $-NH_x$ and $-CO_2$ functional groups of AAs to bind to bicarbonate and sodium is reduced when those groups are replaced by the peptide bond. $6ARG_{HS}$ and $6ASP_{HS}$ retain these functional groups in their side chains; therefore, the changes are minimized in these systems.

The reduction in the number of chemical functional groups able to bind to solute monomers is also evident in the cluster size probability distributions, shown in Figure 7. Again, we note that the choice of cutoff in determining clusters will affect the measured cluster size populations. The largest clusters were found for $6ARG_{HS}$, where the cluster size distribution was shifted to slightly smaller values. In $6ASP_{HS}$, all of the additive molecules were contained in a single cluster, shown in Figure 8, and the size of this cluster is smaller than the largest cluster in $ASP_{HS}$. We also note that, in the case of oligopeptides, there are

more large clusters in the simulations. Figure 8 shows that 6AAs are stabilized by coordination to bicarbonate and sodium. The conformational freedom of the additives allows multiple binding of them to the same solute monomer. At high supersaturations, chelation of solute monomers in close proximity to the additives will be fast (assuming that the energy barriers for this process are accessible), while diffusion of the additive clusters will be relatively slow. In contrast, solute binding to smaller AA additives enables assembly of large clusters. Given sufficient time, there could be further association and restructuring of the 6AA clusters, but this may be beyond reasonable simulation time scales. We can see, particularly in the case of $6GLY_{HS}$, that two 6GLY molecules have folded to coordinate to several cations; these then formed a larger cluster, facilitated by the calcium (bi)carbonate in solution.

Figure 8 shows the distribution of cluster sizes versus the total number of clusters in 6AA simulations, along with the results from $CON_{HS}$. The distribution for $6ARG_{HS}$ is relatively narrow. This large cluster was formed of ~150 solute monomers, which consisted of ~$55Ca^{2+}$, ~$34CO_3^{2-}$, and ~$56HCO_3^-$ ions and all of the additive molecules. Two groups are apparent for $6ASP_{HS}$. Here, a 3(6ASP) cluster containing ~$20Ca^{2+}$, ~$8CO_3^{2-}$, and ~$36HCO_3^-$ (for an average total of 85 monomers) formed separately to a single, liquid $(CaCO_3)_{22}$ ionic network, which dynamically associated with the additive assembly in the simulation. A much more disperse dataset was found for $6GLY_{HS}$, but this was still very different from the control simulation at high salinity. Both the number of solute binding sites and additive size appear to be important to control the level of dynamic clustering in the liquid phase.

RDFs in Figure S11 in the Supporting Information show that the dominant monomer binding to 6ARG was bicarbonate, with also some carbonate binding. Calcium has a tendency to be found beyond the first coordination sphere of atoms in 6ARG. While 6ASP and 6GLY also coordinate to bicarbonate, Na ions are present at smaller coordination distances and there is a reduction in the amount of calcium coordination in the outer sphere of the atoms in these systems. In all of the large 6AA clusters, we find $CaCO_3$ ionic networks adsorbed. Therefore, these additives act in much the same way as AAs at high salinity. Figure S12 in the Supporting Information provides the coordination probabilities for Ca inner-sphere coordination in clusters. Very limited coordination is found between calcium and 6AAs. The coordination profiles for Ca binding to bicarbonate were much the same as for AA simulations (also see Figure S7). The differences between additives for high coordination levels of Ca$-n$CO$_3$ became more apparent when compared to AA simulations. For example, the probabilities of Ca$-3$CO$_3$ were 0.1, 0.14, and 0.19, respectively, for $6ARG_{HS}$, $6ASP_{HS}$, and $6GLY_{HS}$, respectively. In terms of the average total coordination number of Ca to (bi)carbonate, we found an increase in the case of $6ARG_{HS}$ ($n = 1.82$ and 2.01 in $ARG_{HS}$ and $6ARG_{HS}$) and $6GLY_{HS}$ (($n = 1.93$ and 2.22 in $GLY_{HS}$ and $6GLY_{HS}$) and a decrease for $6ASP_{HS}$ ($n = 2.01$ and 2.16 in $ASP_{HS}$ and $6ASP_{HS}$). These could be due to varying levels of bicarbonate available to coordinate to calcium, although the changes are relatively small.

The difference between AAs and 6AAs on calcium (bi)-carbonate solutions is subtle. The crucial difference is that, in the oligopeptides, all the chemical functional groups available for solute association are present together in one molecule. At low salinity, we should expect to see relatively large species forming in solution, especially given the observations of AA effects at low

salinity. This could be the reason why small molecule additives are less effective at stabilizing dense liquid $CaCO_3$ than polymers. No evidence is found that increasing the size of the additive had a significant effect on the formation of $CaCO_3$ ionic networks in water. However, as in simulations of assembled AAs, the additives provide environments where $CaCO_3$ can associate. Association, either through electrostatic interactions or by hydrogen bonding also liberates water from solute monomers, which will have a stabilizing effect for the system, particularly when initial solute concentrations are relatively high.

## CONCLUSIONS

All of the AA additives slightly affected the speciation of calcium (bi)carbonate solutions at pH and concentration levels that targeted typical experimental conditions to study phase separation. AAs provided coordination sites for cations and anions to bind. Association of bicarbonate to $-NH_x$ groups suggests that nucleation might be regulated not just by controlling $c(Ca)$, but also $c((H)CO_3)$. In addition, while AAs did not significantly affect the formation of liquidlike $CaCO_3$ networks, they do provide sites for the dynamic coordination of these entities. This was most apparent on solution decomposition at high salinity. $CaCO_3$ was found predominantly at the interface between extended AA assemblies and surrounding solution. The topologies of these assemblies appear to have some control on the coordination of calcium in the clusters. Furthermore, there was a clear decrease in the mobility of calcium when bound at the edge of these assemblies, even when compared to highly coordinated calcium (by carbonate) in the core of $CaCO_3$ ionic networks. Increasing the size of additives by forming oligopeptides did have some subtle effects on the activities of additives in solution. This was caused by the loss of chemical sites to bind to solute monomers in the formation of the peptide chain. However, generally, there was no significant difference, compared to AA simulations at high supersaturations. We do note that, at concentrations closer to experimental conditions, the high density of solute monomer binding sites in a peptide or polymer will be important and is likely to make these additives more effective than small amino acids, where some degree of assembly will be necessary to control the $CaCO_3$ networks that also form in solution.

The mechanisms for control of calcium (bi)carbonate association by additives presented here are consistent with experiments showing that poly(ASP) stabilizes a dense liquid phase of calcium carbonate before precipitation of an amorphous solid.[32,89] Ion self-diffusion coefficients in extended networks in this work compare favorably with those measured for dense liquid phases.[30] Sebastiani et al.[32] suggested that the position of the liquid−liquid binodal for phase separation in calcium carbonate solutions is unaffected by the addition of poly(ASP), but the presence of the polymer delays the formation of amorphous calcium carbonate. Our work supports this and provides mechanistic insight into how the existence of liquidlike calcium carbonate networks may be prolonged. Bewernitz et al.[30] postulated that the negatively charged carboxylate groups sequester calcium from solution to inhibit the precipitation of a solid phase. Cantaert et al.[90] found that the positively charged additive, poly(allylamine hydrochloride), also stabilized thin films of liquidlike calcium carbonate. They suggested a similar mechanism of "microphase separation", followed by coalescence of droplets driven by an influx of carbonate into a solution containing the polymer and calcium. Our work shows that calcium (bi)carbonate networks can

associate with organic additives in solution. The association appears to be thermodynamically favorable and the resulting reduced mobility of ions means that the additives stabilize these networks, which would otherwise transform more quickly into solid calcium carbonate. The number of solute binding sites and their availability to associate with calcium carbonate networks (requiring some level of conformational flexibility) is likely to be important for stabilizing a dense liquid phase. The charge of the additive appears to be less important, but the charge distribution in additives might help to control the topology of the associating calcium carbonate networks and therefore inhibit solid precipitation.

## ASSOCIATED CONTENT

**ⓈI Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.cgd.9b01693.

> DL_POLY input force field files for the systems simulated in this work (ZIP)

> Movie showing the formation of clusters (MPG)

> Discussion of ion pair−ion pair association in umbrella sampling calculations; supplemental figures and tables (PDF)

## AUTHOR INFORMATION

**Corresponding Author**

**Aaron R. Finney** − *Department of Materials Science and Engineering, University of Sheffield, Sheffield S1 3JD, United Kingdom; Department of Chemical Engineering, University College London, London WC1E 6BT, United Kingdom;* ⓘ orcid.org/0000-0002-1456-5892; Email: a.finney@ucl.ac.uk

**Authors**

**Riccardo Innocenti Malini** − *Laboratory for Biomimetic Membranes and Textiles, EMPA, Swiss Federal Laboratories for Materials Science and Technology, St. Gallen 9014, Switzerland;* ⓘ orcid.org/0000-0003-4366-781X

**Colin L. Freeman** − *Department of Materials Science and Engineering, University of Sheffield, Sheffield S1 3JD, United Kingdom*

**John H. Harding** − *Department of Materials Science and Engineering, University of Sheffield, Sheffield S1 3JD, United Kingdom*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.cgd.9b01693

**Author Contributions**

All authors designed the research. A.R.F. and R.I.M. performed the simulations and analyzed data. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

**Notes**

The authors declare no competing financial interest.

## ■ ABBREVIATIONS

MD, molecular dynamics; PMF, potential of mean force; PNC, prenucleation cluster; US, umbrella sampling; RDF, radial distribution function; ACC, amorphous calcium carbonate

## ■ REFERENCES

(1) Wegst, U. G. K.; Bai, H.; Saiz, E.; Tomsia, A. P.; Ritchie, R. O. Bioinspired Structural Materials. *Nat. Mater.* **2015**, *14*, 23−36.

(2) Zan, G.; Wu, Q. Biomimetic and Bioinspired Synthesis of Nanomaterials/Nanostructures. *Adv. Mater.* **2016**, *28*, 2099−2147.

(3) Mao, L.-B.; Gao, H.-L.; Yao, H.-B.; Liu, L.; Cölfen, H.; Liu, G.; Chen, S.-M.; Li, S.-K.; Yan, Y.-X.; Liu, Y.-Y.; Yu, S.-H. Synthetic Nacre by Predesigned Matrix-Directed Mineralization. *Science* **2016**, *354*, 107−110.

(4) Lowenstam, H. A.; Weiner, S. *On Biomineralization*; Oxford University Press, 1989.

(5) Reimer, T.; Dempster, T.; Wargelius, A.; Fjelldal, P. G.; Hansen, T.; Glover, K. A.; Solberg, M. F.; Swearer, S. E. Rapid Growth Causes Abnormal Vaterite Formation in Farmed Fish Otoliths. *J. Exp. Biol.* **2017**, *220*, 2965−2969.

(6) Wightman, R.; Wallis, S.; Aston, P. Leaf Margin Organisation and the Existence of Vaterite-Producing Hydathodes in the Alpine Plant Saxifraga Scardica. *Flora* **2018**, *241*, 27−34.

(7) Beniash, E.; Aizenberg, J.; Addadi, L.; Weiner, S. Amorphous Calcium Carbonate Transforms into Calcite during Sea Urchin Larval Spicule Growth. *Proc. R. Soc. London, Ser. B* **1997**, *264*, 461−465.

(8) Addadi, L.; Raz, S.; Weiner, S. Taking Advantage of Disorder: Amorphous Calcium Carbonate and Its Roles in Biomineralization. *Adv. Mater.* **2003**, *15*, 959−970.

(9) Weiner, S.; Addadi, L. Crystallization Pathways in Biomineralization. *Annu. Rev. Mater. Res.* **2011**, *41*, 21−40.

(10) Politi, Y.; Metzler, R. A.; Abrecht, M.; Gilbert, B.; Wilt, F. H.; Sagi, I.; Addadi, L.; Weiner, S.; Gilbert, P. U. P. A. Transformation Mechanism of Amorphous Calcium Carbonate into Calcite in the Sea Urchin Larval Spicule. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17362−17366.

(11) Gong, Y. U. T.; Killian, C. E.; Olson, I. C.; Appathurai, N. P.; Amasino, A. L.; Martin, M. C.; Holt, L. J.; Wilt, F. H.; Gilbert, P. U. P. A. Phase Transitions in Biogenic Amorphous Calcium Carbonate. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6088−6093.

(12) Mass, T.; Giuffre, A. J.; Sun, C.-Y.; Stifler, C. A.; Frazier, M. J.; Neder, M.; Tamura, N.; Stan, C. V.; Marcus, M. A.; Gilbert, P. U. P. A. Amorphous Calcium Carbonate Particles Form Coral Skeletons. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E7670−E7678.

(13) Sun, J.; Bhushan, B. Hierarchical Structure and Mechanical Properties of Nacre: A Review. *RSC Adv.* **2012**, *2*, 7617−7632.

(14) Weiner, S.; Addadi, L.; Wagner, H. D. Materials Design in Biology. *Mater. Sci. Eng., C* **2000**, *11*, 1−8.

(15) Peled-Kamar, M.; Hamilton, P.; Wilt, F. H. Spicule Matrix Protein LSM34 Is Essential for Biomineralization of the Sea Urchin Spicule. *Exp. Cell Res.* **2002**, *272*, 56−61.

(16) Wilt, F.; Croker, L.; Killian, C. E.; McDonald, K. Role of LSM34/SpSM50 Proteins in Endoskeletal Spicule Formation in Sea Urchin Embryos. *Invertebr. Biol.* **2008**, *127*, 452−459.

(17) Mann, K.; Wilt, F. H.; Poustka, A. J. Proteomic Analysis of Sea Urchin (Strongylocentrotus Purpuratus) Spicule Matrix. *Proteome Sci.* **2010**, *8*, 33.

(18) Freeman, C. L.; Harding, J. H.; Quigley, D.; Rodger, P. M. Structural Control of Crystal Nuclei by an Eggshell Protein. *Angew. Chem., Int. Ed.* **2010**, *49*, 5135−5137.

(19) Nudelman, F.; Gotliv, B. A.; Addadi, L.; Weiner, S. Mollusk Shell Formation: Mapping the Distribution of Organic Matrix Components Underlying a Single Aragonitic Tablet in Nacre. *J. Struct. Biol.* **2006**, *153*, 176−187.

(20) Keene, E. C.; Evans, J. S.; Estroff, L. A. Matrix Interactions in Biomineralization: Aragonite Nucleation by an Intrinsically Disordered Nacre Polypeptide, N16N, Associated with a β-Chitin Substrate. *Cryst. Growth Des.* **2010**, *10*, 1383−1389.

(21) Ponce, C. B.; Evans, J. S. Polymorph Crystal Selection by N16, an Intrinsically Disordered Nacre Framework Protein. *Cryst. Growth Des.* **2011**, *11*, 4690−4696.

(22) Amos, F. F.; Ponce, C. B.; Evans, J. S. Formation of Framework Nacre Polypeptide Supramolecular Assemblies That Nucleate Polymorphs. *Biomacromolecules* **2011**, *12*, 1883−1890.

(23) Rao, A.; Drechsler, M.; Schiller, S.; Scheffner, M.; Gebauer, D.; Cölfen, H. Stabilization of Mineral Precursors by Intrinsically Disordered Proteins. *Adv. Funct. Mater.* **2018**, *28*, 1802063.

(24) Rao, A.; Seto, J.; Berg, J. K.; Kreft, S. G.; Scheffner, M.; Cölfen, H. Roles of Larval Sea Urchin Spicule SM50 Domains in Organic Matrix Self-Assembly and Calcium Carbonate Mineralization. *J. Struct. Biol.* **2013**, *183*, 205−215.

(25) Pendola, M.; Jain, G.; Huang, Y.-C.; Gebauer, D.; Evans, J. S. Secrets of the Sea Urchin Spicule Revealed: Protein Cooperativity Is Responsible for ACC Transformation, Intracrystalline Incorporation, and Guided Mineral Particle Assembly in Biocomposite Material Formation. *ACS Omega* **2018**, *3*, 11823−11830.

(26) Bahn, S. Y.; Jo, B. H.; Choi, Y. S.; Cha, H. J. Control of Nacre Biomineralization by Pif80 in Pearl Oyster. *Sci. Adv.* **2017**, *3*, e1700765.

(27) Su, J.; Zhu, F.; Zhang, G.; Wang, H.; Xie, L.; Zhang, R. Transformation of Amorphous Calcium Carbonate Nanoparticles into Aragonite Controlled by ACCBP. *CrystEngComm* **2016**, *18*, 2125−2134.

(28) Chen, L.; Huang, F.; Li, S.; Shen, Y.; Xie, A.; Pan, J.; Zhang, Y.; Cai, Y. Biomimetic Synthesis of Aragonite Superstructures Using Hexamethylenetetramine. *J. Solid State Chem.* **2011**, *184*, 2825−2833.

(29) Chen, C.-L.; Qi, J.; Tao, J.; Zuckermann, R. N.; DeYoreo, J. J. Tuning Calcite Morphology and Growth Acceleration by a Rational Design of Highly Stable Protein-Mimetics. *Sci. Rep.* **2015**, *4*, 6266.

(30) Bewernitz, M. A.; Gebauer, D.; Long, J.; Cölfen, H.; Gower, L. B. A Metastable Liquid Precursor Phase of Calcium Carbonate and Its Interactions with Polyaspartate. *Faraday Discuss.* **2012**, *159*, 291−312.

(31) Zou, Z.; Bertinetti, L.; Politi, Y.; Fratzl, P.; Habraken, W. J. E. M. Control of Polymorph Selection in Amorphous Calcium Carbonate Crystallization by Poly(Aspartic Acid): Two Different Mechanisms. *Small* **2017**, *13*, 1603100.

(32) 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. *Angew. Chem., Int. Ed.* **2017**, *56*, 490−495.

(33) De Yoreo, J. J.; Gilbert, P. U. P. A.; Sommerdijk, N. A. J. M.; Penn, R. L.; Whitelam, S.; Joester, D.; Zhang, H.; Rimer, J. D.; Navrotsky, A.; Banfield, J. F.; Wallace, A. F.; Michel, F. M.; Meldrum, F. C.; Cölfen, H.; Dove, P. M. Crystallization by Particle Attachment in Synthetic, Biogenic, and Geologic Environments. *Science* **2015**, *349*, aaa6760.

(34) Gebauer, D.; Völkel, A.; Cölfen, H. Stable Prenucleation Calcium Carbonate Clusters. *Science* **2008**, *322*, 1819−1822.

(35) Demichelis, R.; Raiteri, P.; Gale, J. D.; Quigley, D.; Gebauer, D. Stable Prenucleation Mineral Clusters Are Liquid-like Ionic Polymers. *Nat. Commun.* **2011**, *2*, 590.

(36) 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.

(37) Smeets, P. J. M.; Finney, A. R.; Habraken, W. J. E. M.; Nudelman, F.; Friedrich, H.; Laven, J.; De Yoreo, J. J.; Rodger, P. M.; Sommerdijk, N. A. J. M. A Classical View on Nonclassical Nucleation. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E7882−E7890.

(38) Wolf, S. E.; Leiterer, J.; Kappl, M.; Emmerling, F.; Tremel, W. Early Homogenous Amorphous Precursor Stages of Calcium

Carbonate and Subsequent Crystal Growth in Levitated Droplets. *J. Am. Chem. Soc.* **2008**, *130*, 12342−12347.

(39) Zou, Z.; Habraken, W. J. E. M.; Bertinetti, L.; Politi, Y.; Gal, A.; Weiner, S.; Addadi, L.; Fratzl, P. On the Phase Diagram of Calcium Carbonate Solutions. *Adv. Mater. Interfaces* **2017**, *4*, 1600076.

(40) Gehrke, N.; Cölfen, H.; Pinna, N.; Antonietti, M.; Nassif, N. Superstructures of Calcium Carbonate Crystals by Oriented Attachment. *Cryst. Growth Des.* **2005**, *5*, 1317−1319.

(41) Li, X. Q.; Zeng, H. C. Calcium Carbonate Nanotablets: Bridging Artificial to Natural Nacre. *Adv. Mater.* **2012**, *24*, 6277−6282.

(42) Hu, Q.; Nielsen, M. H.; Freeman, C. L.; Hamm, L. M.; Tao, J.; Lee, J. R. I.; Han, T. Y. J.; Becker, U.; Harding, J. H.; Dove, P. M.; De Yoreo, J. J. The Thermodynamics of Calcite Nucleation at Organic Interfaces: Classical vs. Non-Classical Pathways. *Faraday Discuss.* **2012**, *159*, 509−523.

(43) Henzler, K.; Fetisov, E. O.; Galib, M.; Baer, M. D.; Legg, B. A.; Borca, C.; Xto, J. M.; Pin, S.; Fulton, J. L.; Schenter, G. K.; Govind, N.; Siepmann, J. I.; Mundy, C. J.; Huthwelker, T.; De Yoreo, J. J. Supersaturated Calcium Carbonate Solutions Are Classical. *Sci. Adv.* **2018**, *4*, eaao6283.

(44) Kellermeier, M.; Gebauer, D.; Melero-García, E.; Drechsler, M.; Talmon, Y.; Kienle, L.; Cölfen, H.; García-Ruiz, J. M.; Kunz, W. Colloidal Stabilization of Calcium Carbonate Prenucleation Clusters with Silica. *Adv. Funct. Mater.* **2012**, *22*, 4301−4311.

(45) Finney, A. R.; Rodger, P. M. Probing the Structure and Stability of Calcium Carbonate Pre-Nucleation Clusters. *Faraday Discuss.* **2012**, *159*, 47.

(46) Raiteri, P.; Demichelis, R.; Gale, J. D.; Kellermeier, M.; Gebauer, D.; Quigley, D.; Wright, L. B.; Walsh, T. R. Exploring the Influence of Organic Species on Pre- and Post-Nucleation Calcium Carbonate. *Faraday Discuss.* **2012**, *159*, 61−85.

(47) Picker, A.; Kellermeier, M.; Seto, J.; Gebauer, D.; Cölfen, H. The Multiple Effects of Amino Acids on the Early Stages of Calcium Carbonate Crystallization. *Z. Kristallogr. - Cryst. Mater.* **2012**, *227*, 744−757.

(48) Innocenti Malini, R.; Finney, A. R.; Hall, S. A.; Freeman, C. L.; Harding, J. H. The Water−Amorphous Calcium Carbonate Interface and Its Interactions with Amino Acids. *Cryst. Growth Des.* **2017**, *17*, 5811−5822.

(49) Saharay, M.; Kirkpatrick, R. J. Ab Initio and Metadynamics Studies on the Role of Essential Functional Groups in Biomineralization of Calcium Carbonate and Environmental Situations. *Phys. Chem. Chem. Phys.* **2014**, *16*, 26843−26854.

(50) Tobler, D. J.; Blanco, J. D. R.; Dideriksen, K.; Sand, K. K.; Bovet, N.; Benning, L. G.; Stipp, S. L. S. The Effect of Aspartic Acid and Glycine on Amorphous Calcium Carbonate (ACC) Structure, Stability and Crystallization. *Procedia Earth Planet. Sci.* **2014**, *10*, 143−148.

(51) Orme, C. A.; Noy, A.; Wierzbicki, A.; McBride, M. T.; Grantham, M.; Teng, H. H.; Dove, P. M.; DeYoreo, J. J. Formation of Chiral Morphologies through Selective Binding of Amino Acids to Calcite Surface Steps. *Nature* **2001**, *411*, 775−779.

(52) Wolf, S. E.; Loges, N.; Mathiasch, B.; Panthöfer, M.; Mey, I.; Janshoff, A.; Tremel, W. Phase Selection of Calcium Carbonate through the Chirality of Adsorbed Amino Acids. *Angew. Chem., Int. Ed.* **2007**, *46*, 5618−5623.

(53) Fischer, V.; Landfester, K.; Muñoz-Espí, R. Stabilization of Calcium Oxalate Metastable Phases by Oligo(l-Glutamic Acid): Effect of Peptide Chain Length. *Cryst. Growth Des.* **2011**, *11*, 1880−1890.

(54) Kahlen, J.; Peter, C.; Donadio, D. Molecular Simulation of Oligo-Glutamates in a Calcium-Rich Aqueous Solution: Insights into Peptide-Induced Polymorph Selection. *CrystEngComm* **2015**, *17*, 6863−6867.

(55) Krattiger, P.; Nassif, N.; Völkel, A.; Mastai, Y.; Wennemers, H.; Cölfen, H. Investigation of Active Crystal Morphogenesis Peptide Sequences from Peptide Libraries by Crystallization on Peptide Functionalized Beads. *Colloids Surf., A* **2010**, *354*, 218−225.

(56) Kašparová, P.; Antonietti, M.; Cölfen, H. Double Hydrophilic Block Copolymers with Switchable Secondary Structure as Additives for Crystallization Control. *Colloids Surf., A* **2004**, *250*, 153−162.

(57) Aschauer, U.; Spagnoli, D.; Bowen, P.; Parker, S. C. Growth Modification of Seeded Calcite Using Carboxylic Acids: Atomistic Simulations. *J. Colloid Interface Sci.* **2010**, *346*, 226−231.

(58) ten Wolde, P. R.; Frenkel, D. Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations. *Science* **1997**, *277*, 1975−1978.

(59) Walsh, M. R.; Koh, C. A.; Sloan, E. D.; Sum, A. K.; Wu, D. T. Microsecond Simulations of Spontaneous Methane Hydrate Nucleation and Growth. *Science* **2009**, *326*, 1095−1098.

(60) Matsumoto, M.; Saito, S.; Ohmine, I. Molecular Dynamics Simulation of the Ice Nucleation and Growth Process Leading to Water Freezing. *Nature* **2002**, *416*, 409−413.

(61) Anwar, J.; Zahn, D. Uncovering Molecular Processes in Crystal Nucleation and Growth by Using Molecular Simulation. *Angew. Chem., Int. Ed.* **2011**, *50*, 1996−2013.

(62) Wu, Y.; Tepper, H. L.; Voth, G. A. Flexible Simple Point-Charge Water Model with Improved Liquid-State Properties. *J. Chem. Phys.* **2006**, *124*, 024503.

(63) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.

(64) Kong, C. L. Combining Rules for Intermolecular Potential Parameters. II. Rules for the Lennard-Jones (12−6) Potential and the Morse Potential. *J. Chem. Phys.* **1973**, *59*, 2464−2467.

(65) Nikitin, A.; Milchevskiy, Y.; Lyubartsev, A. AMBER-Ii: New Combining Rules and Force Field for Perfluoroalkanes. *J. Phys. Chem. B* **2015**, *119*, 14563−14573.

(66) Freeman, C. L.; Harding, J. H.; Cooke, D. J.; Elliott, J. A.; Lardge, J. S.; Duffy, D. M. New Forcefields for Modeling Biomineralization Processes. *J. Phys. Chem. C* **2007**, *111*, 11943−11951.

(67) Todorov, I. T.; Smith, W.; Trachenko, K.; Dove, M. T. DL_POLY_3: New Dimensions in Molecular Dynamics Simulations via Massive Parallelism. *J. Mater. Chem.* **2006**, *16*, 1911−1918.

(68) Bush, I. J.; Todorov, I. T.; Smith, W. A DAFT DL_POLY Distributed Memory Adaptation of the Smoothed Particle Mesh Ewald Method. *Comput. Phys. Commun.* **2006**, *175*, 323−329.

(69) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695−1697.

(70) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(71) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(72) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.

(73) Grossfield, A. "*WHAM: the weighted histogram analysis method*", version 2.07, http://membrane.urmc.rochester.edu/wordpress/?page_id=126.

(74) Salvalaglio, M.; Perego, C.; Giberti, F.; Mazzotti, M.; Parrinello, M. Molecular-Dynamics Simulations of Urea Nucleation from Aqueous Solution. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E6−E14.

(75) Raiteri, P.; Demichelis, R.; Gale, J. D. Thermodynamically Consistent Force Field for Molecular Dynamics Simulations of Alkaline-Earth Carbonates and Their Aqueous Speciation. *J. Phys. Chem. C* **2015**, *119*, 24447−24458.

(76) Guskov, A.; Jensen, S.; Faustino, I.; Marrink, S. J.; Slotboom, D. J. Coupled Binding Mechanism of Three Sodium Ions and Aspartate in the Glutamate Transporter Homologue GltTk. *Nat. Commun.* **2016**, *7*, DOI: 10.1038/ncomms13420.

(77) Oh, S.; Boudker, O. Kinetic Mechanism of Coupled Binding in Sodium-Aspartate Symporter GltPh. *eLife* **2018**, *7*, e37291.

(78) De Stefano, C.; Foti, C.; Gianguzza, A.; Rigano, C.; Sammartano, S. Chemical Speciation of Amino Acids in Electrolyte Solutions Containing Major Components of Natural Fluids. *Chem. Speciation Bioavailability* **1995**, *7*, 1−8.

(79) Carino, A.; Testino, A.; Andalibi, M. R.; Pilger, F.; Bowen, P.; Ludwig, C. Thermodynamic-Kinetic Precipitation Modeling. A Case Study: The Amorphous Calcium Carbonate (ACC) Precipitation Pathway Unravelled. *Cryst. Growth Des.* **2017**, *17*, 2006−2015.

(80) Gebauer, D.; Raiteri, P.; Gale, J. D.; Cölfen, H. On Classical and Non-Classical Views on Nucleation. *Am. J. Sci.* **2018**, *318*, 969−988.

(81) Kellermeier, M.; Picker, A.; Kempter, A.; Cölfen, H.; Gebauer, D. A Straightforward Treatment of Activity in Aqueous $CaCO_3$ Solutions and the Consequences for Nucleation Theory. *Adv. Mater.* **2014**, *26*, 752−757.

(82) Finney, A.; Lectez, S.; Freeman, C.; Harding, J.; Stackhouse, S. Ion Association in Lanthanide Chloride Solutions. *Chem. - Eur. J.* **2019**, chem.201900945.

(83) Raiteri, P.; Schuitemaker, A.; Gale, J. Ion Pairing and Multiple Ion Binding in Calcium Carbonate Solutions Based on a Polarizable AMOEBA Force Field and Ab Initio Molecular Dynamics *J. Phys. Chem. B.* **2020**, DOI: 10.1021/acs.jpcb.0c01582.

(84) Perego, C.; Salvalaglio, M.; Parrinello, M. Molecular Dynamics Simulations of Solutions at Constant Chemical Potential. *J. Chem. Phys.* **2015**, *142*, 144113.

(85) Karmakar, T.; Piaggi, P. M.; Parrinello, M. Molecular Dynamics Simulations of Crystal Nucleation from Solution at Constant Chemical Potential. *J. Chem. Theory Comput.* **2019**, *15*, 6923−6930.

(86) Liu, C.; Wood, G. P. F.; Santiso, E. E. Modelling Nucleation from Solution with the String Method in the Osmotic Ensemble. *Mol. Phys.* **2018**, *116*, 2998−3007.

(87) Bushuev, Y. G.; Finney, A. R.; Rodger, P. M. Stability and Structure of Hydrated Amorphous Calcium Carbonate. *Cryst. Growth Des.* **2015**, *15*, 5269−5279.

(88) Finney, A. R. Nucleation and Dehydration of Calcium Carbonate. Ph.D. Thesis, University of Warwick, 2014.

(89) Gower, L. B.; Odom, D. J. Deposition of Calcium Carbonate Films by a Polymer-Induced Liquid-Precursor (PILP) Process. *J. Cryst. Growth* **2000**, *210*, 719−734.

(90) Cantaert, B.; Kim, Y.-Y.; Ludwig, H.; Nudelman, F.; Sommerdijk, N. A. J. M.; Meldrum, F. C. Think Positive: Phase Separation Enables a Positively Charged Additive to Induce Dramatic Changes in Calcium Carbonate Morphology. *Adv. Funct. Mater.* **2012**, *22*, 907−915.