Structural Features That Distinguish Kinetically Distinct Biomineralization Polypeptides

Sebastiano Collino and John Spencer Evans*

Center for Biomolecular Materials Spectroscopy, Laboratory for Chemical Physics, New York University, 345 East 24th Street, New York, New York 10010

Received January 6, 2007; Revised Manuscript Received February 7, 2007

AP7 and AP24 are mollusk shell proteins which are responsible for aragonite polymorph formation and stabilization within the nacre layer of the Pacific red abalone, *Haliotis rufescens*. It is known that the 30-AA N-terminal mineral modification domains of both proteins (AP7N, AP24N) possess identical multifunctional mineralization capabilities within in vitro assays but differ in terms of rate kinetics, with AP24N > AP7N. In this report, we identify previously unreported molecular features of AP24N and contrast the lowest energy polypeptide backbone structures of AP24N (planar configuration) with that of AP7N ("bent paper clip" configuration) using NMR data and simulated annealing molecular dynamics structure refinement. Like AP7N, we find that AP24N possesses an unfolded conformation, can sequester Ca(II) and other multivalent metal ions, can adsorb onto or within calcite crystals, and possesses anionic and cationic electrostatic "pocket" regions on its molecular surfaces. However, AP24N has some unique features: greater conformational responsiveness to Ca(II), the tendency to form a more planar backbone configuration, and longer anionic and hydrogen-bonding donor/acceptor sequence blocks. We conclude that the presence of unfolded polypeptide conformation, electrostatic surface pockets, and interactive sequence clustering endow both AP7N and AP24N with similar features that lead to comparable effects on crystal morphology and nucleation. However, AP24N possesses longer anionic and hydrogen-bonding sequence clusters and exhibits a tendency to adopt a more planar backbone configuration than AP7N does. We believe that these features facilitate peptide−mineral, peptide−ion, or water cluster interactions, thereby enhancing the mineralization kinetics of AP24N over AP7N.

Introduction

Biomineralization is a dynamic model for composite material processing, for here organic macromolecules are exploited by nature to construct exquisitely structured inorganic solids with fascinating functions, such as protection, magnetoreception, and light transmission and amplification.[1−4] Of the more than 60 different inorganic biominerals produced by organisms,[5,6] calcium carbonate biominerals are perhaps the most abundant in nature and also the most extensively studied. In several species, calcium carbonates are observed to exist not only as different crystalline polymorphs (calcite, vaterite, aragonite)[7−12] but also as an amorphous phase (amorphous calcium carbonate or ACC).[13,14] Under ambient conditions, calcite is the thermodynamically stable polymorph, with aragonite and vaterite both exhibiting greater thermodynamic instability.[7−12] Polymorph selection, stabilization, and transformation are among the least understood aspects of biomineralization, and thus, the ultimate challenge is to deduce the mechanisms utilized by nature for phase transformation/transition, such that they can be transferred successfully to the materials science realm.

In relevant model systems of polymorph stabilization, such as the Pacific red abalone (*Haliotis rufescens*),[11,12] sawtooth penshell (*Atrina serrata*),[15] and Japanese pearl oyster (*Pinctada fucata*),[16] aragonite is routinely synthesized to form mechanically strong nacre layers. It now appears that aragonite formation is under the control of protein superfamilies,[11,12,15−17] and recent efforts have been devoted to the isolation, sequencing, and characterization of these interesting proteins. As part of this overall effort, we initiated molecular studies of the conformation and in vitro mineralization properties of 30 AA N-terminal sequences from the nacre-specific proteins AP7 and AP24 (see Figure 8).[17−21] These two domains are dissimilar in primary sequence and exist as conformationally unfolded macromolecules with subtle structural differences in solution.[18] Surprisingly, both N-terminal sequences are multifunctional and induce identical events within in vitro calcite hillock overgrowth experiments, that is, pinning of acute growth steps, increasing obtuse step velocities, and formation of amorphous-appearing deposits on terrace surfaces of calcite hillocks.[21] However, what distinguishes these two domains is kinetics: obtuse step velocities are 4× faster in the presence of AP24N versus AP7N, and AP24N induces more extensive acute step pinning and nucleation deposits on calcite surfaces compared to AP7N.[21]

The question that we pose is the following: what molecular features allow the same basic mineralization tasks to be performed by both peptides in vitro but "tune" the kinetics of these tasks? To address this question, we approached the problem in two steps. Previously, we demonstrated that AP7N responds on a local structural level to Ca(II) interactions and adsorbs onto exposed calcite surfaces within in vitro mineralization assays.[19] However, no equivalent data set has been available for AP24N. Hence, we first explored the Ca(II) sequestration and mineral adsorption capabilities of AP24N and found that this sequence behaves in a very similar fashion to AP7N. Following this, we wanted to compare both sequences with respect to three-dimensional structural and electrostatic features. To achieve this, we performed simulated annealing molecular dynamics (SA/MD) structure refinement simulations of AP24N and AP7N, using atomic restraints derived from newly acquired

* To whom correspondence should be addressed. E-mail: jse1@nyu.edu.

10.1021/bm0700183 CCC: $37.00
Published on Web 03/24/2007

high-resolution 600 MHz ^{1}H NMR polypeptide data sets. From these simulation studies, we find that AP24N and AP7N both share the following common features: similar unfolded global conformations, surface-accessible electrostatic pockets, and putative anionic and hydrogen-bonding interactive sequence blocks. However, our studies also reveal a critical difference between these two mineralization domains: although both sequences appear to be conformationally similar on a qualitatively level, the potential to form a planar backbone configuration appears to be higher for AP24N than it is for AP7N. We believe that this structural feature, along with the presence of longer anionic and hydrogen bonding donor/acceptor sequence blocks, may be responsible for the enhanced mineralization kinetics of AP24N.

Materials and Methods

Peptide Synthesis and Purification. Free amino termini, C-amide-capped AP24N and AP7N, were synthesized and purified by Dr. Janet Crawford at the Yale University Wm. Keck Biotechnology Peptide Synthesis Facility as described in our earlier work.[18] For our studies, the α-^{15}N-labeled N$^{\alpha}$ Fmoc-L-Asp β-*t*-butyl ester precursor was incorporated at positions D3 and D6 in AP24N during solid-phase synthesis to allow selective identification of Asp residues during Ca(II) ion-binding NMR experiments.[19] Procedures for resin cleavage and peptide purification followed the protocol described in our earlier work.[18,19] The experimental molecular mass for α-^{15}N-labeled AP24N was determined by matrix-assisted laser desorption ionization/time-of-flight-mass spectroscopy (MALDI/TOF−MS) (Yale University) to be 3374.4 Da in agreement with the theoretical molecular mass of 3373.5 Da.

Ion Trap Mass Spectrometry (MS) Metal Ion-Binding Studies. AP24N was dissolved in deionized distilled water and was titrated with microliter volumes of Tris-HCl to make a stock solution (741 μM, pH 7.4). This stock solution was then diluted with deionized distilled water to a final concentration of 10 μM to which appropriate volumes of $CaCl_2$, $CdCl_2$, $LaCl_3$, and $EuCl_3$ stock solutions (each 5 M, 99.99% pure, Sigma/Aldrich, dissolved in deionized distilled water) were added to create peptide/metal ion ratios of 1:10 for each multivalent metal ion series.[19] Ion trap MS experiments were performed on an Agilent LC/MSD 1100 equipped with electrospray ionization source using the following procedure: 1000 μL of each sample solution was manually injected directly into the nebulizer at the rate of 10 μL/min, using a N_2 nebulizing gas pressure of 16 psi, with N_2 drying gas flow of 5 L/min at a temperature of 325 °C. Negative ionization mode was employed with skimmer voltage of −45 V and capillary exit voltage of −70 V.

Solution-State NMR Spectroscopy. Homonuclear ^{1}H−^{1}H and heteronuclear ^{15}N−^{1}H solution-state NMR experiments were performed on a narrow bore Varian INOVA 600 z-PFG digital NMR spectrometer (HCN 5 mm z-PFG Nalorac probe) at 293 K. Pulse field gradient (PFG) total correlation (TOCSY) experiments (t_{mix} = 60, 70, 80 ms) were used for proton scalar coupling assignments, and PFG Nuclear Overhauser Spectroscopy (NOESY) (t_{mix} = 50, 100, 150, 200 ms),^{15}N−^{1}H heteronuclear single quantum coherence (HSQC), ^{15}N−^{1}H−^{1}H NOESY-HSQC (t_{mix} = 150, 200 ms) experiments were used for sequential assignments. In TOCSY, NOESY, and NOESY-HSQC experiments, water suppression was achieved using 3-9-19 *z*-axis PFG gradient schemes.[19] Parallel ^{1}H NMR NMR experiments were performed on a 500 μM AP24N and 450 μM AP7N apo sample[19] under identical solution conditions (i.e., 90% deionized distilled water/10% D_2O, pH 7.5, with each peptide acting as its own buffer).[18,19] In addition, for AP24N, a single-point Ca(II) titration sample [Ca(II):peptide = 2:1] was created and studied for NMR using a stock solution consisting of 99.99% pure $CaCl_2$*$2H_2O$ (Sigma/Aldrich Chemicals) dissolved in deionized distilled water and neutralized to pH 7.5.[19] NMR data were visualized using Sparky software[22] and were processed using VNMRj and NMRPipe[23] software. NMR acquisition and process parameters are reported in the corresponding figure legends.

Solid-State ^{1}H Magic Angle Spinning (MAS) NMR Spectroscopy. To ascertain whether AP24N interacts with calcium carbonate crystals during the assay, we utilized ^{1}H NMR MAS experiments to detect the presence of bound peptide on geological calcite crystals that were "rescued" from the calcite overgrowth assays.[19] Using the same assay conditions described in earlier reports,[17,18] we ran three parallel assays: a negative control (i.e., no polypeptide added) and two parallel assays run in the presence of 1.7×10^{-4} M AP24N polypeptide. Each assay condition utilized 200 mg of calcite fragments, obtained from the same geological parent crystal, and all mineralization assays were 90 min in duration.[17−19] At the conclusion of the assay, all calcite fragments were washed and treated using the differential washing procedure outlined for AP7N;[19] typically, each washing step had a duration of 5 min.[19] After the washing procedures were completed, the crystals were centrifuged at 5000*g* for 3 min at room temperature, the supernatant was aspirated off, and 10 mg wet crystals was packed into individual 3.2-mm ZrO_3 pencil MAS rotors. One-dimensional ^{1}H MAS NMR experiments were conducted at 25 °C on a Varian INOVA 600 narrow bore digital NMR spectrometer, using a Chemagnetics HXY 3.2 mm MAS Balun T3 probehead and a rotor spinning rate of 10 kHz. A WALTZ-16 presaturation pulse was applied during the recovery period to suppress the residual HOD solvent signal. Parallel ^{1}H MAS NMR experiments were also run on a 2.5-mg lyophilized AP24N peptide sample that was exposed to D_2O vapor (25 °C, 48 h in an airtight container). This created a hydrated AP24N sample that was determined to be 20% w/w D_2O compared to a dehydrated lyophilized polypeptide sample.[19] NMR data were processed using the VNMRj software. Acquisition and processing parameters are described in the figure legends.

Simulated Annealing Molecular Dynamics (SA/MD) Structure Refinement. Initial structure coordinates for AP7N and AP24N were generated in the Biopolymer module (Biosym InsightII package, Accelrys, San Diego, CA) using the CVFF force field and extended conformation ($\varphi = 180°$ $\Psi = 180°$). Side chain and main chain partial atomic charges (protonated Lys, Arg, deprotonated Asp, Glu, α-NH_3^+, and COO^- termini) were explicitly represented to emulate a neutral pH scenario. NOE values were obtained from NOESY spectra (t_{mix} = 50, 100, 150, and 200 ms) for each polypeptide, and these restraints were used to set corresponding H−H distances at specific locations within each sequence. NMR distance restraints were applied as either strong (1.8−2.5 Å), medium (2.5−3.5 Å), or weak (3.5−5.0 Å) soft-square potentials with force constant of 50 kcal/mol-Å^2. Inter- and intraresidue interproton distances were calculated using a reference distance of 2.4 Å for Ala α-CH−β-CH protons.[23] Pseudoatoms defining the centroids of the methyl and methylene groups were defined in the NOE restraint file. Because of the conformational lability of each sequence, only a total of 75 and 96 NOE restraints were identified for AP24N and AP7N, respectively. The remaining H−H distances were left unrestrained in the starting structures. SA/MD runs (10 total) were conducted using the XPLOR software package. Specifically, all simulations employed a distance-dependent dielectric of 78.5 Debye (representing implicit solvation), a nonbonding cutoff of 4.5 Å (to determine electrostatic interactions using a switching function), and the Verlet algorithm integrator (1 femtosecond per step) along with a temperature-coupling option. Essentially, this approach is a Langevin-type dynamics method with zero random forces and a scaled friction coefficient. Following initial minimization to convergence (Fletcher−Powell algorithm), 30 psec of equilibrium dynamics was performed at 300 K, and then the system was incrementally ramped to 2000 K to sample conformational space. The system was cooled from 2000 to 300 K at 100 K per step, and the resulting conformers were then minimized to convergence. Each SA/MD run generated a pool of 1000 structures in the initial ensemble, with 100 structures chosen from this pool on the basis of the fewest NOE violations. From this 100 member subset, 10 lowest energy

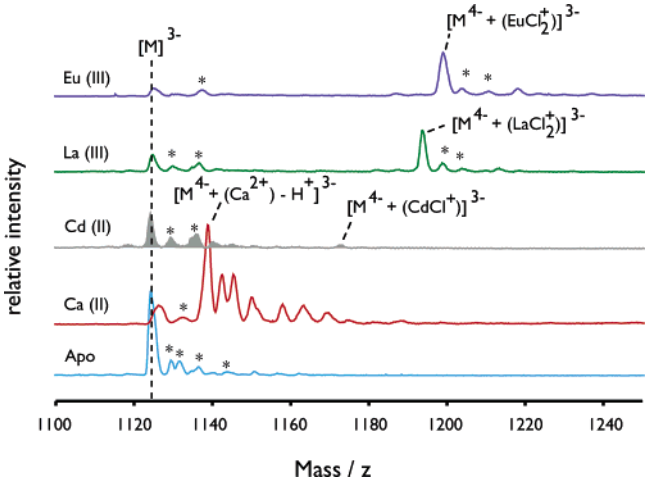

Figure 1. Ion trap mass spectra (negative ionization mode) of apo-AP24N and AP24N in the presence of 10 fold molar excess of $CaCl_2$, $CdCl_2$, $LaCl_3$, and $EuCl_3$. Major AP24N:metal ion adduct species are noted on the spectra, and the observed and theoretical mass/Z values are as follows: apo $[M]^{3-}$, observed = 1124.4 Da, theoretical = 1123.8 Da; $[M^{4-} + (Ca^{2+}) - H^{+}]^{3-}$, observed = 1138.6 Da, theoretical = 1136.5 Da; $[M^{4-} + (CdCl^{+})]^{3-}$, observed = 1173.8 Da, theoretical = 1172.7 Da; $[M^{4-} + (LaCl_2^{+})]^{3-}$, observed = 1193.7 Da, theoretical = 1193.4 Da; $[M^{4-} + (EuCl_2^{+})]^{3-}$, observed = 1199.1 Da, theoretical = 1197.8 Da. The * symbol denotes AP24 N − Na ion adducts.

structures were selected to form the conformer library for further analyses.

Lowest energy coordinate files were visualized using the MolMol and PyMOL packages, and solvation energies (vacuum to solvent transfer) and surface plots were created using the Adaptive Poisson Boltzmann Solver (APBS) solvation calculation module[24] within the Python Molecular Viewer software package v. 1.4.3. Solvation parameters include a linearized Poisson−Boltzmann method using single Debye−Huckel boundary conditions and a spline-based surface-smoothing method, a protein dielectric of 2.0, a solvent dielectric of 78.5, grid dimensions of 65 Å × 65 Å × 65 Å, grid spacing of 0.48 Å × 0.57 Å × 0.63 Å, a solvent radius of 1.4 Å, and a system temperature of 298 K. A system ionic strength of 0.01 M was employed to mimic the low ionic strength conditions of our NMR samples.

Results

Evidence for AP24N−Metal Ion Complexation. Previously, atomic force microscopy (AFM) investigations demonstrated that both AP24N and AP7N subdomains induce mineral precipitation on calcite hillock terraces, suggesting that both sequences may sequester ion clusters and promote calcium carbonate nucleation.[21] We investigated AP24N interactions with Ca(II) and Ca(II) multivalent ion analogues [i.e., Cd(II)(112.4 amu), La(III)(138.9 amu), Eu(III) (157.9 amu)] using ion trap mass spectrometry and the metal ion-induced mass shift method.[9,19,25,26] These larger mass Ca(II) ion analogues generate resolvable, unambiguous mass shifts (≈100 Da) for peptide:metal ion complexes[19] and have been successfully employed in studying Ca(II) interaction sites within several Ca(II) binding proteins.[19,27−29]

As shown in Figure 1, apo AP24N is detected as a quadruply $[M^{4-}]$ and triply $[M^{3-}]$ charged adduct species, but because the signal intensities for the triply charged species were higher in our system, we focused our attention on these species. In the presence of 10 fold excess of each metal salt, we observe both the apo $[M^{3-}]$ and the corresponding 1:1 adduct species (Figure 1), with excellent agreement to theoretical mass/Z values for the metal−chloride salt:peptide complexes without associated waters of hydration (see legend to Figure 1). Other unidentified adduct species which appear in the ion trap MS spectra most likely correspond to adduct species which vary with regard to stoichiometric levels of associated water, residual sodium ions, or the presence of peptide fragmentation species.[19,21,25,26] We were unable to detect metal ion:peptide complexes > 1:1, either because of the more complex adduct species being inherently unstable in the gas phase or because of the limitations in our instrumentation detection range.[19,21,25,26] Again, as noted in our earlier metal ion-induced mass shift experiments, the adduct signal intensities varied for each type of metal ion under investigation, most likely because of the relative stabilities of each type of metal ion adduct in the gas phase, which varies from sequence to sequence.[19,21] Thus, our ion trap experiments reveal that AP24N can form complexes with Ca(II) and Ca(II) metal ion analogues, and this verifies that AP24N, like AP7N, is capable of promoting metal ion sequestration in solution.

The Conformation of AP24N in the Presence of Ca(II). Previous studies of mineral modification polypeptides reveal that these sequences persist in an unfolded state in the presence of excess metal ions.[17−21,30−33] Presumably, this feature allows these macromolecules to persist in an unfolded state in supersaturated solution and to interact with exposed interfaces or to promote ion clustering. To determine if this molecular trait is also shared by AP24N, we investigated the secondary structure of AP24N in the presence (i.e., 2:1 Ca(II) AP24N complex) and absence (i.e., apo state) of Ca(II) using NMR spectroscopy. First, with the use of NMR digital filtering processing techniques and higher field magnet (600 MHz), we were able to detect and resolve additional ^{1}H NMR resonances for the apo-AP24N subdomain that were not originally reported in our earlier 500 MHz studies, thereby improving upon the original sequence-specific ^{1}H, ^{15}N NMR backbone and side chain assignments for apo-AP24N (Figures 2, 3, see Supporting Information).[18] A comparison of the apo- and 2:1 Ca(II):AP24N polypeptide NMR data sets reveals the absence of α-helix or β-hairpin specific long-range $d_{NN(i,j)}$ backbone or $d_{sc(i,j)}$ side chain NOEs[34−36] in either sample (Figures 2, 3). These findings indicate that, similar to AP7N, the AP24N polypeptide does not convert into an α-helix or β-strand conformation in the presence of Ca(II); rather, the peptide persists in a conformation qualitatively similar to the apo state (i.e., combination of random coil and extended secondary structure).[18,19]

Evidence for Ca(II)-Induced Motional Perturbations within AP24N. Although AP24N does not experience global conformational change in the presence of Ca(II) ions, we cannot rule out the possibility that local conformational events occur within the sequence in response to Ca(II) binding. In fact, these phenomena have been detected in other mineral modification subdomains, including AP7N[19] and the Lustrin A D4 polyanionic segment,[33] and may be relevant to the function of each subdomain sequence within a Ca(II) rich environment such as that found within the developing mollusk shell.

To address this issue, we examined structurally sensitive regions of the AP24N NOESY spectra (Figures 2, 3). An overlay comparison of the NH−CH_α and NH−NH fingerprint regions[19,32] from the NOESY spectra of apo- and 2:1 Ca(II):AP24N polypeptide samples indicates that Ca(II) addition led to observable changes in $d_{\alpha N(i,i)}$ intraresidue cross-peak intensities or resonance frequency shifts for A1, D4, S8, S9, G10, L11, Y15, N16, N25, and K28. Moreover, we also note intensity or resonance frequency shifts for D2−A1, D6−E5, C12−L11, N16−Y15, V19−N18, R22−T21, N24−P23, K28−P27, F30−

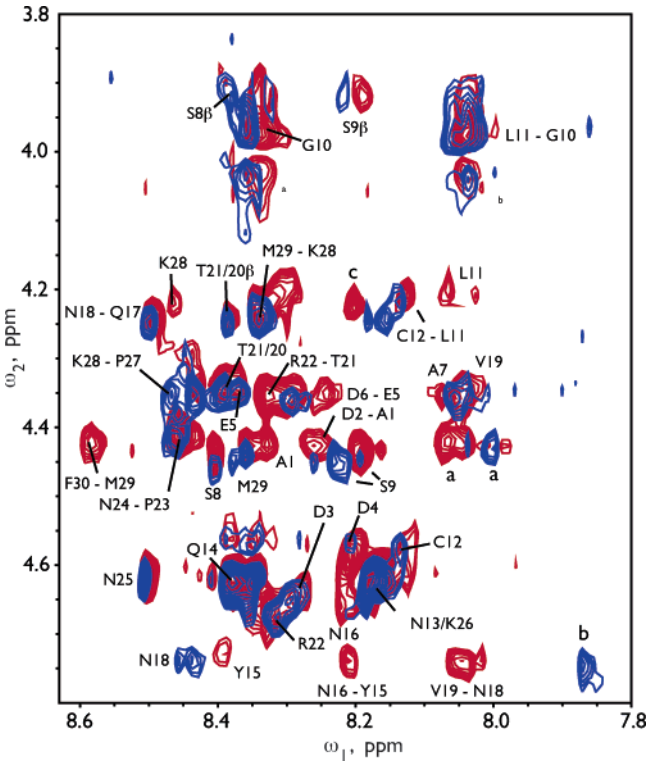

Figure 2. Overlay of 200-ms mixing time PFG-NOESY NH−CH_α fingerprint spectra obtained for apo-AP24N (red) and 2:1 Ca(II):AP24N (blue), 293 K. Both spectra were acquired using identical acquisition parameters, sample volumes, and peptide concentrations. Acquisition parameters include a spectral window of 10 ppm, ^{1}H 90° pulse = 5.1 μs, 64 transients per experiment, with 1024 and 512 complex points acquired in t_2 and t_1, respectively. The carrier is centered on the water resonance. *Z*-gradient parameters: $G_1 = 10.6$ G/cm, $G_2 = 10.6$ G/cm, each with a duration of 1 ms and a stabilization time of 500 μs. The hypercomplex States-TPPI phase-sensitive method was utilized to process both 2D spectra, with zero filling in the F2 dimension. Proton chemical shifts are referenced from internal d_4-TSP. Lower case letters a, b, and c indicate chemical or conformational exchange cross-peaks.

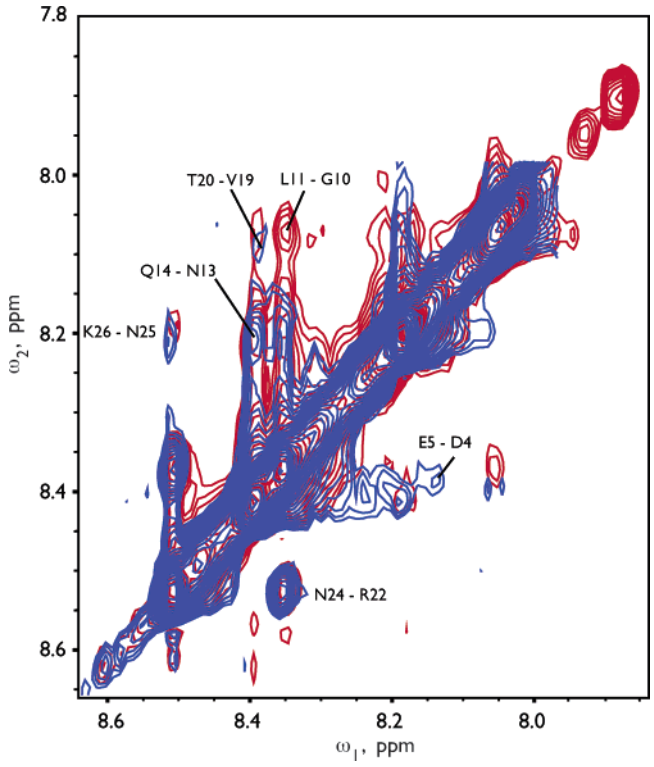

Figure 3. Overlay of 200-ms mixing time PFG-NOESY NH−NH fingerprint spectra obtained for apo-AP24N (red) and 2:1 Ca(II):AP24N (blue), 293 K. Acquisition and processing parameters are identical to those presented in Figure 2.

M29 $d_{\alpha N(i,i+1)}$ and E5−D4, L11−G10, T20−V19 $d_{NN(i,i+1)}$ sequential interresidue connectivities. These findings are significant in that these Ca(II)-induced changes in AP24N NOE cross-peak intensities occur throughout the length of the polypeptide molecule, a feature not observed in AP7N.[19] As described in our previous work,[19,33] these NOE intensity changes most likely reflect alterations in polypeptide backbone motion and dynamics or backbone intra- or interresidue H−H distances[33−35] in response to metal ion interactions within the AP24N sequence.

Another unique, previously unreported feature[18] is the presence of conformational exchange cross-peaks in the NH−CHα fingerprint NOESY spectra for AP24N in either the apo or 2:1 Ca(II) state (Figure 2, note letters a, b, c). We believe that the conformational exchange cross-peaks for apo AP24N were previously undetected because of their inherently weak intensities and, hence, were unobservable without the use of digital filtering processing techniques and higher magnetic fields. These cross-peaks correlate with residues A1, A7, S8, Q17, N18, and V19 (Figures 2, 3), and it appears that these residues are experiencing exchange between similar backbone torsion angle states in either apo-AP24N or 2:1 Ca(II):AP24N on the NMR time scale (milliseconds to seconds). This finding supports our previous observations regarding the conformational lability of the apo-AP24N sequence.[18] What is interesting about these conformational exchange NOE cross-peaks is that their intensities are either diminished or enhanced by the presence or absence of Ca(II) (Figure 2). Similar results were noted for the Lustrin A D4 mineral modification domain,[33] but slightly different results were obtained for AP7N, where the presence of Ca(II) led to a reduction in NOE cross-peak intensities exclusively and did not lead to cross-peak enhancement.[19] Thus, we conclude that Ca(II) interactions within AP24N have an impact not only on residue-specific motion- or distance-related NMR parameters but also on the conformational exchange equilibrium involving specific amino acid residues.

Because of ^{1}H chemical shift overlap, we found that PFG ^{1}H−^{1}H NOESY experiments were unsatisfactory with regard to unambiguously detecting expected changes in NOE cross-peak intensities for putative Ca(II) interaction sites within the DDDED anionic sequence cluster. Thus, we employed 3-D HSQC-NOESY experiments to observe α-^{15}N and ^{1}H NMR resonances arising from overlapping residues D3 and D6 in selectively labeled ^{15}N apo-AP24N and 2:1 Ca:AP24N (Figure 4). Here, with the addition of Ca(II) and relative to the apo-state, we observe the following: (a) the disappearance of the D3 $d_{\beta N(i,i)}$ intraresidue cross-peak and (b) the disappearance of the D3 $d_{\alpha N(i,i)}$ cross-peak and the upfield ^{1}H resonance frequency shift (≈0.2 ppm) of the D6 $d_{\alpha N(i,i)}$ cross-peak. The reduction in Asp proton NOE intensities, particularly those associated with the beta carbon side chain, most likely arises from motional or distance-dependent perturbations[19,34,35] that occur when Ca(II) interacts with Asp COO^- side chain groups.

To summarize, our NMR experiments indicate the following: (a) Ca(II) ion interactions occur within the DDDED anionic cluster of AP24N (Figures 2−4); (b) Ca(II) interactions apparently involve the D3 residue and to a lesser extent, the D6 residue (Figure 4); (c) Ca(II) ion interactions within the AP24N sequence lead to extensive changes in backbone motion or exchange at nonanionic regions upstream of the DDDED anionic

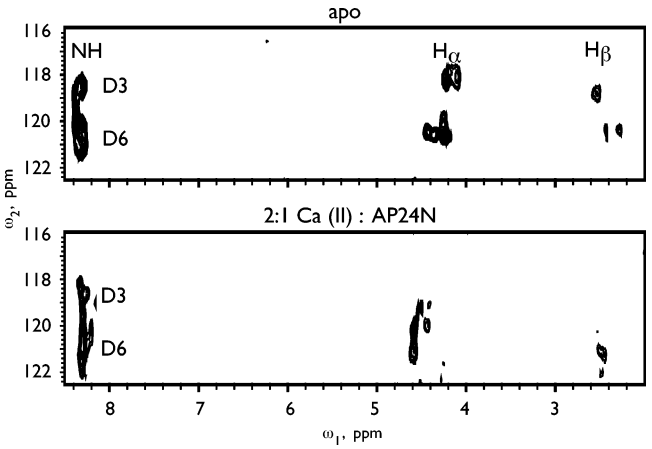

Figure 4. $^{15}N-^{1}H$ NOESY-HSQC spectra of apo-AP24N and 2:1 Ca-(II):AP24N polypeptide samples, 293 K. ^{1}H acquisition parameters: 90° pulse = 5.1 μs, spectral window =10 ppm; ^{15}N acquisition parameters: 90° pulse = 36.9 μs, spectral window = 40 ppm. Other acquisition parameters included a recovery delay of 1.5 s, 64 transients per experiment, with a total of 128 experiments with 1024 and 256 data points acquired in t_2 and t_1. For the HSQC coherence transfer period, a $^{1}J_{NH}$ delay of 2.3 ms was utilized and ^{15}N GARP decoupling was applied during the acquisition period. A mixing time of 200 ms was utilized for the NOESY cross-relaxation transfer period. Gradients parameters: $G_1 = 5$ G/cm, $G_2 = 12$ G/cm, $G_3 = 22.4$ G/cm, using sinebell gradient shape and a duration of 1 ms with 500 μs for each gradient pulse. The hypercomplex States-TPPI phase-sensitive method was applied for processing the 2D spectra, with zero filling in F1 and F2 dimension. ^{1}H NMR chemical shift are referenced from internal d_4-TSP; ^{15}N NMR chemical shifts are referenced from external $^{15}NH_4Cl$, with the center of the ^{15}N frequency axis set to 116.7 ppm.

cluster (i.e., the -SSGLC-, -YN-, -NVTTRPNN-, and -PKMF sequence regions, Figures 2, 3), and this phenomenon exceeds what was initially observed within AP7N in the presence of Ca(II).[19]

Detection of a Calcite-Adsorbed AP24N Species. In earlier studies,[17,21] we observed that AP24N modifies the morphology of calcite dislocation hillocks, which suggests that AP24N interacts with exposed calcite surfaces. To monitor the interactions of AP24N with calcium carbonate crystals that occur during the calcite overgrowth assay period, we analyzed geological calcite fragments that were "rescued" from negative control assays and from assays containing AP24N (Figure 5) using ^{1}H MAS NMR.[19,33] First, the differentially washed negative control calcite sample (spectrum E) reveals the absence of methyl and methylene resonances in the 4.0−0.0 ppm frequency range. By comparison, calcite crystals that are "rescued" from assays containing AP24N possess broad proton resonances in this frequency range (spectra D, C), indicating the presence of an absorbed species. Specifically, the 3.2- and 2.2-ppm peaks correlate with the solution-state AP24N Asp CH_β and Q14/V19/M29 CH_β resonances, respectively (spectrum A, also see Supporting Information). What is intriguing about these peaks is their dependency on the washing technique: we observe either two broad resonances (3.2 ppm, 2.2 ppm, spectrum D) after washing with 10 mM $CaCl_2$ solution or a single, narrow resonance (3.2 ppm, spectrum C) after washing with 10 mM $CaCl_2$ followed by 0.5 M $CaCl_2$ solutions. Finally, the lyophilized form of AP24N (spectrum B) possesses three resolvable resonances centered at 3.9, 1.43, and 0.9 ppm, which correspond to the Gly α-CH/Ser β-CH_2, the Thr γ-CH_3/Arg δ-CH_2/Leu γ-CH_2, and the Leu protons of solution-state AP24N, respectively (spectrum A, also Supporting Information). Unusually, the methylene resonances within the 3.5−2.0 ppm frequency range are not observed in the lyophilized AP24N sample, and

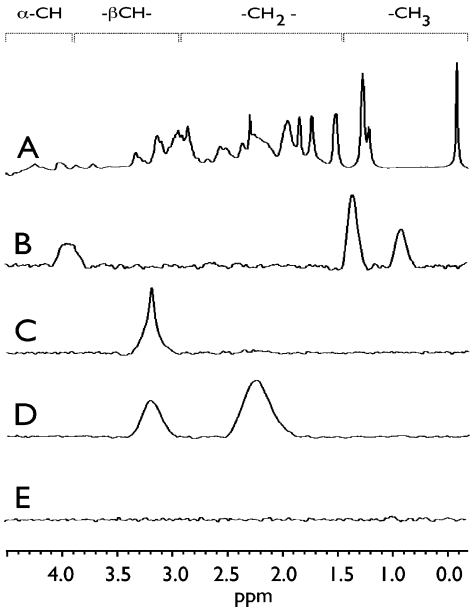

Figure 5. 600 MHz MAS ^{1}H NMR identification (aliphatic region) of AP24N adsorbed species on rescued calcite assay fragments. (A) 1-D solution spectra of 500 μM AP24N polypeptide, pH 7.5; (B) solid-state MAS spectra of 2.5 mg hydrated, lyophilized AP24N sample; (C) solid-state MAS spectra of calcite fragments rescued from AP24N containing assays that have been washed with 10 mM $CaCl_2/D_2O$ and 0.5 M $CaCl_2/D_2O$ prior to data acquisition; (D) solid-state MAS spectra of calcite fragments rescued from AP24N containing assays that have been washed with only 10 mM $CaCl_2/D_2O$ prior to data acqusition; (E) MAS spectra of calcite assay fragments rescued from negative control assays and washed with 10 mM $CaCl_2/D_2O$ and 0.5 M $CaCl_2/D_2O$. ^{1}H NMR acquisition and processing parameters for all solid-state spectra were identical; rotor spinning rate = 10 kHz, spectral window = 12 ppm, data size = 2K with zero filling to 4K points, number of transients = 2048, recovery delay = 1.5 s, 90° pulse = 2.8 μs. Spectra were processed using an apodization function of 10 Hz with exponential multiplication filtering. For the solution-state spectra, 90° = 5.1 μs, spectral width = 10 ppm, data size = 2K, with zero-filling to 4K, recovery delay = 1 s, number of transients = 256, and a 3-9-19 Watergate z-PFG scheme was used for water suppression (see legend to Figure 4 for gradient parameters). Proton chemical shifts were referenced from d_4-TSP.

these resonances appear to be broadened and are not resolvable at the MAS spinning rate or temperature utilized in this study. At this time, we do not understand why differences in Ca(II) washing procedures lead to observed differences in polypeptide-specific ^{1}H NMR chemical shifts and peak line width and line shape. Nonetheless, the persistence of calcite-adsorbed AP24N indicates that some fraction of adsorbed AP24N cannot be released from the calcite fragments via electrostatic neutralization or competitive Ca(II) displacement (Figure 5). We attribute this result to either a population of strongly adsorbed AP24N molecules or to some population of AP24N molecules which become occluded within the calcium carbonate overgrowth formed during the assay period and thus are not accessible for displacement. These findings are similar to those obtained in our earlier MAS NMR studies with AP7N.[19]

Structure Refinement of AP24N and AP7N. With the acquisition of digital 600 MHz ^{1}H NMR data for AP24N for this present study, we realized a significant improvement in cross-peak overlap and identification compared to our earlier 500 MHz studies.[18] Thus, we extended our 600 MHz ^{1}H NMR studies to apo AP7N as well (Supporting Information) and employed H−H distance restraints obtained from these ^{1}H NMR NOESY experiments as input for first-time structure refinement simulations of AP7N and AP24N. As shown in Figure 6, we

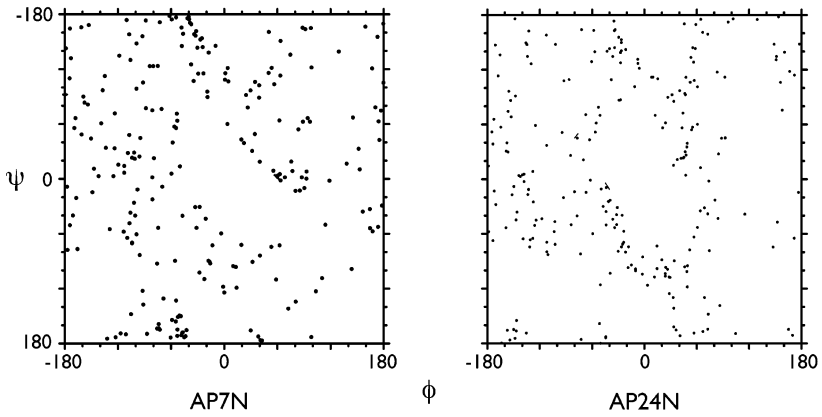

Figure 6. Ramachandran φ, ψ dihedral angle scatterplots for AP24N and AP7N conformer libraries. Scatterplots represent dihedral distributions for XPLOR SA/MD generated conformer ensemble ($n = 10$ for each sequence).

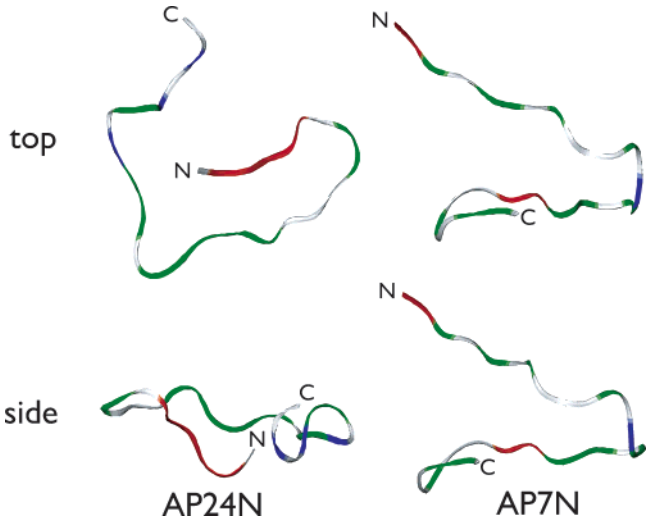

Figure 7. SA/MD XPLOR lowest energy backbone, only conformations of AP7N and AP24N. Ribbon representations are shown from top and side perspectives, with N and C termini labeled. Red = anionic, blue = cationic, green = hydrogen bonding donor−acceptor, white = nonpolar sequence blocks.

obtained plausible lowest energy conformational ensembles ($n = 10$ for each peptide) that match the NMR data sets. As expected, the conformer library for each 30 AA sequence is conformationally heterogeneous and does not possess strong α helix or beta-strand characteristics as evidenced by the broad distribution of φ, ψ backbone torsion angles obtained for both lowest energy structures (Figure 6). This result is not unexpected, given the reported structural lability of these polypeptides;[18,19] hence, both polypeptides are probably very similar in conformation on a qualitative level.

However, there is a structural tendency that distinguishes AP24N from AP7N. The lowest energy backbone conformation obtained for AP24N appears to be flat or planar in profile (side view, Figure 7). Planar or "flat" surfaces are also featured in some of the three-dimensional structures obtained for ice-binding antifreeze proteins[37−39] and are a requirement for alcohol monolayer headgroups that successfully nucleate ice.[40] In contrast, the lowest energy AP7N polypeptide chain has a tendency to adopt a bent paper clip configuration, consisting of a smaller planar region that is offset by the presence of an out-of-plane N-terminal sequence region (side view, Figure 7). Thus, although both sequences appear to be conformationally similar on a qualitatively level (Figure 6), the potential to form a planar backbone configuration, as represented by the lowest energy structure, appears to be higher for AP24N than it is for AP7N (Figure 7).

Both AP7N and AP24N polypeptides are suspected of facilitating displacement of waters of hydration at crystal surfaces or perturbing water around ionic clusters in solution.[21] Although the calculated solvent-accessible areas of the lowest energy structures for AP7N and AP24N are similar (AP24N, 3347.5 Å^2 versus AP7N, 3275 Å^2, a 2% difference), we felt that the observed differences in structural planarity and interactive sequence block length might have a greater impact on the ability of either polypeptide to interact with its solvent environment. Hence, we calculated the Poisson−Boltzmann solvation energies for each lowest energy structure and generated solvation electrostatic surfaces for both peptides (Figure 8, full atom representation). An analysis of the solvation-accessible electrostatic molecular surfaces of both domains reveal that both AP24N and AP7N feature similar traits, that is, anionic and cationic pocket regions that might be important for proposed peptide−Ca(II) and CO_3^{-2} interactions, respectively, in solution or on crystal surfaces (Figure 8). However, the major difference noted for each sequence is in the solvation energy value: −2873 kJ/mol for AP7N versus −3466 kJ/mol for AP24N (i.e., a 17% difference). We believe that the presence of longer anionic and polar sequence blocks in AP24N, along with the tendency to adopt a more planar configuration, promotes more favorable interactions with the solvent grid compared to AP7N (Figure 8). This, in turn, would lead to a lower value in calculated solvation energy.

Discussion

The ability of polypeptides to modulate crystal growth and nucleation is perhaps one of the most poorly understood areas of biology and material science. Clearly, nature has evolved strategies wherein sequence selection, and structural features, are chosen to elicit specific effects on crystal growth. In this current study, under in vitro laboratory conditions, we have determined that AP24N can bind multivalent metal ions (Figure 1) and that the primary site for metal ion binding is within the pentaanionic DDDED cluster (involving D3 and to a lesser extent, D6) (Figure 4). Moreover, this polypeptide experiences Ca(II)-dependent perturbation in backbone motion and dynamics (Figures 2, 3) and adsorbs onto calcite crystal surfaces and cannot easily be displaced from these surfaces via differential washing treatments (Figure 5). Although we lack sufficient details regarding the AP polypeptide−mineral adsorption process, it is probable that the inability to fully displace AP24N

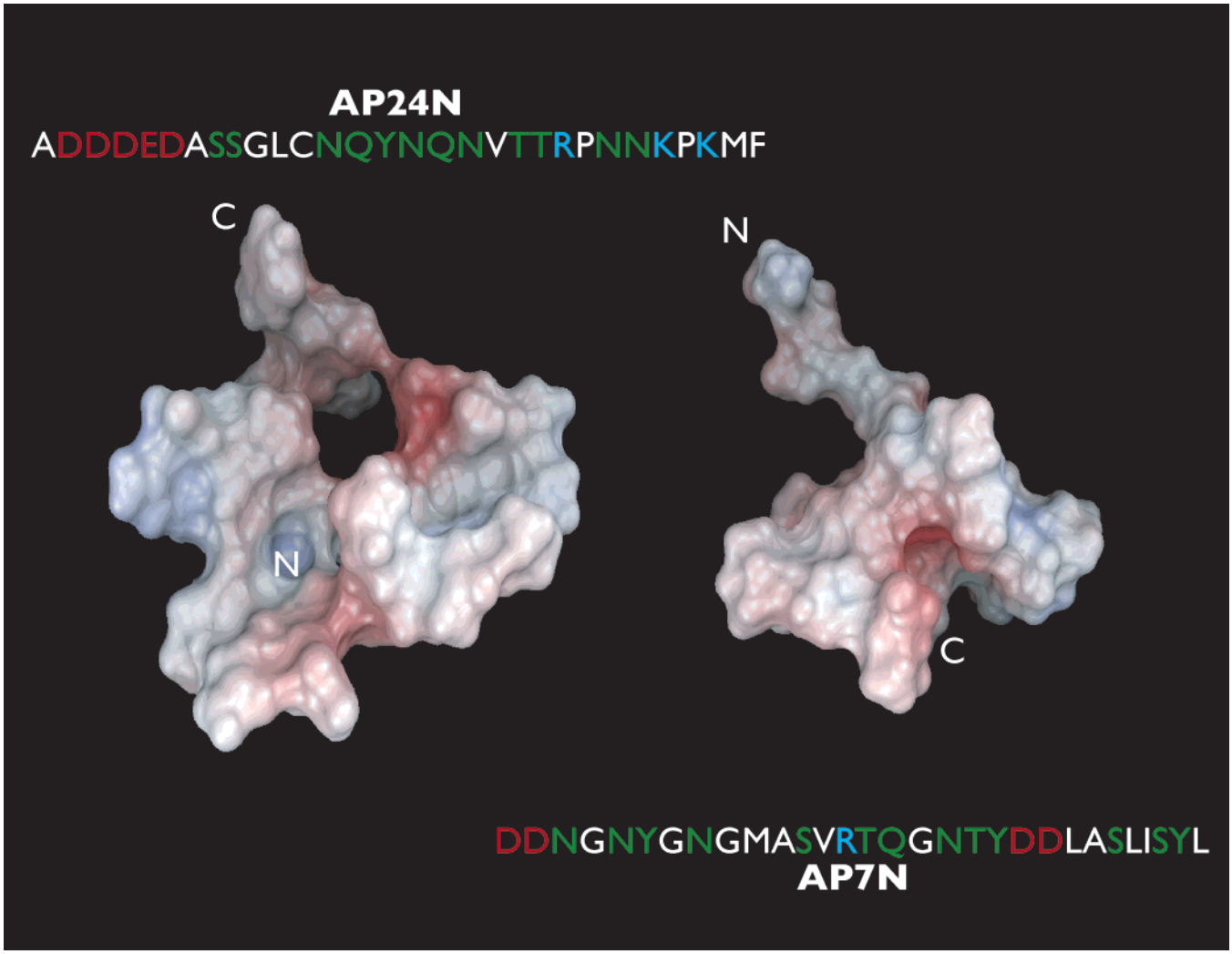

Figure 8. Calculated APBS solvent electrostatic molecular surface of AP7N and AP24N lowest energy conformers (full atom representation). Termini are labeled as N or C. Electrostatic surface rendering is as follows: Red = anionic; blue = cationic. Both peptides are shown in the same orientation as per the top perspectives in Figure 7. Primary sequences are presented next to each image; important sequence blocks are denoted by color coding (red = anionic, blue = cationic, green = hydrogen bonding donor−acceptor).

from mineral surfaces is due to either (a) strong polypeptide−mineral interactions that are not easily overcome by the presence of excess Ca(II) or (b) the presence of an occluded AP24N species within the mineral phase itself. Clearly, additional studies will be required to determine the correct location of AP24N on growing calcium carbonate crystals.

AP7N and AP24N represent two nacre-specific polypeptides where an interesting dichotomy exists: Although dissimilar in primary sequence, both polypeptide affect in vitro calcium carbonate crystal morphology and nucleation in very similar ways but with differing kinetics.[17,21] We have identified several shared and unique molecular features with regard to AP7N and AP24N that fall into two categories, one dealing with crystal morphology and nucleation and the other with crystal growth kinetics. We propose that the following features (not necessarily listed in order of importance) are responsible for the similarities in AP7N, AP24N-mediated crystal morphology, and nucleation.

(a) Unfolded Conformation. AP24N adopts an unfolded structure that is qualitatively similar to AP7N (Figure 6), and this unfolded form permits adaptation with exposed mineral interfaces or promotes ion clustering.

(b) Presence of Interactive Sequence Clusters. Both domains possess qualitatively similar types of sequence blocks that range from 2 to 6 AA in length. The first type, Asp, Glu, or anionic, is considered to be putative site for ion sequestration and for interaction with exposed Ca(II) ions at crystal surfaces (Figure 8). The second type, hydrogen-bonding donor/acceptor (i.e., Arg, Lys, Ser, Tyr, Gln, and Asn), is considered to be putative site for displacement of waters of hydration or for orientation of carbonate ions, either of which would change the activation barrier for incorporation during crystal growth and promote growth step acceleration (Figure 8).[21]

(c) Presence of Surface-Accessible Pocket Regions. As shown in Figure 8, the electrostatic surfaces of both AP7N and AP24N possess accessible anionic (red) and cationic (blue) pockets. These regions could act as putative sites for ion cluster or water sequestration during the mineralization process.

Collectively, the presence of similar unfolded conformation, two types of interactive sequence blocks, and surface-accessible pocket regions in each polypeptide would ensure similar in vitro effects on crystal morphology, that is, acute step pinning, acceleration of obtuse step velocities, and nucleation deposition.[21]

Having identified common molecular traits within AP7N and AP24N domains, there are molecular features that distinguish AP24N from AP7N and, probably, that dictate the observed differences in polypeptide-mediated mineralization kinetics.

These are as follows (again, not necessarily listed in order of importance).

(**a) Tendency to Adopt a Planar Polypeptide Backbone Configuration.** It has been shown in antifreeze proteins and with alcohol monolayer headgroups that planar molecular surfaces facilitate interaction with water clusters.[37−40] Although AP7N and AP24N are conformationally similar to one another (Figure 6), our structure refinement data indicates that both AP7N and AP24N can, at some point, structurally deviate from one another and adopt different degrees of planar configuration in their backbone structures (Figures 7, 8). Although we do not know how often the lowest energy structures are adopted by either sequence in solution, the tendency to form a more planar configuration, even if transient, would convey AP24N with improved capabilities over AP7N with regard to adaptation to mineral surfaces or to promote ion/water sequestration over time.

(b) Critical Length and Location of Anionic Sequence Clusters. Here, the major difference between AP7N and AP24N lies in the length and location of anionic sequence clusters. AP7N possesses a pair of -DD- diads at sequence positions 1,2, and 21,22, whereas AP24N possesses a larger pentaanionic DDDED cluster located at the N-terminus.[17] It is likely that the concentration of five anionic groups together (AP24N), as opposed to the dispersal of four anionic groups in pairs (AP7N), provides AP24N with higher anionic density, which, in turn, would improve Ca(II) ion sequestration or polypeptide adsorption at Ca(II) sites on crystal surfaces.

(c) Critical Length and Location of Hydrogen-Bonding Sequence Clusters. AP24N possesses longer sequence blocks of hydrogen bonding donor/acceptor residues (2, 3 AA in length) than AP7N (1, 2, 3 AA in length) and also features a unique 6-AA -NQYNQN- block (Figure 8). This longer sequence block may facilitate peptide−water or peptide−carbonate ion interactions to a greater degree than the corresponding shorter sequence clusters found in AP7N.

(d) Local Conformational Response to Ca(II). As shown in Figures 2 and 3, Ca(II) ion interactions within the AP24N sequence lead to extensive motional, exchange, or distance-related local changes at other nonanionic regions upstream of the DDDED anionic cluster (i.e., involving the -SSGLC-, -YN-, -NVTTRPNN-, and -PKMF sequence regions). This Ca(II) responsiveness exceeds what we originally observed for AP7N[19] and is probably attributable to the unique planar configuration of AP24N (Figure 8) which features the DDDED anionic cluster in close proximity to the other sequence regions in AP24N. Hence, Ca(II) interactions within the DDDED anionic cluster affect noncontiguous sequence regions to a greater degree within AP24N than in AP7N, particularly since AP7N features out-of-plane positioning of the -DD- diads within the bent paper clip structure (Figures 7, 8). Collectively, the length, linear arrangement, and three-dimensional positioning of interactive sequence blocks are the features which distinguish AP24N, and we believe that these features convey enhanced mineralization kinetics upon this sequence. Obviously, additional studies will be required to fully determine the specific sequence and structural features of AP7N and AP24N which are responsible for morphology and kinetics.

With the identification of structural features that correlate with polypeptide-mediated mineralization behavior and kinetics, we believe that the concept of polypeptide "tuning" of calcium carbonate crystal growth[21] could reach a more sophisticated level in the laboratory setting. We propose that it may be possible to create artificial biomineralization proteins that possess the requisite features for crystal growth morphology selection, namely, introduce unfolded conformations, Ca(II), carbonate, and water-binding sites, and tune the kinetics of crystal growth of these proteins via further manipulation of interactive sequence block length, position, and composition. In addition, kinetic control could also be tuned via sequence engineering of given percentages of planar polypeptide backbone configuration at interactive sequence blocks. These concepts should be testable via the use of amino acid substitutions within AP7N and AP24N. These types of experiments are currently in progress.

Finally, although we do not know what specific tasks are appointed to the AP7N and AP24N sequences within the realm of the mollusk shell nacre layer, we do know that the parent proteins AP7 and AP24 both limit the growth of calcite, presumably to allow other proteins to promote the growth of the metastable polymorph, aragonite.[17] Moreover, it is also known that both AP7 and AP24 proteins are purified from the nacre layer as a noncovalent complex.[17] It is intriguing that two proteins which coexist within the same layer of the mollusk shell would possess N-terminal sequences that are capable of performing similar tasks but with different kinetics. This suggests that the inhibition of calcite crystal growth within the nacre layer may require the participation of an AP7−AP24 protein complex, with each protein contributing similar effects on morphology but different kinetic effects on crystal growth. Although the necessity of this unusual scenario is not understood at present, it does point to the potential complexity that exists within the aragonite polymorph selection process of the mollusk shell, and it suggests that biological control over crystal growth is tuned by a specific combination of proteins and their sequence domains to a degree that we have not yet fully appreciated.

Acknowledgment. This work was supported by funding from the Department of Energy (DE-FG02-03ER46099, to J.S.E.). The INOVA 600 NMR spectrometer was supported by a Defense University Research Instrumentation Program award from the Army Research Office (W911NF-04-1-0214), and this paper represents contribution number 33 from the Laboratory for Chemical Physics, New York University.

Supporting Information Available. 600 MHz ^{1}H chemical shifts obtained for apo-AP7N and apo-AP24N C-amide polypeptides and 600 MHz ^{1}H 200-ms mixing time PFG-NOESY NH−CH_α and NH−NH fingerprint spectra of apo-AP7N at 293 K. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Weiner, S.; Addadi, L. *J. Mater. Chem.* **1997**, *7*, 689−702.
(2) Almqvist, N.; Thomson, N. H.; Smith, B. L.; Stucky, G. D.; Morse, D. E.; Hansma, P. K. *Mater. Sci. Eng.* **1999**, *C7*, 37−43.
(3) Aizenberg, J.; Tkachenko, A.; Weiner, S.; Addadi, L.; Hendler, G. *Nature* **2001**, *412*, 819−822.
(4) Bazylinski, D. A. *Chem. Geol.* **1996**, *132*, 191−198.
(5) *Biomineralization: Chemical and Biochemical Perspectives*; Mann, S., Webb, J., Williams, R. J. P., Eds.; VCH: Weinheim, Germany, 1989.
(6) Lowenstam, H. A.; Weiner, S. *On Biomineralization;* Oxford Press: New York, 1989.
(7) Litvin, A. L.; Valiyaveettil, S.; Kaplan, D. L.; Mann, S. *Adv. Mater.* **1997**, *9*, 124−127.
(8) D'Souza, S. M.; Alexander, C.; Carr, S. W.; Waller, A. M.; Whitcombe, M. J.; Vulfson, E. N. *Nature* **1999**, *398*, 312−316.
(9) Lakshminarayanan, R.; Valiyaveettil, S.; Loy, G. L. *Cryst. Growth Des.* **2003**, *3*, 953−958.
(10) Kim, I. W.; Robertson, R. E.; Zand, R. *Adv. Mater.* **2003**, *15*, 709−712.
(11) Belcher, A. M.; Wu, X. H.; Christensen, R. J.; Hansma, P. K.; Stucky, G. D.; Morse, D. E. *Nature* **1996**, *38*, 56−58.

(12) Thompson, J. B.; Paloczi, G. T.; Kindt, J. H.; Michenfelder, M.; Smith, B. L.; Stucky, G.; Morse, D. E.; Hansma, P. K. *Biophys. J.* **2000**, *79*, 3307−3312.
(13) Pontoni, D.; Bolze, J.; Dingenouts, N.; Narayanan, T.; Ballauff, M. J. *Phys. Chem. B* **2003**, *107*, 5123−5125.
(14) Aizenberg, J.; Lambert, G.; Weiner, S.; Addadi, L. *J. Am. Chem. Soc.* **2002**, *124*, 32−39.
(15) Levi, Y.; Albeck, S.; Brack, A.; Weiner, S.; Addadi, L. *Chem. Eur. J.* **1998**, *4*, 389−396.
(16) Samata, T.; Hayashi, N.; Kono, M.; Hasegawa, K.; Horita, C.; Akera, S. *FEBS Lett.* **1999**, *462*, 225−229.
(17) Michenfelder, M.; Fu, G.; Weaver, J. C.; Wustman, B. A.; Taranto, L.; Evans, J. S.; Morse, D. E. *Biopolymers* **2003**, *70*, 522−533; errata **2004**, *73*, 291.
(18) Wustman, B. A.; Morse, D. E.; Evans, J. S. *Biopolymers* **2004**, *74*, 363−376.
(19) Kim, I. W.; Morse, D. E.; Evans, J. S. *Langmuir* **2004**, *20*, 11664−11673.
(20) Kim, I. W.; Collino, S.; Morse, D. E.; Evans, J. S. *Cryst. Growth Des.* **2006**, *6*, 1078−1082.
(21) Kim, I. W.; Darragh, M. R.; Orme, C.; Evans, J. S. *Cryst. Growth Des.* **2006**, *5*, 5−10.
(22) Goddard, T. D.; Kneller, D. G. *SPARKY 3*, version 3.110; University of California, San Francisco: San Francisco, CA, 2004.
(23) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(24) Baker, N. H.; Sept, D.; Joseph, S.; Holst, N. J.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037−10041.
(25) Shields, S. J.; Bluhm, B. K.; Russell, D. H. *Int. J. Mass Spectrom.* **1999**, *182*/*183*, 185−191.
(26) Wong, K. C.; Chan, T. D. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 513−516.
(27) Prigodich, R. V.; O'Connor, T.; Coleman, J. E. *Biochemistry* **1985**, *24*, 6291−6298.
(28) Wojcik, J.; Goral, J.; Pawlowski, K.; Bierzynski, A. *Biochemistry* **1997**, *36*, 680−687.
(29) MacManus, J. P.; Hogue, C. W.; Marsden, B. J.; Sikorska, M.; Szabo, A. *J. Biol. Chem.* **1990**, *265*, 10358−10366.
(30) Collino, S.; Kim, I. W.; Evans, J. S. *Cryst. Growth Des.* **2006**, *6*, 839−842.
(31) Kim, I. W.; DiMasi, E.; Evans, J. S. *Cryst. Growth Des.* **2004**, *4*, 1113−1118.
(32) Evans, J. S. *Curr. Opin. Colloid Interface Sci.* **2003**, *8*, 48−54.
(33) Wustman, B. A.; Weaver, J. C.; Morse, D. E.; Evans, J. S. *Langmuir* **2003**, *19*, 9373−9381; errata **2004**, *20*, 277.
(34) Fiebig, K. M.; Schwalbe, H.; Buck, M.; Smith, L. J.; Dobson, C. M. *J. Phys. Chem.* **1996**, *100*, 2661−2666.
(35) Logan, T. M.; Theriault, Y.; Fesik, S. W. *J. Mol. Biol.* **1994**, *236*, 637−648.
(36) Smith, L. J.; Bolin, K. A.; Schwalbe, H.; MacArthur, M. W.; Thornton, J. M.; Dobson, C. M. *J. Mol. Biol.* **1996**, *255*, 494−506.
(37) Cheng, Y.; Yang, Z.; Tan, H.; Liu, R.; Chen, G.; Jia, Z. *Biophys. J.* **2002**, *83*, 2202−2210.
(38) Liou, Y. C.; Tocilj, A.; Davies, P. L.; Jia, Z. *Nature* **2000**, *406*, 322−324.
(39) Yang, D. S.; Hon, W. C.; Bubanko, S.; Xue, Y.; Seetharaman, J.; Hew, C. L.; Sicheri, F. *Biophys. J.* **1998**, *74*, 2142−2151.
(40) Dai, Y.; Evans, J. S. *J. Phys. Chem. B* **2001**, *105*, 10831−10837.

BM0700183