Identification and Structural Characterization of an Unusual RING-Like Sequence within an Extracellular Biomineralization Protein, AP7[†,‡]

Sebastiano Collino, Il Won Kim, and John Spencer Evans*

Laboratory for Chemical Physics, Center for Biomolecular Materials Spectroscopy, New York University, 345 E. 24th Street, Room 1007, New York, New York 10010

Received September 26, 2007; Revised Manuscript Received December 16, 2007

ABSTRACT: The RING or Really Interesting New Gene represents a family of eukaryotic sequences that bind Zn (II) ions and participate in intracellular processes involving protein−protein interaction. Although found in over 400 different proteins, very little is known regarding the structure–function properties of these domains because of the aggregation problems associated with RING sequences. To augment this data set, we report an unusual 36 AA C-terminal sequence of an extracellular matrix mollusk shell protein, AP7, that exhibits partial homology to the RING family. This Cys, His-containing sequence, termed AP7C, binds Zn (II) and other multivalent ions, but does not utilize a tetracoordinate complexation scheme for binding such as that found in Zn (II) finger polypeptides. Moreover, unlike Zn (II) finger and RING domains, this 36 AA can fold into a relatively stable structure in the absence of Zn (II). This folded structure consists of three short helical segments (A, B, and C), with segments A and B separated by a 4 AA type I β-turn region and segments B and C separated by a 7 AA loop-like region. Interestingly, the putative RING-like region, -RRPFHECALCYSI-, experiences slow conformational exchange between two structural states in solution, most likely in response to imido ring interconversion at P8 and P21. Poisson–Boltzmann solvation calculations reveal that the AP7C molecular surface possesses a cationic region near its N-terminus, which lies adjacent to the 30 AA mineral modification domain in the AP7 protein. Given that the AP7C sequence does not influence mineralization, it is probable that this cationic pseudo-RING region is utilized by the AP7 protein for other tasks such as protein−protein interaction within the mollusk shell matrix.

The formation of biominerals in nature often requires the participation of specialized proteins that direct matrix assembly, nucleation, and crystal growth (*1, 2*). In many cases, the sequences of these specialized proteins are highly unique and often feature regions that do not correspond to the sequences of other known globular proteins (*3–7*) and with good reason. The process of nucleation and crystal growth control require unique sequences that are not common to other proteins, such as repetitive and nonrepetitive clusters of anionic, cationic, and hydrogen bonding donor/acceptor amino acids that play important roles in ionic cluster formation, hydration, and mineral surface interaction processes (*8–17*). Although we do not fully appreciate how biomineralization protein sequences function within the context of inorganic solid phase synthesis, we are beginning to understand that these proteins are multifunctional and capable of executing simultaneous tasks such as deposit formation, crystal growth acceleration and deceleration, and occlusion within the mineral phase itself (*6, 7, 12*).

The multifunctional capabilities of biomineralization proteins actually may run deeper than we suspect. There have been observations of imperfect homology between certain regions of biomineralization proteins and nonmineral globular proteins (*4, 5, 7*). Some recent examples can be found in mollusk shell mineralization proteins, where regions homologous to acetylcholine-binding domains (*5*) glycine loops (*18*), and disulfide core domains (*18*) have been identified. There is speculation that these globular-like domains have functions that are not directly related to nucleation or crystal growth processes (*5, 18*). As an example, in our recent studies with AP7[1], a 66 AA, 7.5 kDa protein of the nacreous layer of the Pacific Red abalone, *Haliotis rufescens* (*4, 14*), we identified that the 30 AA N-terminal sequence (AP7N) acted as a calcite blocker in vitro but that the remaining 36 AA C-terminal region of this protein (AP7C) had no observable effect on the mineralization process in vitro (*14*). Intriguingly, this 36 AA C-terminal region contains Zn (II)-coordinating residues Cys and His (Figure 1) (*4*). Recent sequence comparisons reveal that small regions of the AP7C domain

[†] This work was supported by funding from the Department of Energy (DE-FG02-03ER46099). The INOVA 600 NMR spectrometer was supported by a Defense University Research Instrumentation Program award from the Army Research Office (W911NF-04-1-0214).

[‡] The coordinates of the lowest energy structure (2jyp) have been deposited in the Protein Database (PDB).

* To whom correspondence should be addressed. Tel: 2129989605. Fax: 2129954087. E-mail: jse1@nyu.edu.

[1] Abbreviations: BLAST, basic local alignment search tool; RNF32, ring finger protein 32; RKR1, nuclear ring domain protein-1, TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser spectroscopy; PFG, pulsed field gradient; AP7, aragonite protein 7; AP7C, 36 AA C-terminal domain, AP7; AP7*N*, 30 AA N-terminal domain, AP7; AP24, aragonite protein 24; RING, really interesting new gene; DTT, dithiothreitol; HHARI, human homologue ariadne protein.

10.1021/bi701949p CCC: $40.75
Published on Web 02/26/2008

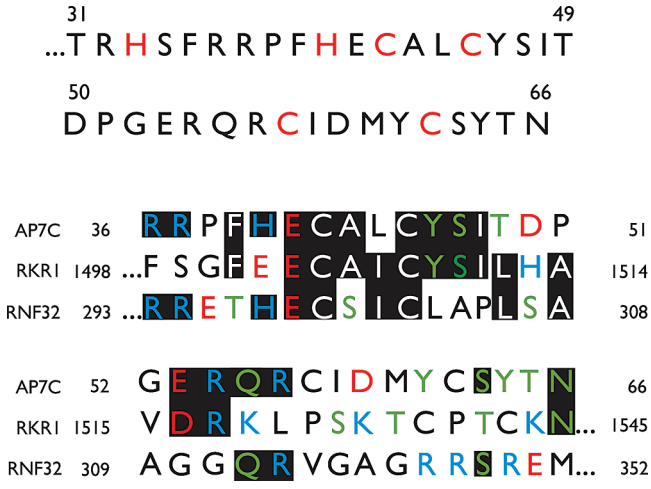

FIGURE 1: Top: Primary sequence of AP7 protein, C-terminal region (AP7C fragment). Sequence numbering corresponds to the original AP7 sequence (*4*), and red color indicates the location of His and Cys putative Zn (II) binding residues. Note that in this article, we use the sequence numbering of the AP7C fragment as 1 through 36. Bottom: Partial sequence homology alignment between the AP7 C-terminal domain sequence and two known RING motifs found in Human RNF32 (*(19a)*) and Yeast YMR247C (*(19b)*). The Basic Local Alignment Search Tool (BLAST) was utilized for protein–protein search inquiries and included the Swiss Protein Sequence Database (Swissprot) with BLASTP (Protein–protein BLAST) algorithm. Color legend: Red = anionic residues; blue = cationic residues; green = hydrogen-bonding donor/acceptor residues. Black bars indicate where homologies exist.

exhibit partial homology to two intracellular proteins, human RNF32 (*(19a)*) and yeast RKR1 (*(19b)*), both of which possess an interesting subset of carboxy-terminal Cys-rich Zn-binding motifs known as RING (really interesting new gene domain, Figure 1) (*19–23*). Reported variants of the RING domain family occur in over 400 proteins and bind either one (*21, 23*) or two (*20, 22*) Zn (II) ions per motif (*20, 21*); in general, these domains are involved in mediating protein−protein interactions within eukaryotic cells (*20–23*). Interestingly, the AP7 protein has been reported to copurify from abalone nacre with another nacre-specific protein, AP24 (*4*), which suggests that protein−protein recognition and docking may be another attribute of the AP7 protein.

Because of their inherent propensity to aggregate (*20–23*), very little is known regarding the RING domains. Given the general lack of information on RING domains and the solubility problems associated with 66 AA AP7, we elected to pursue the characterization of the 36 AA AP7C fragment, similar to the approach utilized in other RING domain studies (*20–23*). We felt that the fragment characterization approach was valid because of the following observation: although the AP7N fragment adopts an unfolded structure, both the AP7 protein and the AP7C fragment adopt an α-helical structure in solution in the absence of Zn (II) (*14*), thus demonstrating that the α-helical conformation of AP7 arises from the AP7C sequence. Therefore, we chemically synthesized the 36 AA AP7C polypeptide, and using UV–vis spectroscopy, mass spectometry, NMR, and CD spectrometry, we investigated the structure and Zn (II) binding capabilities of the apo form of AP7C under reducing conditions. We find that the apo-AP7C fragment folds without the assistance of Zn (II) to form a global conformation consisting of three short helical segments, A, B, and C, that are interconnected by nonhelical regions. The AP7C sequence binds Zn (II) in either 1:1 or 1:2 peptide:Zn (II) complexes but does not appear to utilize a tetracoordinate binding scheme (*20–29*). Interestingly, the three-dimensional surface of AP7C features electrostatic regions that may play a role in docking with other nacre proteins such as AP24 (*4*).

MATERIALS AND METHODS

Polypeptide Synthesis and Purification. N^{α}-acetyl capped, free C-terminal 36-AA AP7C polypeptide was synthesized in the free Cys-SH form and purified at the 100 μmol level at the Wm. Keck Biotechnology Peptide Synthesis Facility, Yale University, by Dr. Janet Crawford, using protocols described in our earlier work (*14*). After resin cleavage and reverse-phase HPLC purification (Waters C-18 column, >95% pure) (*14*), the experimental molecular mass of AP7C was determined by matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) to be 4398.7 Da, in agreement with the theoretical value of 4399.9 Da (*14*).

Sample preparation of AP7C for subsequent experiments involved the dissolving of lypohilized polypeptide (final peptide concentration, 450 μM) in H_2/high purity N_2 flushed deionized–distilled water containing 30 equivalents of dithiothreitol (DTT, Sigma-Aldrich) and heating of this sample at 90 °C for 30 min, followed by slow cooling to room temperature (*24–26*). Aliquots were then stored in airtight sealed vials that were preflushed with high purity N_2 gas. These stock solutions were then stored at −20 °C until needed. To ensure reducing conditions for AP7C Cys residues, all solutions utilized in the experiments described below were degassed and purged with high purity N_2/H_2 gas for at least 15 min immediately prior to each experiment. In addition, containers with AP7C samples were flushed with high purity N_2 gas upon opening and before closing (*24–26*). Note that the reduced apo-AP7C polypeptide was found to exhibit excellent solubility in the 8 to 450 μM range in neutral pH, low ionic strength aqueous buffers.

Ion Trap Mass Spectrometry. AP7C stock solution was dissolved in degassed deionized–distilled water and titrated with microliter volumes of Tris-HCl to a final peptide concentration of 10 μM at pH 7.4. To create AP7C:metal ion complexes, appropriate volumes of degassed, N_2-flushed $CaCl_2$, $CdCl_2$, $ZnCl_2$, $LaCl_3$, and $EuCl_3$ stock solutions (each 10 mM, 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 (*8*). Ion trap MS experiments were performed on an Agilent LC/MSD 1100 equipped with an electrospray ionization source using the following procedure: 1 mL 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 a skimmer voltage of −45 V and a capillary exit voltage of −70 V.

UV–Vis Spectroscopy. We utilized Co (II) as an analogue of Zn (II) to probe metal ion binding within the putative RING-like motif of AP7C, using the protocols published earlier for Zn (II) finger polypeptides (*24–26*). Peptide

concentrations varied from 10 to 50 μM in 100 μM Tris-HCl at pH 7.5, and spectra were obtained using a 1-cm path length (quartz cells, Sarna, Inc., NY). Co (II) titrations of AP7C titrations were performed by adding stoichiometric amounts of 1 M $CoCl_2$ stock solutions (99.99% pure, Sigma-Aldrich, in deionized–distilled water) to the apo-AP7C peptide; peptide:Co (II) stoichiometries of 1:5, 1:2, 1:1, 1:2, 1:10, 1:20, 1:100, 1:200, and 1:300 were utilized. Ultraviolet and visible light spectra of Co (II):AP7C complexes were obtained at 20 °C on an Agilent 8453 spectrophotometer, with each spectrum scanned from 800 to 200 nm at a rate of 500 nm/min.

CD Spectrometry. Zn (II) titrations were performed using 8 μM AP7C polypeptide in 100 μM Tris-HCl at pH 7.5 and microliter additions of $ZnCl_2$ stock solution (99.9% pure, Sigma-Aldrich, deionized–distilled water) over a range of 2:1, 1:1, 1:2, 1:4, 1:10, 1:20 AP7C:metal ion ratios. The peptide samples were scanned from 185 to 260 nm at 20 °C with an AVIV 60 CD spectrometer, running 60 DS software version 4.1. The CD spectrometer was previously calibrated with d-10 camphorsulfonic acid. Spectra were obtained as an average of 3 to 5 scans using 1 nm bandwidth and a scan rate of 1 nm/s, with appropriate background buffer subtraction performed (where applicable DTT, Tris-HCl). Final data are reported as mean residue ellipticity [θ_M].

^{1}H NMR Spectroscopy. Homonuclear 2-D ^{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 NOESY (t_{mix}= 50, 100, 150, 200 ms) experiments were used for sequential assignments (*8*). In TOCSY and NOESY experiments, water suppression was achieved using 3-9-19 *z*-axis PFG gradient schemes. NMR experiments were performed under reducing conditions on a 450 μM AP7C sample containing 50 μM dithiothreitol (DTT), 90% deionized–distilled water/10% D_2O at pH 7.5 (adjusted with microliter volumes of either 1 N NaOH or 1 N HCl), with the peptide acting as its own buffer. Three-bond *J*-coupling constants ($^3J_{HNHalpha}$, abbreviated as ^3J) were determined using z-PFG-TOCSY experiments and spectral processing methods that permit coupling constant measurements from TOCSY spectra (*8, 15*). NMR data were visualized using Sparky software (*30*) and processed using VNMRj and NMRPipe software (*31*). NMR acquisition and process parameters are reported in the corresponding figure legends.

Simulated Annealing Molecular Dynamics (SA/MD) Structure Refinement. Initial structure coordinates for AP7C were generated in the Biopolymer module (Biosym Insight II package, Accelrys, San Diego, USA) using the CVFF force field and extended conformation (φ = 180°, ψ = 180°). Partial atomic charges (protonated His+, Arg+, alpha-NH_3^+; deprotonated Asp-, Glu-, and alpha-COO^-) were applied to emulate a neutral pH scenario. NOE values obtained from ^{1}H NMR NOESY (t_{mix} = 50, 100, 150, 200 ms) were used as restraints in setting corresponding H−H distances. 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) categories (*8, 32*). These distances were employed using soft-square potentials with a force constant of 50 kcal/mol-Å (*2*). Inter and intraresidue interproton distances were calculated using a reference distance of 2.47 Å for the delta-CH, epsilon-CH Tyr aryl ring protons (*8, 32*). Pseudoatoms defining the centroids of the methyl and methylene groups were defined in the NOE restraint file. The ^3J coupling constants were used where applicable to calculate φ angles by application of the Karplus equation following the Pardi parameters (*33*). Using the XPLOR software package a total of 150 restraints (129 NOE and 21 φ angles) were used for the SA/MD. Specifically, all simulations employed a distance dependent dielectric (ϵ) = 78.5 (representing implicit solvation), a nonbonding cutoff of 4.5 Å (to determine electrostatic interactions using a switching function) and the Verlet algorithm integrator (1 fs 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 (*8*).

Following initial minimization to convergence (Fletcher−Powell algorithm), 30 ps of equilibrium dynamics were performed at 300 K, and then the system was incrementally ramped to 2000 K to sample conformational space. The system was cooled from 2000 K to 300 K at 100 K per step, and the resulting conformers were then minimized to convergence. A total of 1000 structures were generated in the initial ensemble. About 100 structures were then chosen on the basis of the number of NOE and angle restraint violations. Out of these structures, 10 were selected on the basis of an energy cutoff, thus selecting the lowest energy structure. Lowest energy coordinate files were visualized using the MolMol and PyMOL packages. Solvation energies (vacuum to solvent transfer) and surface plots were created for lowest energy conformers with the Adaptive Poisson–Boltzmann Solver (APBS) solvation calculation module (*34*) 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 78.54, 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 10 mM was employed to mimic the low ionic strength conditions of our NMR samples.

RESULTS

Metal Binding Studies. We initiated our study of the Zn (II) binding properties of AP7C by using the ion trap MS and the metal ion-induced mass shift method (*8, 15*), an approach that has successfully determined the metal binding properties of polypeptides. In order to displace ubiquitous sodium ions and to ensure sufficient peptide:metal ion complex formation, the method requires the use of mole excess of metal ions (i.e., peptide:metal ion = 1:10 or greater) (*8, 15*). For polypeptide MW >1500 Da, the atomic masses of both Zn (II) and Ca (II) are rough multiples of the mass of Na (I) (= 22.99 a.m.u.), and this can complicate adduct identification. To circumvent this problem, we employed large atomic mass metal ion analogues of Zn (II) [i.e., Cd (II) (112.4 a.m.u.)] or Ca (II) [i.e., La (III) (138.9 a.m.u) and Eu (III) (157.9 a.m.u)], such that resolvable, unambiguous mass shifts (≈100 Da) were induced for peptide:metal ion complexes (*8, 15*). Note that La (III) and Eu (III) have been successfully employed in studying Ca

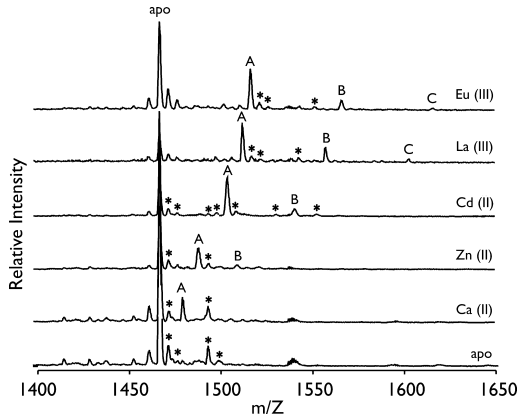

FIGURE 2: Ion trap mass spectra (negative ionization mode) of apo-AP7C and AP7C in the presence of 10-fold molar excess of $CaCl_2$, $ZnCl_2$, $CdCl_2$, $LaCl_3$, and $EuCl_3$. Major AP7C:metal ion adduct species are noted on the spectra. See Table 1 for the observed and theoretical mass/Z values of denoted adducts in each spectra. The * symbol denotes AP7C-Na ion adducts.

(II) interaction sites within a number of Ca (II) binding sequences (*8, 15, 35–37*). As shown in Figure 2 and Table 1, we note the formation of 1:1 and 1:2 AP7C:metal ion adduct species in the presence of Zn (II), Cd (II), and in the presence of La (III) and Eu (III), 1: 3 AP7C:metal ion adducts. Tris buffer penta-complexation with AP7C was also noted. Because of the low atomic mass of Ca (II) and its hypothetical overlap with 2 sodium ions, only the 1:1 AP7C: Ca (II) adduct was confirmed. Nonetheless, our ion trap experiments reveal that AP7C can minimally form 1:1 and 1:2 complexes with a variety of multivalent metal ions, including Zn (II).

Subsequent UV–vis spectrophotometric characterization of Co (II) binding by AP7C reveals that AP7C does not exhibit the coordination characteristics of a Zn (II) finger polypeptide (Figure 3). Typically, aqueous $CoCl_2$ possesses two bands in the visible region at 460 and 510 nm whose intensities are concentration-dependent (Figure 3A). In the presence of Zn (II) finger polypeptides, the binding of Co (II) ion produces the characteristic strong d−d transitions seen between 550 and 725 nm in tetrahedrally coordinated complexes of this metal ion (*24–26*). However, as shown in Figure 3B, the stoichiometric addition of Co (II) to reduced AP7C does not reveal this d− d transition; instead, it retains the same spectral features as aqueous $CoCl_2$ without the appearance of the expected red shift. The only spectral perturbation that is noted is the appearance of a narrow UV ligand–metal charge transfer (LMCT) band at 270 nm, indicating that AP7C is interacting with Co (II) via peptide-specific thiol-SH and/or imidazole N ligand atoms (*24–26*). However, note that the LMCT band for typical Zn (II) finger sequences has been reported to be much broader and exhibits a shoulder region at 310–350 nm (*24–26*), which is not evident in our AP7C UV–vis spectra (Figure 3). Hence, the UV–vis experiments confirm that AP7C forms a complex with Co (II) but not in a manner consistent for tetracoordinate Zn (II) binding. This implies that the ligand coordination number for peptide:Zn (II) interactions must be either mono, bi, or tridendate.

Qualitative Structural Features of AP7C in the Presence and Absence of Zn (II). One of the major Zn (II) finger characteristics is the adoption of random coil or unfolded polypeptide structure in the apo form and the formation of α-helical structure with the introduction of Zn (II) ion (*19, 24–28*). With this folding scenario in mind, we performed a Zn (II) titiration of AP7C and monitored the conformation of this sequence using CD spectrometry (Figure 4). Here, under reducing conditions, we confirm our original finding (*14*) that the apo form of AP7C does not adopt a random coil conformation, as evidenced by the absence of a negative ellipticity near 195–198 nm. Rather, we observe two negative ellipticity bands associated with α-helix: a π-π* transition band at 208 nm and a n-π* transition band at approximately 222 nm (*8, 15, 24–28*). With the stoichiometric addition of Zn (II), these two bands persist and exhibit increased negative intensity changes, particularly for the 1:2 and 1:1 Zn:AP7C samples. This suggests that the helical content of AP7C may be increasing in response to Zn (II) binding within the pseudo-RING domain. Hence, we conclude that at neutral pH both apo- and Zn (II)-bound AP7C exist in a helical conformation.

Evidence of Conformational Exchange in Apo-AP7C. Following up from our CD investigations, we initiated solution-state NMR experiments at neutral pH to quantitate the sequence-specific structure of AP7C at neutral pH. Because of Zn (II)-induced aggregation problems that occur at higher polypeptide concentrations (i.e., 400–500 μM) for NMR spectroscopy, we were forced to limit our experiments to the study of the apo-polypeptide. Surprisingly, during this process, we uncovered the presence of conformational exchange within apo-AP7C. This phenomena is clearly seen in the PFG-TOCSY (Figure 5) and NOESY (Figure 6) amide fingerprint spectra, where we observe duplicate cross-peaks for H3, R6, Y16, and S17 that represent two nonequivalent populations (one denoted by *) of alpha-N(i, i) and/or beta-N(i, i) intraresidue protons. These nonequivalent populations reflect slow exchange between two distinct conformational states for H3, R6, Y16, and S17. This phenomenon mirrors the situation that we observed in a partially helical carbon nanohorn-binding 12 AA sequence, where conformational cis–trans interconversion of a centrally located Pro residue introduced unique conformational backbone motion in this peptide (*38*). This interconversion led to the appearance of two-state exchange cross-peaks in the corresponding TOCSY and NOESY spectra for specific residues that resided downstream of the Pro residue (*38*).

It is known that Pro residues can function as a molecular hinge within protein sequence segments and influence the dynamics, conformation, and/or motion of these segments in response to intra or intermolecular interactions (*39, 40*). Moreover, it is known that Pro cis–trans interconversion does occur in small polypeptides (*38–40*). Because of ^{1}H NMR chemical shift overlap in the frequency region of the imido α, β, γ, and δ NOE cross-peaks and α, β, γ, and δ methine/methylene protons, we were unable to detect discernible differences in the cis−trans ratios for P8 or P21. Nonetheless, given that the H3, R6, and Y16, S17 residues of the pseudo-RING motif lie downstream of P8 and P21, respectively (Figure 1), it is likely that P8, P21- specific cis–trans intercoversion or other imido ring dynamics are responsible for the observed slow conformational exchange within the RING-like motif region.

Solution Structure of the AP7C Fragment. Qualitative NMR parameters, such as proton conformational shifts (*41*)

Table 1: AP7C:Metal Ion Adduct Species

sample	adduct	adduct species	observed MW (Da)	theoretical MW (Da)
Apo AP7C	$[M]^{-3}$	$[M^{3-} + 5(\text{Tris-HCl})]^{3-}$	1466.2	1466.7
AP7C:Ca (II)	$[M]^{-3}$	$[M^{3-} + 5(\text{Tris-HCl})]^{3-}$	1466.2	1466.7
	A	$[M^{5-} + 5(\text{Tris-HCl}) + Ca^{2+}]^{3-}$	1478.8	1479.3
AP7C:Zn (II)	$[M]^{-3}$	$[M^{5-} + 5(\text{Tris-HCl})]^{3-}$	1466.2	1466.7
	A	$[M^{-5-} + 5(\text{Tris-HCl}) + Zn^{2+}]^{3-}$	1487.8	1487.8
	B	$[M^{7-} + 5(\text{Tris-HCl}) + 2(Zn^{2+})]^{3-}$	1508.8	1508.9
AP7C:Cd (II)	$[M]^{-3}$	$[M^{3-} + 5(\text{Tris-HCl})]^{3-}$	1466.2	1466.7
	A	$[M^{5-} + 5(\text{Tris-HCl}) + Cd^{2+}]^{3-}$	1503.4	1503.5
	B	$[M^{7-} + 5(\text{Tris-HCl}) + 2(Cd^{2+})]^{3-}$	1540.0	1540.3
AP7C:La (III)	$[M]^{-3}$	$[M^{3-} + 5(\text{Tris-HCl})]^{3-}$	1466.2	1466.7
	A	$[M^{6-} + 5(\text{Tris-HCl}) + La^{3+}]^{3-}$	1511.4	1512.0
	B	$[M^{9-} + 5(\text{Tris-HCl}) + 2(La^{3+})]^{3-}$	1557.0	1557.3
	C	$[M^{12-} + 5(\text{Tris-HCl}) + 3(La^{3+})]^{3-}$	1602.6	1602.6
AP7C:Eu (III)	$[M]^{-3}$	$[M^{-3} + 5(\text{Tris-HCl})]^{3-}$	1466.2	1466.7
	A	$[M^{6-} + 5(\text{Tris-HCl}) + Eu^{3+}]^{3-}$	1515.8	1516.3
	B	$[M^{9-} + 5(\text{Tris-HCl}) + 2Eu^{3+}]^{3-}$	1565.6	1566.0
	C	$[M^{12-} + 5(\text{Tris-HCl}) + 3Eu^{3+}]^{3-}$	1615.6	1615.7

and ΔJ values (*42, 43*), reveal that the AP7C sequence possesses a partially folded structure (see Supporting Information). Subsequently, we quantitated NOE's and ^3J coupling constants to identify the location of helical and nonhelical structure within the 36 AA sequence. As shown in Table 2, we note that the ^3J coupling constants for H3, S4, R6, H10, C12, C15, T19, D20, G22, Q25, S33, and T35 are <6 Hz, which are inconsistent for extended or random coil structures but are consistent with α-helical structure (*32, 42, 43*). The presence of folded structure is also supported by PFG-NOESY data, which reveal the presence of (a) sequential backbone and $d_{\text{alpha-beta}}$ side chain NOE's (Figures 6, 7, and 8, and Table 2); (b) medium and long-range backbone NOE's; (c) medium and long-range backbone−side chain NOE's; and (d) sequential, medium, and long-range side chain−side chain NOE's. The presence of sequential, medium and long-range NOE values are generally associated with α-helical structure (*32, 33, 38*). As shown in Figure 8, these NOE's do not consistently appear throughout the entire AP7C sequence, and this indicates that the overall helical content of AP7C arises from the collective presence of short, dispersed helical regions as opposed to the presence of a continuous helical conformation.

If we use the $d_{\text{alpha-N(i, i + 4)}}$ NOE's and ^3J coupling constants as a guide to α-helix occurrence (*32*), then we conclude that there exist three short α-helical regions within the sequence: S4−R7 (segment A), C12−G22 (segment B), and M30−N36 (segment C)(Figure 8 and Table 2). These regions also possess side chain and backbone−side chain NOE's, indicating a compact structure. Note that within each of the three helical segments, we note that there are residues with $^3J >$ 5 Hz, which are outside the typical range for α-helical structure (*32*). From this, we conclude that there exists some degree of disordering or helix destabilization within each segment. This is also supported by our observations of Pro-induced conformational exchange/motion involving residues H3 and R6 of segment A and Y16 and Y17 of segment B, which would be consistent with regions that are experiencing interconversion between two conformational states. Thus, it is likely that Pro-induced chain dynamics plays an important role in the destabilization of helical structure within apo-AP7C. It is not known if similar dynamics and conformational exchange also occur within the 66 AA AP7 protein.

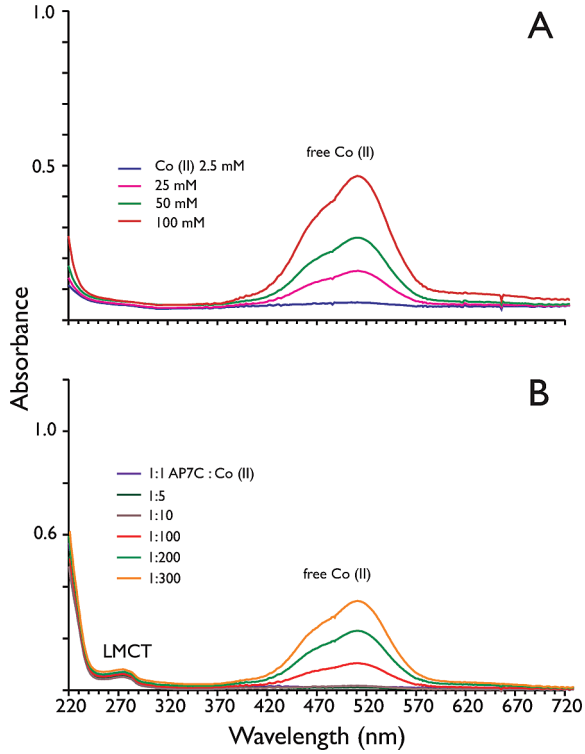

FIGURE 3: Absorption spectra of $CoCl_2$ in (A) 100 μM Tris-HCl at pH 7.5, and (B) in the presence of stoichiometric amounts of the AP7C polypeptide, 100 μM Tris-HCl at pH 7.5.

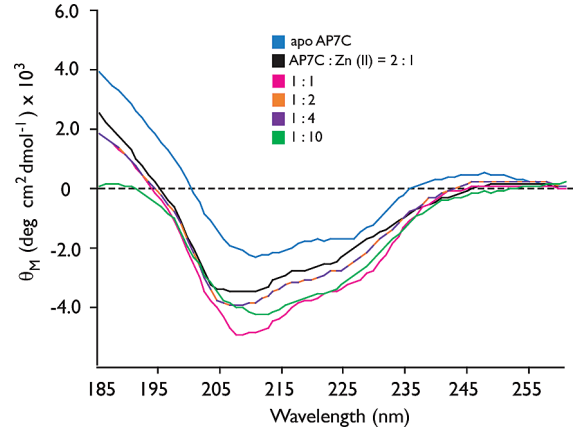

FIGURE 4: CD spectra of 8 μM AP7C in 100 μM Tris-HCl at pH 7.5, in the apo form and in the presence of stoichiometric amounts of $ZnCl_2$. Note that the 1:2 and 1:4 stoichiometry points give rise to identical, overlapping ellipticity curves, which we present here as a hybrid curve.

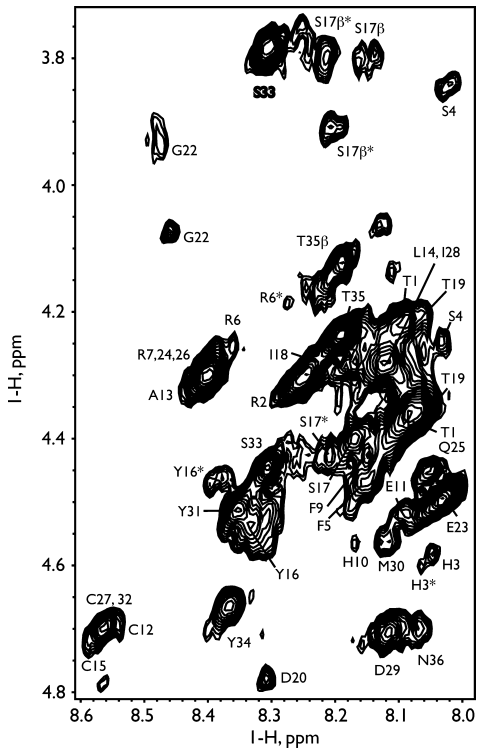

FIGURE 5: ^{1}H PFG-TOCSY spectra (600 MHz) obtained for apo-AP7C at pH 7.5 and 293 K. Acquisition parameters include a 75 ms mixing time, spectral window of 10 ppm, ^{1}H 90° pulse = 5.1 μs, 128 transients per experiment, with 1024 and 512 complex points acquired in t_2 and t_1, respectively, with the carrier centered on the water resonance. PFG Z-gradient parameters: $G_1 = G_2 =$ 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 the 2D spectra, with zero filling in the F2 dimension. Chemical exchange cross-peaks are denoted by the * symbol. Proton chemical shift are referenced from internal d_4-TSP.

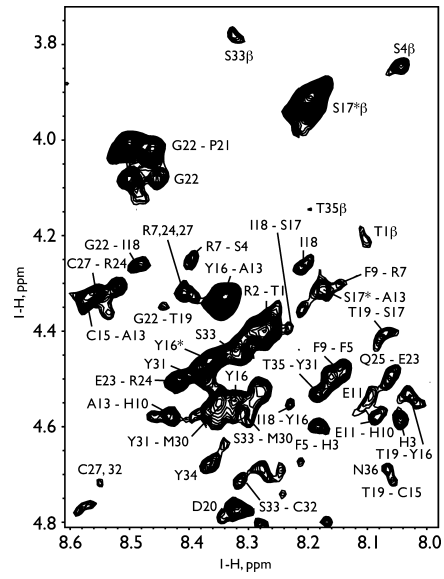

FIGURE 6: ^{1}H PFG-NOESY spectra (600 MHz) obtained for apo-AP7C at pH 7.5 and 293 K. Acquisition parameters include a mixing time of 200 ms, spectral window of 10 ppm, 90° pulse = 5.1 μs, 128 transients per experiment, with 1024 and 512 complex points acquired in t_2 and t_1, respectively, with the carrier centered on the water resonance. PFG Z-gradient parameters: $G_1 = G_2 =$ 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 the 2D spectra, with zero filling in the F2 dimension. Chemical exchange cross-peaks are denoted by the * symbol. Proton chemical shifts are referenced from internal d_4-TSP.

As to the remaining regions of the AP7C sequence, we tentatively assign secondary structure characteristics as follows. Downstream of segment A, the N-terminal region T1−H3 appears to be disordered, as evidenced by the absence of side chain−side chain and side chain−backbone NOE's, and the appearance of sequential and medium-to-short-range backbone NOE's. However, we note that the 4-AA P8−E11 region, which spans segments A and B, and the 7-AA E23−D29 region, which spans segments B and C, do possess medium range backbone NOE's and side chain NOE's. This data suggests that within these short intervening regions there is some degree of folded structure that is not consistent with either α-helix, β-sheet, or 3_{10}-helix (*32*). We suspect that these short intervening regions are either β-turn and/or loop in nature (Figure 8). Evidence for this hypothesis can be found in Table 2, where we note that the coupling constants for the third (H10) and fourth (E11) residues of the -PFHE- segment are approximately 2 and 8 Hz, respectively, which is consistent with a β-turn type I structure (*32*). For the heptapeptide -ERQRCID- sequence, ^3J values of 5, 9, and 11 Hz were obtained for the second (R24), sixth (I28), and seventh (D29) residues of this sequence block; unfortunately, we were unable to unambiguously determine coupling constants for the remaining residues. Nonetheless, the sequence length and coupling constants of -ERQRCID- are not consistent with a β-turn (*32*), but are consistent with a loop region (*44*).

In conclusion, our NMR studies confirm the results obtained from our earlier CD studies (*14*) and reveal that the AP7C polypeptide sequence consists of three short α-helical segments, A, B, and C, along with β-turn and loop-like intervening segments (Figure 8). Interestingly, the RING putative Zn (II) binding motif, -RRPFHECALCYSI-, encompasses portions of helical segments A and B in AP7C and includes the β-turn type I region which unites these two partially helical domains. Given that the NOE's, ^3J values, and ^{1}H NMR conformational shifts (see Supporting Information) do not correspond exactly to α-helical specifications, we believe that the structure of this 36 AA polypeptide has some degree of inherent lability or instability, most likely as a result of P8, P21 cis–trans interconversion events occurring within the sequence.

Structure Refinement of AP7C. For structure refinement simulations of AP7C, we employed H−H distance restraints obtained from ^{1}H NMR PFG NOESY experiments and available ^3J coupling constants to define φ angle ranges. We obtained plausible lowest energy conformational ensembles ($n = 10$ for each peptide) that match the NMR data sets and exhibit torsion angle distributions that are distributed within the α-helical and β-turn regions of the Ramachandran map as well as the β-strand, 3_{10} helix, and other undefined regions (see Supporting Information). These nonhelical dihedral angles arise from the conformationally labile N-terminal and loop regions of AP7C (Figure 8) and likewise represent the sequence regions downstream of P8 and P21 that experience conformational exchange (Figures 5 and 6). The Pro-induced conformational exchange phenomena has a significant impact on the distribution of φ, ψ dihedral pairs for AP7C, as evidenced by the significant degree of scatter that we observe

Table 2: ^{1}H NMR (600 MHz) Chemical Shifts and ^3J Coupling Constants for Apo-AP7C at pH 7.5 and 293 K[a]

residue	NH	CH_{alpha}	CH_{beta}	CH_{gamma}	CH_{delta}	$CH_{epsilon}$	ring	side NH	^3J
T1	8.09	4.38	4.23	1.16					7.6
R2	8.28	4.33	1.74, 1.65	N/O	1.74, 1.65	3.04, 3.00		7.10, 6.77	9.6
H3	8.05	4.58	3.08, 2.92				7.68, 6.87		4.9
H3*	8.06	4.61	3.09, 2.98				7.68, 6.87		12.3
S4	8.03	4.25	3.85						3.3
F5	8.18	4.50	2.92, 2.81				7.30, 7.00, 6.77		13.0
R6	8.36	4.26	1.75, 1.65	N/O	1.75, 1.65	3.01, 2.81		7.13, 6.83	2.8
R6*	8.27	4.18	1.75, 1.65	N/O	1.76, 1.65	N/O		7.10, 6.85	9.8
R7	8.40	4.3	1.78, 1.70	1.38	1.78, 1.70	3.05, 2.92		7.08, 6.78	N/D
P8		4.43	2.56	2.02	3.93				
F9	8.15	4.43	2.96, 2.81				7.24, 6.99, 6.73		11.1
H10	8.17	4.57	3.10, 2.94				7.64, 6.87		2.3
E11	8.09	4.70	2.74	2.81, 2.69					8.0
C12	8.54	4.70	3.16, 3.11						3.9
A13	8.43	4.32	1.39						7.2
L14	8.13	4.28	1.89	1.60	0.87				N/D
C15	8.59	4.71	3.17, 3.11				7.41, 6.78		2.8
Y16	8.32	4.54	3.18, 3.06				7.42, 6.76		13.1
Y16*	8.37	4.46	3.11, 3.01				7.45, 6.79		9.9
S17	8.16	4.40	3.80						8.0
S17*	8.22	4.43	3.91, 3.80						10.7
I18	8.23	4.28	1.19	0.87	N/O				6.0
T19	8.06	4.35	4.21						5.0
D20	8.10	4.70	2.78, 2.71						4.9
P21	-	4.42	2.61	2.11	3.99				N/A
G22	8.46	4.07, 3.93							5.4
E23	8.03	4.49	2.79, 2.69	2.79					7.4
R24	8.40	4.30	1.78, 1.70	1.38	1.78, 1.70	3.05, 2.92		7.08, 6.78	N/D
Q25	8.06	4.45	2.80	2.96				7.49, 6.81	5.4
R26	8.40	4.30	1.78, 1.70	1.38	1.78, 1.70	3.05, 2.92		7.08, 6.78	N/D
C27	8.56	4.70	3.10, 2.97						N/D
I28	8.12	4.29	1.17	N/O	N/O				N/D
D29	8.12	4.70	2.92, 2.75						8.8
M30	8.11	4.56	1.89	2.81, 2.65					10.6
Y31	8.35	4.51	3.18, 3.06				7.36, 6.78		7.1
C32	8.56	4.70	3.10, 2.97						N/D
S33	8.31	4.45	3.78						5.0
Y34	8.37	4.66	3.11, 2.91				7.41, 6.78		8.8
T35	8.19	4.24	4.12	1.16					5.5
N36	8.06	4.67	2.96, 2.71					7.56, 6.88	7.5

[a] Proton chemical shifts were determined from PFG-TOCSY and NOESY experiments and are referenced from internal d_4-TSP. Diastereotopic protons are separated by commas. ^3J coupling constants were determined using z-PFG-TOCSY experiments and spectral processing methods that permit coupling constant measurements from TOCSY spectra (*8, 15*). N/O = not observed; N/D = not determined due to cross-peak overlap.

in the Ramachandran plot and in the mean backbone rmsd value for this 10 conformer library (5.54 ± 1 Å). Consequently, there is an above-average degree of variation in this structural ensemble.

The lowest energy backbone conformation for AP7C (Figure 9) conforms to our NMR data set in that segments A(S4−R7), B (C12−G22), and C (M30−N36) clearly exhibit partial helical structure, and the N-terminal TRH-region is unstructured. Similarly, the P8−E11 region appears to adopt a turn- or loop-like structure, whereas the larger E29−D29 region exhibits a more extended loop-like structure. Interestingly, both P8 and P21 are located in regions where helical segments transition into either turn- or loop-like structures, and these Pro residues appear to be responsible for affecting the mobility and/or flexibility of downstream polypeptide segments A and B (Figures 5−7). We speculate that Pro-induced motion may affect the ability of AP7C to bind Zn (II) as well as modulate the structure of the AP7 C-terminal domain in a manner that fosters interaction with other nacre proteins (*4*).

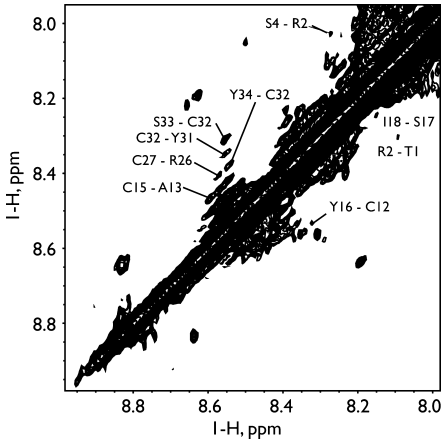

FIGURE 7: ^{1}H PFG-NOESY NH-NH fingerprint spectra (600 MHz) obtained for apo-AP7C at 293 K. Acquisition and processing parameters are identical to those presented in Figure 6.

At present, the structural coordinates for RNF32 and RKR1 RING domains are not available. However, the solution structure of the Zn (II)-bound 60 AA C-terminal RING domain taken from the RING−IBR−RING triad motif of the ubiquitin-conjugating enzyme E2-binding protein-1, HHARI (*21*) is available (1WD2, Figure 10). Here, we observe that the 60 AA backbone structure of the HHARI C-RING domain is somewhat similar to AP7C (Figure 10). First, HHARI also features two helical domains that are

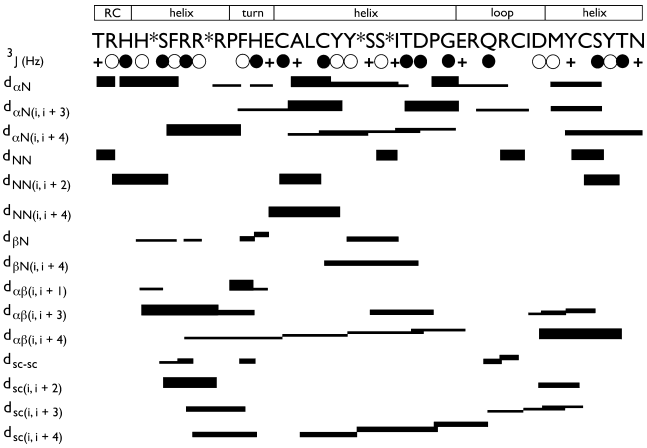

FIGURE 8: Summary of sequence-specific NMR parameters for apo-AP7C at pH 7.5 and 293 K. The summary includes intraresidue, interresidue sequential, and medium to long-range backbone NOE's, and ^3J coupling constants (^3J). ^3J coupling constants were determined using z-PFG-TOCSY experiments and spectral processing methods that permit coupling constant measurements from TOCSY spectra (*8, 15*). Below the amino acid sequence, filled circles identify residues with $^3J_{HNHalpha} < 6$ Hz, indicative of local α-helical-type conformation; open circles correspond to $^3J_{HNHalpha} > 8.0$ Hz, indicative of residues in extended chain conformation; crosses identify residues with $^3J_{HNHalpha} = 6.0-8.0$ Hz, which corresponds to other conformations, including random coil (RC). For the sequential proton–proton NOE connectivities, the three different thicknesses of connecting bars indicate strong, medium, and weak NOE intensities. Note that $d_{alphaN} = d_{alphaN(i+1)}$, $d_{alphaN(i+2)}$; $d_{NN} = d_{NN(i+1)}$; $d_{betaN} = d_{betaN(i+1)}$, $d_{betaN(i+2)}$. The terminology and presentation of this figure follows the standardization set forth in ref *45*.

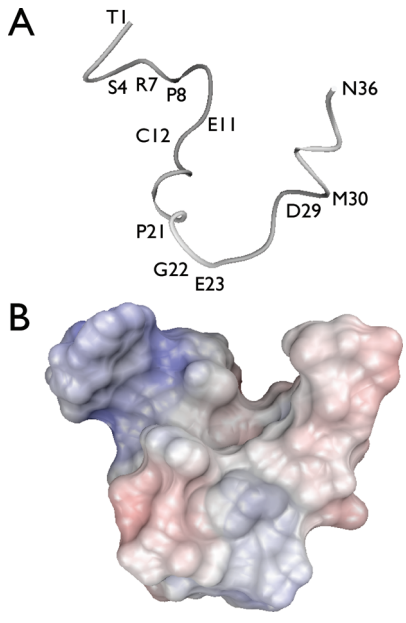

FIGURE 9: (A) SA/MD XPLOR lowest energy backbone only conformation of AP7C. (B) Calculated APBS solvent electrostatic molecular surface of AP7C lowest energy conformer (full atom representation). Electrostatic surface rendering is as follows: Red = anionic; blue = cationic; and the polypeptide molecule is shown in the same orientation as that in A.

linked by loop regions, and second, there are unfolded regions as well within the HHARI C-RING domain (*21*). However, the important difference to note is that the AP7C RING-like domain structure folds without the assistance of Zn (II), whereas HHARI and other RING domains appear

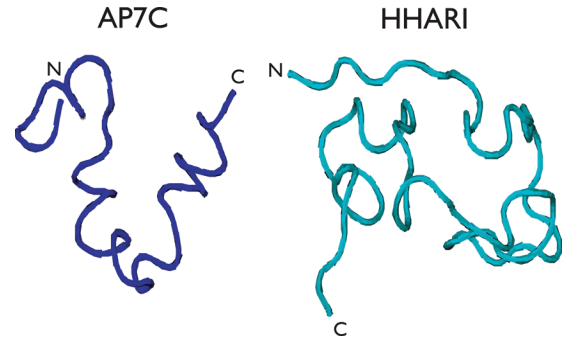

FIGURE 10: Comparison of backbone structures (tube representation) of 36 AA AP7C and the solution structure of the Zn (II)-bound 60 AA C-terminal RING domain of the ubiquitin-conjugating enzyme E2-binding protein-1, HHARI (*21*). Given the sequence length discrepancies and the lack of meaningful sequence overlap between the two polypeptides, rmsd backbone fitting procedures were not performed.

to require Zn (II) to achieve a folded state (*21*). At the present time, we do not know the exact location of Zn (II) binding site(s) within the AP7C domain, although we speculate that the His, Cys-containing -RRPFHECALCYSI- sequence region is one putative site for Zn (II) interactions. However, additional studies will be necessary to establish the true site(s) of Zn (II) binding within this sequence, and these are currently in progress.

Solvation Electrostatics of the AP7C RING-Like Motif. Given that RING domains are associated with protein−protein interactions (*19–23*), these motifs may exhibit molecular surface features, such as electrostatics, that are conducive to fostering interactions between AP7 and other nacre proteins. To evaulate this potential within the AP7C sequence, we calculated the Poisson–Boltzmann solvation energies (*8, 34*) for this structure and generated a corresponding solvation electrostatic surface (Figure 9, full atom representation). An analysis of the solvation electrostatic molecular surface of AP7C reveals the presence of a significant cationic region near the N-terminus. This surface represents the pseudo-RING motif and includes contributions from the alpha amino group of T1, the imidzole groups of H3 and H10, and the guanidinium groups of R6 and R7 (Figures 1 and 9). We also note the presence of anionic regions located at E23 and D29 and at the C-terminal end of the AP7C segment. Thus, the AP7C sequence possesses several solvent-accessible, charged surfaces that could serve at sites of complementary electrostatic interaction with other nacre proteins such as AP24, which does possess electrostatic clusters downstream of its mineral modification domain (*4*).

DISCUSSION

The 36 AA C-terminal sequence of the nacre protein, AP7 (*4*), was previously demonstrated to have no significant interactions with calcium carbonate crystals in vitro (*14*). We have now uncovered evidence that this 36 AA sequence may have evolved for other purposes. This sequence exhibits partial homology to the Zn (II) binding RING sequence family found in a number of intracellular proteins (*19–23*). This C-terminal sequence can bind Zn (II), Co (II), and other multivalent metal ions (Figures 2 and 3, and Table 1) in 1:1 and 2:1 metal ion:peptide stoichiometries, and the ligand coordination number for polypeptide:Zn (II) interaction is either mono, bi, or tridentate in nature. Furthermore, the

AP7C sequence distinguishes itself from typical Zn (II) finger motifs (*24–29*) in the following way. First, the sequence does not appear to form a tetracoordinate complex with Co (II) (Figure 3). Second, the apo-form of this sequence does not exist as an unfolded, random-coil conformation (Figure 4) but instead adopts a partially folded state that consists of three short α-helical segments, A, B, and C, along with a 4 AA type I β-turn and 7 AA loop-like intervening segments (Figures 8 and 10). The RING-like region, -RRPFHECAL-CYSI- comprises a portion of helical segments A and B and the type I β-turn region that unites these two helical segments. This structure is qualitatively similar to the solution structure of the Zn (II)-bound HHARI C-RING (*21*). Interestingly, our CD data suggests that the helical content of apo-AP7C does increase when Zn (II) ions are introduced (Figure 4). Unfortunately, because of aggregation problems in the presence of Zn (II) at higher peptide concentrations, we were unable to utilize NMR spectroscopy to verify if additional helical folding occurred in AP7C (data not shown). Similarly, although we know that AP7C conveys helical properties to the AP7 protein (*14*), we do not know if the extent of helical content or the presence of intervening nonhelical regions is the same in AP7. This issue will be revisted once appropriate experimental conditions for working with AP7 are identified.

Two unexpected structural features were noted regarding the RING-like domain in AP7C. First, we identified the presence of Pro hinge-like motion (P8, P21) and two-state conformational exchange involving H3, R6, Y16, and S17 within the RING-like -RRPFHECALCYSI- region (Figures 5 and 6). Pro ring interconversion and resulting polypeptide backbone motion has also been noted in other polypeptide sequences (*11, 38*), and we believe that P8 and P21 contribute to helix destabilization within apo-AP7C and generate conformational heterogeneity within this sequence. However, it is not clear if this phenomena is also present in the native AP7 protein or if it plays an important role in nacre protein function, and thus the significance of this phenomenon must await further studies. Second, our Poisson–Boltzmann solvation partial atomic charge calculations indicate that the T1, H3, R6, R7, and H10 residues of the RING-like region generate a significant cationic surface near the N-terminal region of AP7C. Given that RING motifs are associated with protein−protein interaction, this solvent-accessible cationic region within AP7 may represent a putative site for interaction with other nacre proteins such as AP24 that possess electrostatic sequence clusters (*4*). The fact that this cationic region lies immediately upstream of the neighboring mineral modification domain in the AP7 sequence (*4*) may also be significant with regard to AP7 interacting simultaneously with calcium carbonates and other nacre macromolecules or to the positioning of individual nacre protein domains in proximity to each other or to nucleating minerals.

Although RING domain variants have been identified (*19–23*), there exists uncertainty as to their true function and how the structure of these unusual domains contributes to function. To a certain extent, our understanding of these domains has been compromised by solubility and aggregation problems (*19–23*). Our study now reveals the existence of another RING variant that offers two twists on the existing story of the RING domain. First, a RING-like domain has now been identified within an extracellular biomineralization protein, AP7. This is a departure from the typical location of RING domains within intracellular proteins (*19–23*). Second, this RING variant does not require Zn (II) to adopt a folded structure in solution, a peculiar feature that distinguishes this sequence from all previously reported RING motifs (*19–23*). We should note that the Zn (II) ion is not totally without effect on the structure of AP7C; as shown in Figure 4, there is some evidence that the helical content of AP7C increases upon the addition of Zn (II) at low polypeptide concentrations. However, because of Zn (II)-inducd aggregation problems that occur at higher polypeptide concentrations for NMR spectroscopy, we were unable to verify whether Zn (II) addition leads to further stabilization of helical segments within the AP7C structure, and thus the full effect of Zn (II) ion on the folding of AP7C remains to be determined.

Given that the AP7C sequence exhibits only a partial homology to other RING domains (Figure 1) (*19*), it is likely that the pseudo-RING motif within the AP7 sequence evolved along different lines from other RING family sequences for specific reasons that are not completely understood at present. However, the fact that the apo AP7C sequence can fold suggests that structural stability in the absence of Zn (II) conveyed some functional advantage in an extracellular matrix environment that is rich in other cations such as Ca (II) and Mg (II). Unfortuantely, since AP7 exists within an intracrystalline environment and mineral dissociative procedures (i.e., use of chelation agents or acid conditions) are used for nacre matrix protein extraction (*4*), there is no reliable method for isolating a Zn (II)-bound form of AP7. Moreover, the metal ion content of the nacre matrix has yet to be established, and thus the presence of extracellular Zn (II) ions remains unknown. Clearly, additional studies will be needed to determine what role, if any, Zn (II) has on AP7 structure, function, and on nacre mineralization. Moreover, it would be interesting to learn the existence of other RING-like domains in extracellular proteins and how these domains function in other protein-directed processes.

ACKNOWLEDGMENT

This paper represents contribution number 40 from the Laboratory for Chemical Physics, New York University.

SUPPORTING INFORMATION AVAILABLE

Determination and evaluation of proton conformational shifts and coupling constants, and Ramachandran φ, ψ dihedral angle distribution plot for AP7C conformer libraries. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Mann, S., Webb, J., and Williams, R. J. P., Eds. (1989) *Biomineralization: Chemical and Biochemical Perspectives*, VCH, Weinheim, Germany.
2. Lowenstam, H. A., and Weiner, S. (1989) *On Biomineralization*, Oxford Press, New York, NY.
3. Samata, T., Hayashi, N., Kono, M., Hasegawa, K., Horita, C., and Akera, S. (1999) A new matrix protein family related to the nacreous layer formation of Pinctada fucata. *FEBS Lett. 462*, 225–232.
4. Michenfelder, M., Fu, G., Lawrence, C., Weaver, J. C., Wustman, B. A., Taranto, L, and Evans, J. S. (2003) Characterization of two molluscan crystal-modulating biomineralization proteins and identification of putative mineral binding domains. *Biopolymers 70*, 522–533.

5. Ma, Z., Huang, J., Sun, J., Wang, G., Li, C., Xi, L., and Zhang, R. (2007) A novel extrapallial fluid protein controls the morphology of nacre lamellae in the pearl oyster, *Pinctada fucata*. *J. Biol. Chem. 282*, 23253–23260.
6. Fu, G., Qiu, S. R., Orme, C. A., Morse, D. E., and DeYoreo, J. J. (2007) Acceleration of calcite kinetics by nacre proteins. *Adv. Mater. 17*, 2678–2683.
7. Mann, K., Siedler, F., Treccani, L., Heinemann, F., and Fritz, M. (2007) Perlinhibin, a cysteine, histidine, and arginine-rich mini-protein from abalone (*Haliotis laevigata*) nacre, inhibits in vitro calcium carbonate crystallization. *Biophys. J. 93*, 1246.
8. Collino, S., and Evans, J. S. (2007) Structural features that distinguish kinetically distinct biomineralization polypeptides. *Biomacromolecules 8*, 1686–1694.
9. Delak, K., Collino, S., and Evans, J. S. (2007) Expected and unexpected effects of amino acid substitutions on polypeptide-directed crystal growth. *Langmuir 23*, 11951–11955.
10. Seker, O. U. S., Wilson, B., Dincer, I. S., Kim, I. W., Oren, E. E., Evans, J. S., Tamerler, C, and Sarikaya, M. (2007) Adsorption behavior of linear and cyclic genetically engineered platinum binding polypeptides. *Langmuir 23*, 7895–7900.
11. Kulp, J. L., III, Minamisawa, T., Shiba, K., and Evans, J. S. (2007) Conformational properties of an artificial protein that regulates the nucleation of inorganic and organic crystals. *Langmuir 23*, 3857–3863.
12. Kim, I. W., Darragh, M. R., Orme, C., and Evans, J. S. (2006) Molecular "tuning" of crystal growth by nacre-associated polypeptides. *Cryst. Growth Des. 6*, 5–10.
13. Collino, S., Kim, I. W., and Evans, J. S. (2006) Identification of an "acidic" C-terminal mineral modification sequence from the mollusk shell protein, Asprich. *Cryst. Growth Des. 6*, 839–842.
14. Kim, I. W., Collino, S, Morse, D. E., and Evans, J. S. (2006) A crystal modulating protein from molluscan nacre that limits the growth of calcite in vitro. *Cryst. Growth Des. 6*, 1078–1082.
15. Kim, I. W., Morse, D. E., and Evans, J. S. (2004) Molecular characterization of the 30-AA N-terminal mineral interaction domain of the biomineralization protein AP7. *Langmuir 20*, 11664–11673.
16. Kim, I. W., DiMasi, E., and Evans, J. S. (2004) Identification of mineral modulation sequences within the nacre-associated oyster shell protein, n16. *Cryst. Growth Des. 4*, 1113–1118.
17. Wustman, B. A., Morse, D. E., and Evans, J. S. (2004) Structural characterization of the N-terminal mineral modification domains from the molluscan crystal-modulating biomineralization proteins, AP7 and AP24. *Biopolymers 74*, 363–376.
18. Shen, X., Belcher, A. M., Hansma, P. K., Stucky, G. D., and Morse, D. E. (1997) Molecular cloning and characterization of Lustrin A, a matrix protein from shell and pearl nacre of *Haliotis rufescens*. *J. Biol. Chem. 272*, 32472–32481.
19. (a) van Baren, M. J., van der Linde, H. C., Breedveld, G. J., Barrends, W. M., Rizzu, P., de Graaf, E., Oostra, B. A., and Heutink, P. (2002) A double RING-H2 domain in RNF32, a gene expressed during sperm formation. *Biochem. Biophys. Res. Commun. 292*, 58–86. (b) Braun, M. A., Costa, P. J., Crisucci, E. M., and Arndt, K. M. (2007) Identification of Rkr1, a nuclear RING domain protein with functional connections to chromatin modification in Saccharomyces cerevisiae. *Mol. Cell. Biol. 27*, 2800–2811.
20. (a) Saurin, A. J., Borden, K. L. B., Boddy, M. N., and Freemont, P. S. (1996) Does this have a familiar RING? *TIBS 21*, 208–215. (b) Freemont, P. S. (2000) Ubiquitination: RING for destruction? *Curr. Biol. 10*, R84–R87.
21. Capili, A. D., Edghill, E. L., Wu, K., and Borden, K. L. B. (2004) Structure of the C-terminal RING finger from a RING-IBR-RING/TRIAD motif reveals a novel zinc-binding domain distinct from a RING. *J. Mol. Biol. 340*, 1117–1129.
22. (a) Borden, K. L. B. (1998) RING fingers and B-boxes: zinc-binding protein-protein interaction domains. *Biochemistry and Cell Biology 76*, 351–358. (b) Borden, K. L. B. (2000) RING domains: Master builders of molecular scaffolds? *J. Mol. Biol. 295*, 1103–1112.
23. Eisenhaber, B., Chumak, N., Eisenhaber, F., and Hauser, M.-T. (2007) The ring between RING fingers (RBR) protein family. *Genome Biol. 8*, 209–215.
24. Worthington, M. T., Amann, B. T., Nathans, D., and Berg, J. M. (1996) Metal binding properties and secondary structure of the zinc-binding domain of Nup475. *Proc. Natl. Acad. Sci. U.S.A. 93*, 13754–13759.
25. Michael, S. F., Kilfoil, V. J., Schmidt, M. H., Amann, B. T., and Berg, J. M. (1992) Metal binding and folding properties of a minimalist Cys_2His_2 zinc finger peptide. *Proc. Natl. Acad. Sci. U.S.A. 89*, 4796–4800.
26. Frankel, A. D., Berg, J. M., and Pabo, C. O. (1987) Metal-dependent folding of a single zinc finger from transcription factor I11A. *Proc. Natl. Acad. Sci. U.S.A. 84*, 4841–4845.
27. Nomura, A, and Sugiura, Y. (2002) Contribution of individual zinc ligands to metal binding and peptide folding of zinc finger peptides. *Inorg. Chem. 41*, 3693–3698.
28. Westman, B. J., Perdomo, J., Matthews, J. M., Crossley, M., and Mackay, J. P. (2004) Structural studies on a protein-binding zinc-finger domain of Eos reveal both similarities and differences to classical zinc fingers. *Biochemistry 43*, 13318–13327.
29. Parraga, G., Horvath, S., Hood, L., Young, E. T., and Klevit, R. E. (1990) Spectroscopic studies of wild-type and mutant "zinc finger" peptides: Determinants of domain folding and structure. *Proc. Natl. Acad. Sci. U.S.A. 87*, 137–141.
30. Goddard, T. D., Kneller, D. G. (2004) *SPARKY 3*, version 3.110, University of California, San Francisco, San Francisco, CA.
31. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR. 6*, 277–293.
32. (a) Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York, NY. (b) Berger, S., and Braun, S. (2004) *200 and More NMR Experiments*, Wiley-VCH GmbH and Co., Morlenbach, Germany.
33. Pardi, A., Billeter, M., and Wuthrich, K. (1984) Calibration of the angular dependence of the amide proton-C-alpha proton coupling constants, $^3J_{HNalpha}$, in a globular protein. Use of $^3J_{HNalpha}$ for identification of helical secondary structure. *J. Mol. Biol. 180*, 741–751.
34. Baker, N. H., Sept, D., Joseph, S., Holst, N. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037–10041.
35. Prigodich, R. V., O'Connor, T., and Coleman, J. E. (1985) Proton, cadmium-113, and phosphorus-31 NMR of osteocalcin (bovine gamma.-carboxyglutamic acid containing protein). *Biochemistry 24*, 6291–6298.
36. Wojcik, J., Goral, J., Pawlowski, K., and Bierzynski, A. (1997) Isolated calcium-binding loops of EF-Hand proteins can dimerize to form a native-like structure. *Biochemistry 36*, 680–687.
37. MacManus, J. P., Hogue, C. W., Marsden, B. J., Sikorska, M., and Szabo, A. (1990) Terbium luminescence in synthetic peptide loops from calcium-binding proteins with different energy donors. *J. Biol. Chem. 265*, 10358–10366.
38. Kulp, J. L., Shiba, K., and Evans, J. S. (2005) Probing the conformational features of a phage display polypeptide sequence directed against single wall carbon nanohorn surfaces. *Langmuir 21*, 11907–11914.
39. (a) Dumy, P., Keller, M., Ryan, D. E., Rohwedder, B., Wohr, T., and Mutter, M. (1997) Pseudo-prolines as a molecular hinge: Reversible induction of cis amide bonds into peptide backbones. *J. Am. Chem. Soc. 119*, 918–925. (b) Butcher, D. J., Nedved, M. L., Neiss, T. G., and Moe, G. R. (1996) Proline pipe helix: Structure of the Tus proline repeat determined by 1-H NMR. *Biochemistry 35*, 698–703. (c) Bhattacharyya, R, and Chakrabarti, P. (2003) Stereospecific interactions of proline residues in protein structures and complexes. *J. Mol. Biol. 331*, 925–940.
40. (a) Cammers-Goodwin, A., Allen, T. J., Oslick, S. L., McClure, K. F., Lee, J. H., and Kemp, D. S. (1996) Mechanism of stabilization of helical conformations of polypeptides by water containing trifluoroethanol. *J. Am. Chem. Soc. 118*, 3082–3090. (b) Krieger, F., Moglich, A., and Kiefhaber, T. (2005) Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J. Am. Chem. Soc. 127*, 3346–3352.
41. (a) Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol. 222*, 311–333. (b) Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) 1-H, 13-C and 15-N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR 6*, 135–140. (c) Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) 1-H, 13-C and 15-N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR 5*, 67–81.
42. Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., and Dobson, C. M. (1996) Analysis of main chain

torsion angles in proteins: Prediction of NMR coupling constants for native and random coil conformations. *J. Mol. Biol. 255*, 494–506.

43. Serrano, L. J. (1995) Comparison between the phi distribution of the amino acids in the protein database and NMR data iIndicates that amino acids have various phi propensities in the random coil conformation. *J. Mol. Biol. 254*, 322–333.
44. (a) Wustman, B., Santos, R., Zhang, B., and Evans, J. S. (2002) Identification of a glycine loop-like coiled structure in the 34-AA Pro, Gly, Met repeat domain of the biomineral-associated protein, PM27. *Biopolymers 65*, 1305–1318. (b) Zhang, B., Wustman, B., Morse, D. E., and Evans, J. S. (2002) Model peptide studies of sequence regions in the elastomeric biomineralization protein, Lustrin A. I. The C-domain consensus -PG-, -NVNCT- motif. *Biopolymers 64*, 358–369.
45. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wuthrich, K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Biomol. NMR 12*, 1–23.

BI701949P