A Solution NMR Investigation into the Early Events of Amelogenin Nanosphere Self-Assembly Initiated with Sodium Chloride or Calcium Chloride[†]

Garry W. Buchko,[‡] Barbara J. Tarasevich,[‡] Jacky Bekhazi,[‡] Malcolm L. Snead,[§] and Wendy J. Shaw*[,‡]

Pacific Northwest National Laboratory, Richland, Washington, and Center for Craniofacial Molecular Biology, University of Southern California, Los Angeles, California

Received July 1, 2008; Revised Manuscript Received October 21, 2008

ABSTRACT: Using solution-state NMR spectroscopy, new insights into the early events governing amelogenin supramolecular self-assembly have been identified using sodium chloride and calcium chloride to trigger the association. Two-dimensional $^{1}H-^{15}N$ HSQC spectra were recorded for ^{15}N- and ^{13}C-labeled murine amelogenin as a function of increasing NaCl and $CaCl_2$ concentration beginning with solution conditions of 2% acetic acid at pH 3.0, where amelogenin was monomeric. Residue specific changes in molecular dynamics, manifested by the reduction in intensity and disappearance of $^{1}H-^{15}N$ HSQC cross-peaks, were observed with the addition of either salt to the protein. With increasing NaCl concentrations, residues between T21 and R31 near the N-terminus were affected first, suggesting that these residues may initiate amelogenin dimerization, the first step in nanosphere assembly. At higher NaCl concentrations, more residues near the N-terminus (Y12−I51) were affected, and with further additions of NaCl, residues near the C-terminus (L141−T171) began to show a similar change in molecular dynamics. With increasing $CaCl_2$ concentrations, a similar stepwise change in molecular dynamics involving essentially the same set of amelogenin residues was observed. As the concentration of either salt was increased, a concomitant increase in the estimated overall rotational correlation time (τ_c) was observed, consistent with assembly. Self-assembly into a dimer or trimer was established with dynamic light scattering studies under similar conditions that showed an increase in diameter of the smallest species from 4.1 nm in the absence of salt to ~10 nm in the presence of salt. These results suggest a possible stepwise interaction mechanism, starting with the N-terminus and followed by the C-terminus, leading to amelogenin nanosphere assembly.

Dental enamel, the layer of elongated carbonated hydroxyapatite on the outer layer of the tooth, is the hardest tissue in the human body (*1*). It needs to be strong because it is exposed to repeated masticatory, parafunctional, and occasional impact loading, but unlike the dominant biomineral in the human body (*2*), mesenchyme-derived bone, enamel cannot self-repair nor can it undergo remodeling, and therefore, it must last a lifetime. Enamel's resistance to wear and deformation are due to a combination of high mineral content and unique three-dimensional structural organization (*3*). Ninety-five percent of mature enamel consists of long and narrow crystals of carbonated hydroxyapatite packed into parallel arrays, called enamel rods, that are intricately interwoven into an unique lattice architecture (*4, 5*). There is little, if any, matrix protein remaining in mature enamel, another feature differentiating enamel markedly from dentin and bone.

The nucleation, growth, and organization of enamel (amelogenesis) are extracellular processes orchestrated by one predominant enamel matrix protein, amelogenin (*5*−*8*), that is secreted by the inner enamel epithelium (ameloblasts) and found exclusively in the region of newly formed enamel (*9*). Specific mutations to amelogenin and amelogenin deficient mice both result in dramatic enamel phenotypes similar to amelogenesis imperfecta in humans (*5*), observations that strongly suggest amelogenin plays an essential role in amelogenesis. The amino- and carboxy terminal regions of this low-molecular-weight (~20 kDa, depending on species) protein are highly conserved among species, suggesting these regions serve important functional roles in mediating enamel biomineralization (*10*). Except for the hydrophilic and charged C-terminal region (*11*), amelogenin is hydrophobic in nature (*10*) and contains a central region enriched in the amino acids P, L, H, and Q (HQP-rich region). Indeed, it is variations in the length of this central HQP-rich region that largely account for the differences in the size of amelogenin among various species. The secondary structure of amelogenin has been studied extensively by circular dichroism (CD[1]) (*12*−*14*) and small-angle X-ray spectroscopy (SAXS) (*15*). The general consensus is that in monomeric form, the tyrosine-rich N-terminal region forms an extended β-sheet/β-strand structure, while the HQP-rich

[†] This work was supported by NIH-NIDCR Grant DE-015347. The research was performed at the Pacific Northwest National Laboratory (PNNL), a facility operated by Battelle for the U.S. Department of Energy and at the W.R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the U.S. DOE Biological and Environmental Research program.

* To whom correspondence should be addressed. Tel: 509-375-5922. Fax: 509-375-6660. E-mail: Wendy.shaw@pnl.gov.

[‡] Pacific Northwest National Laboratory.

[§] University of Southern California.

[1] Abbreviations: CD, circular dichroism; DLS, dynamic light scattering; NNLS, non-negatively constrained least squares; rp(H)M180, recombinant mouse amelogenin with N-terminal polyhistidine tag; SEC, size exclusion chromatography.

10.1021/bi8018288 CCC: $40.75
Published on Web 11/14/2008

region adopts a β-spiral/polyproline II (PPII) structure (*5, 14, 15*). This structure appears to be very sensitive to pH, temperature, ionic strength, and metal ions. Of the many interesting features of amelogenin, perhaps the most remarkable is its ability, under the correct conditions, to self-assemble into a unique quaternary structure, termed nanosphere, that appears to be an essential functional and structural component of the secretory-stage enamel matrix (*8, 16, 17*).

Amelogenin nanospheres composed of multiple amelogenin molecules are 20−50 nm in radius and have been observed both *in vivo* (*8*) and *in vitro* (*7*). *In vitro*, amelogenin molecules will spontaneously self-associate to form nanospheres under many different types of conditions (*17*−*19*). Mutants of amelogenin that have disrupted nanosphere formation *in vitro* (*20*) are observed to cause malformed enamel in knock-in mice *in vivo* (*21*) supporting the importance of this quaternary structure in enamel biomineralization. Protein proteolysis studies (*22*) and yeast two-hybrid assays (*23*) suggest that the regions of the protein essential in nanosphere self-assembly are M1-M42 and S157-K173. It is hypothesized that nanosphere assembly occurs through the progressive accretion of amelogenin molecules with monomers first associating into dimer and trimer building blocks (*19, 24*). However, molecular details of the self-assembly process are still not well understood. This is partially because amelogenin self-association is a complex process and depends on the interplay between protein concentration and the many properties of the solution (ionic strength, pH, solutes, and temperature) (*17*−*19*) that, in turn, affect the physical properties of the protein (hydrophobicity and electrostatics) (*5, 25, 26*).

To better understand details of the initial events of nanosphere self-assembly, we have used solution NMR spectroscopy and dynamic light scattering to probe the behavior of mouse amelogenin (*11*), at a fixed concentration, as a function of increasing ionic strength using the salts sodium chloride and calcium chloride. The smaller monomeric (~20 kDa) and dimeric (~40 kDa) species are accessible to solution state NMR experiments unlike the much larger, fully assembled, nanospheres (~2,000 kDa). Because we have assigned the ^{1}H, ^{13}C, and ^{15}N chemical shifts for amelogenin under acidic conditions where it is a monomer in solution (*27*), it is possible to identify the specific residues involved in the early stages of nanosphere assembly by following perturbations to their amide chemical shifts in the ^{1}H−^{15}N HSQC spectrum as a function of salt concentration. Changes in the chemical environment of the nuclei at interaction interfaces are often accompanied by perturbations, typically chemical shift and/or intensity reductions to the backbone $^{1}H^{N}$ and ^{15}N resonances (*28, 29*). In this instance, the major changes observed in the ^{1}H−^{15}N HSQC spectra were the disappearance of cross-peaks, an observation that indicates a change in dynamics at or near the affected residues into a regime of intermediate motion (microsecond to millisecond) that is unfavorable in the NMR time scale (*28, 30*).

EXPERIMENTAL PROCEDURES

All chemicals and enzymes were purchased from the Sigma Chemical Company (St. Louis, MI) except when indicated.

Protein Synthesis. Recombinant mouse amelogenin (*31*) containing a 12-residue, N-terminal histidine tag (MRGSH-HHHHHGS-) to assist protein purification (rp(H)M180) was expressed and purified from *Escherichia coli* BL21(DE3) cells using methods previously described and exchanged into the final NMR buffer (2% CD_3CO_2D and 5% D_2O/95% H_2O, pH 3.0) (*27*) used as the starting point for the titration studies.

Sodium Chloride and Calcium Chloride NMR Titrations. Stock solutions of 1 M NaCl and 1 M $CaCl_2$ were prepared in NMR buffer and the pH measured to ensure that there was no significant variation from the starting condition. Aliquots of the sodium or calcium chloride solutions were added directly to the rp(H)M180 samples (0.14 mM (210 μL) and 0.13 mM (260 μL), respectively), and following gentle agitation, a high resolution ^{1}H−^{15}N HSQC spectrum (20 °C) was immediately acquired at a ^{1}H resonance frequency of 800 MHz (NaCl) or 750 MHz ($CaCl_2$) on a Varian Inova spectrometer. Spectra were recorded in the absence of salt and at NaCl/protein molar ratios of 65, 330, 500, 670, 940, 1000, 1330, 1670, 2330, and 3000 to one and $CaCl_2$/protein molar ratios of 60, 120, 360, 560, 700, 925, 1150, 1300, and 1500 to one. At the end of the titration, due to dilution, the final concentration of rp(H)M180 was 0.10 mM (NaCl) and 0.11 mM ($CaCl_2$), the final concentration of NaCl was 300 mM, and $CaCl_2$ was 165 mM, and the final volumes were 300 (NaCl) and 310 ($CaCl_2$) μL. Prior to the titration, at the end of the titration, and at various titration points in between, an overall rotational correlation time (τ_c) for rp(H)M180 was rapidly estimated from backbone amide ^{15}N $T_{1\rho}/T_1$ ratios (*32*) measured using a modified ^{1}H−^{15}N HSQC experiment to record an ^{15}N-edited one-dimensional spectrum. Both ni and phase were set equal to 1, D1 = 1.5, relaxT set to 0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0, 1.5, and 2.0 s with T1 = y and relaxT = 0.01, 0.02, 0.03, 0.05, 0.07, 0.1, 0.13, 0.17, 0.21, and 0.25 s with T1rho = y. Only the intensity of the region containing the amide resonances (9.0−7.5 ppm, the side chain resonances were excluded) was integrated in the 1D ^{15}N-relaxation spectra using the Varian macro intav. This data was then analyzed with the Varian t1a and t2a macros and a τ_c value estimated from the ratios.

pH NMR Titrations. Two 0.14 mM samples of double-labeled rp(H)M180 were prepared by dissolving lyophilized protein into 5% D_2O/95% H_2O and adjusting the pH to 4.4 and 5.5 with small volumes of dilute deuterated acetic acid (no salt). Following the acquisition of ^{1}H−^{15}N HSQC spectra at a ^{1}H resonance frequency of 750 MHz, an overall rotational correlation time (τ_c) was estimated. Similar data were also collected on two samples of double-labeled rp(H)M180 dissolved in 2% deuterated acetic acid with the pH adjusted to 4.5 and 5.0 with 1 M NaOH.

NMR Data Processing. All NMR data was processed using Felix2007 (Felix NMR, Inc., San Diego, CA) software, and the ^{1}H, ^{13}C, and ^{15}N chemical shifts were referenced to DSS (DSS = 0 ppm) using indirect methods (*33*).

Dynamic Light Scattering. Approximately 3 mg/mL rp(H)M180 (Veritas, Rockville, MD) samples (0.14 mM) were prepared by directly dissolving the protein into eight different prefiltered buffers (0.02 μm): 2% acetic acid (pH 3.0), 50 mM $CaCl_2$ in 2% acetic acid (pH 3.0), 100 mM NaCl in 2% acetic acid (pH 3.0), 300 mM NaCl in 2% acetic acid (pH 3.0), 2% acetic acid adjusted to pH values of 4.0

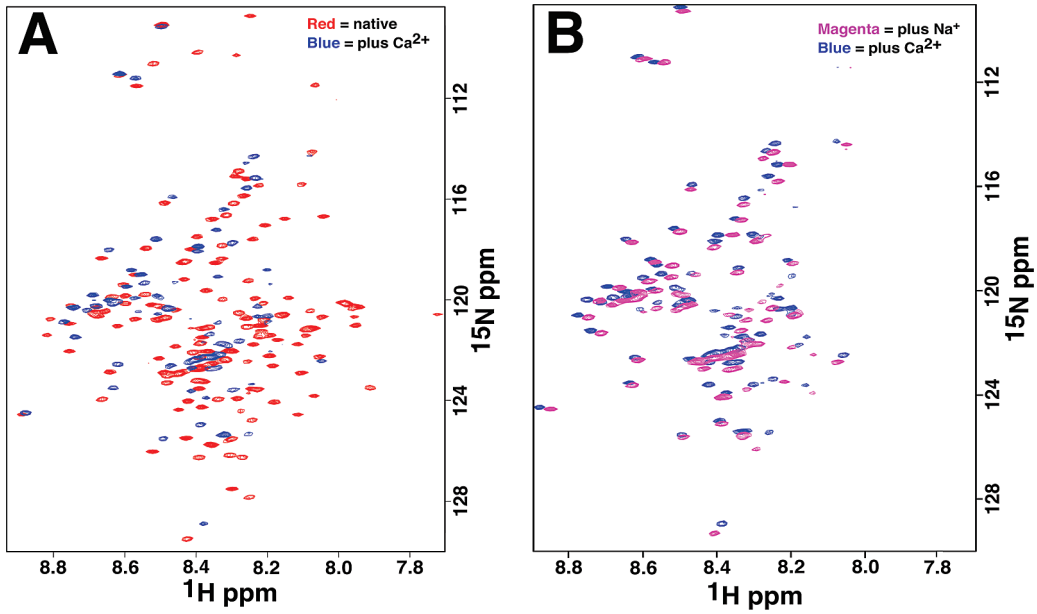

FIGURE 1: (A) Overlay of the $^{1}H-^{15}N$ HSQC spectra of rp(H)M180 in 2% acetic acid, pH 3.0, in the absence of salt (red) and in the presence of 1500:1 molar excess $CaCl_2$ (blue). (B) Overlay of the $^{1}H-^{15}N$ HSQC spectra of rp(H)M180 in the presence of 1500:1 (blue) and 3000:1 (magneta) molar excess $CaCl_2$ and NaCl, respectively. Spectra recorded at 20 °C at a ^{1}H resonance frequency of 750 ($CaCl_2$) and 800 (NaCl) MHz.

and 5.0 with 1 M NaOH, and protein solutions adjusted to pH 4.0 and 5.0 using low concentrations of acetic acid. Samples were transferred into disposable polystyrene cells and placed into the temperature controlled sample chamber (24 °C) of a Brookhaven Instruments 90 Plus (Brookhaven Instruments Corporation, Holtsville, NY) DLS instrument equipped with a 657 nm 35 mW laser. Time dependent fluctuations in the scattered intensity were measured using a Bl-APD digital correlator. Protein solutions were analyzed in triplicate using a 90° scattering angle at 24.0 °C with acquisition times of 1 to 3 min. Buffer solutions without protein were run through 0.02 μm filters and were also analyzed by DLS prior to dissolution of protein. Standard NIST traceable polystyrene 92 ± 3.7 nm latex standards and a blank, 0.02 μm filtered, deionized, ultrapure water (VWR International, West Chester, PA) were also run as standards. The autocorrelation functions were deconvoluted to obtain size distributions using both the non-negatively constrained least-squares fit (multiple pass NNLS) and the regularized LaPlace inversion (Contin) algorithms. The volume weighted size distributions obtained from the NNLS algorithms are presented.

RESULTS AND DISCUSSION

NMR Spectroscopy: Salt Titrations. Figure 1A is an overlay of the $^{1}H-^{15}N$ HSQC spectra of rp(H)M180 in the absence (red) of salt and in the presence of 1500:1 molar excess $CaCl_2$ (blue). The most significant observation is that a large number of the amide cross-peaks in the $^{1}H-^{15}N$ HSQC spectrum of rp(H)M180 disappear in the presence of an excess of $CaCl_2$. For the cross-peaks that do not disappear, most show some chemical shift perturbation, although these perturbations are not significant given the narrow chemical shift scale, approximately 1 ppm, in the proton dimension. Together, the disappearing and shifting amide cross-peaks suggest that the original structure of rp(H)M180 is being perturbed with the addition of $CaCl_2$. Similar behavior was observed with the titration of NaCl into rp(H)M180 as shown in Figure 1B, an overlay of the $^{1}H-^{15}N$ HSQC spectrum of rp(H)M180 at 3000:1 NaCl:rp(H)M180 (magenta) and 1500:1 $CaCl_2$:rp(H)M180 (blue). The $^{1}H-^{15}N$ HSQC spectra are very similar (and again, note that the differences are minor given the small scale of the proton dimension), suggesting that rp(H)M180 is being affected similarly by the presence of either salt at high molar ratios. The most likely explanation for the disappearance of these $^{1}H-^{15}N$ HSQC cross-peaks at high salt concentrations is restricted motion at a protein−protein interface.

The disappearing cross-peaks of amide resonances in the $^{1}H-^{15}N$ HSQC spectrum can be understood in relation to the time scale of motion observed by solution-state NMR (*28, 30*). Nuclei in molecules or regions of molecules that undergo rapid changes in their local chemical environment on the NMR time scale ($<\mu$s) generate observed resonances that represent the weighted average of all chemical environments sampled by the nuclei. Nuclei that undergo slow motion in their local chemical environment on the NMR time scale (>ms) generate observed resonances for each chemical environment sampled, and if there are many different chemical environments in slow exchange, it may not be possible to detect any of these resonances above the background noise. In between these two extremes is motion on the intermediate NMR time scale (ms to μs) where no resonances are observed because the resonances have broadened beyond detection (*30, 34*). At pH 3.0 in 2% acetic acid, one set of sharp, assignable resonances was observed in the $^{1}H-^{15}N$ HSQC spectrum for rp(H)M180 (*27*) indicating that the molecule was in fast exchange on the NMR time scale ($<\mu$s). Because of the narrow line shapes of the cross-peaks in the $^{1}H-^{15}N$ HSQC spectrum and the successful acquisition of all 3D backbone assignment data for protonated rp(H)M180 (poor magnetization transfer would be expected for a ~40 kDa dimer), rp(H)M180 was determined to be a monomer in solution under these conditions (*27*).

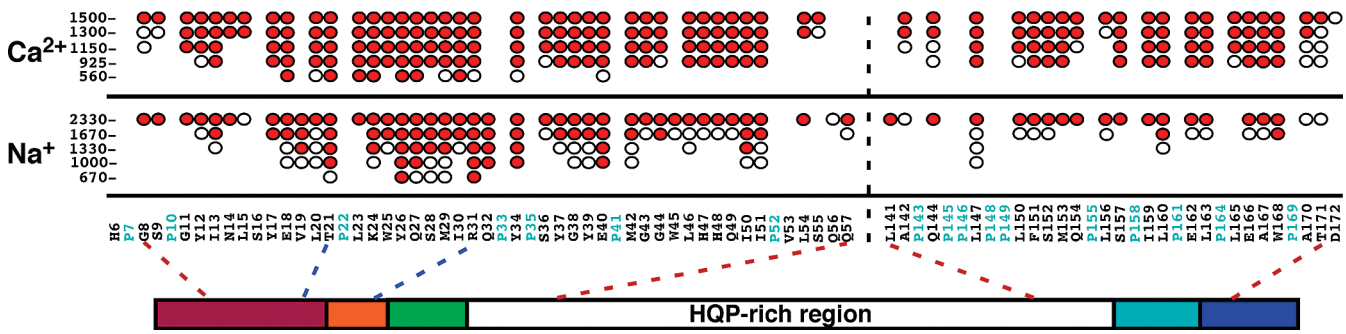

FIGURE 2: Plot of the amide resonances that start to disappear (open circles) or completely disappear (red circles) in the $^{1}H-^{15}N$ HSQC spectrum of ^{15}N-labeled rp(H)M180 at different salt/rp(H)M180 molar ratios (indicated on the left). Underneath is a schematic illustration of murine amelogenin with various regions of the protein highlighted (*26*). The first three N-terminal regions comprise the N-terminal TRAP (tyrosine-rich amelogenin peptide) region (magenta = protein−protein interaction region, green = lectin-like binding tri-tyrosine domain). The cyan colored region is a hydrophobic segment cleaved by enamelysin, and the blue colored region is the hydrophilic C-terminal mineral-binding domain. The dashed blue lines point to the region first affected by the addition of Ca^{2+}, and the dashed red lines are the regions affected at higher salt concentrations. Not shown on the plot are six residues that start to disappear (T63, F116, Q117, and H136) or have disappeared (A75 and Q83) at the highest sodium ion concentrations. Also not shown are two residues that disappear (T63 and M129) at the highest $CaCl_2$ concentration.

Such a conclusion is further supported by the absence of long-range NOEs (only sequential NOEs were observed) (*27*) and independently by SAXS data collected under similar solution conditions (*15*). As Na^+ or Ca^{2+} was added to rp(H)M180, amide resonances disappeared, and these disappearing amide resonances indicate that the molecular dynamics in these regions of the protein have changed, going from a state of fast motion to one of intermediate motion. There are many possible explanations for the observed change in dynamics, including restricted motion at a protein−protein or protein−ligand (Na^+ or Ca^{2+}) interface, an intermediate rate of chemical exchange between the free and associated states, or a salt induced change in the local monomer conformation. We will demonstrate that the explanation most consistent with our data is a protein−protein interaction. Regardless of the physical cause for the alteration in molecular dynamics, motion in the intermediate time scale range is often associated with regions involved in binding and catalysis in many proteins (*28, 30, 35*).

Figure 1 shows the end points of the NMR titrations with NaCl and $CaCl_2$. Because the $^{1}H-^{15}N$ HSQC spectrum of rp(H)M180 has been assigned in the absence of salt (*27*) and the chemical shift perturbations to the spectra following the addition of salt are small, it is possible to track the sequential disappearance of amide cross-peaks as a function of the molar ratio of salt present, and these assignments are shown in the Supporting Information and summarized in Figure 2.

With the titration of NaCl, only minor chemical shift perturbations are observed until a Na^+/protein molar ratio of 670:1 is reached and two amide cross-peaks disappear, and four are reduced in intensity for residues all located near the N-terminus between T21 and R31. Upon further addition of NaCl, additional cross-peaks disappear or begin to disappear around this N-terminal region. At a Na^+/protein molar ratio of 1670:1, a total of 23 cross-peaks disappear, and ten are reduced in intensity between Y12 and Q57. Accompanying these later changes near the N-terminus is the disappearance of cross-peaks for residues near the C-terminus. At a Na^+/protein molar ratio of 2330:1, the changes have reached a plateau with 39 out of 44 and 17 out of 21 cross-peaks absent between G8-Q57 and L141-T171, respectively. Not shown in Figure 2 are six residues in between these terminal regions where the cross-peaks disappear (A75, and Q83) or become reduced in intensity (T63, F116, Q117, and H136) at a Na^+/protein molar ratio of 2330:1. Also not shown in Figure 2 is the result for a Na^+/protein molar ratio of 3000:1 that is identical to 2330:1, e.g., the $^{1}H-^{15}N$ HSQC spectra were the same at both molar ratios except for an overall reduction in the intensity of all the cross-peaks at the higher salt ratio. In summary, adding increasing amounts of NaCl to rp(H)M180 resulted in the disappearance of amide cross-peaks primarily between two regions of rp(H)M180, G8-Q57 and L141-T171, regions near the N- and C-termini, respectively.

Figure 2 shows a similar progression of events with the titration of $CaCl_2$. Only minor chemical shift perturbations are observed until a Ca^{2+}/protein molar ratio of 560:1 is reached and seven amide cross-peaks disappear, and five are significantly reduced in intensity for residues all located near the N-terminus, between E18 and Y34. Upon further addition of calcium (925:1) additional cross-peaks for residues near both the N-terminus (27 completely and three partially) and the C-terminus (12 completely and five partially) disappear: regions Y12−I51 and Q144-T171, respectively. At the highest calcium concentration studied, a Ca^{2+}/protein molar ratio of 1500:1, there is a marked reduction in the intensity of all the $^{1}H-^{15}N$ cross-peaks and a further increase in the number of cross-peaks for residues near the C- (G8−S55) and N-termini (A142-D172) that disappear. Additionally, two residues (T63 and M129) between the N- and C-terminal regions now disappear completely (data not shown in Figure 2). In summary, adding increasing amounts of $CaCl_2$ to rp(H)M180 resulted in the disappearance of amide cross-peaks in essentially the same two N- and C-terminal regions of rp(H)M180 affected by high concentrations of NaCl, G8-S55, and A142-D172, respectively, suggesting that dimer self-assembly is also occurring at high $CaCl_2$ concentrations. Higher molar ratios of NaCl were necessary to produce similar changes in the $^{1}H-^{15}N$ spectra observed with $CaCl_2$, and this is likely related, at least partially, to the three-times greater ionic strength of $CaCl_2$ over NaCl (at equal concentrations). Multivalent metal ions have been observed to influence the secondary structure of the HQP-rich region of amelogenin (*36*), and hence, some of the differences in the details between the NaCl and $CaCl_2$ titrations may be ion

related (Na^+ versus Ca^{2+}) (*5*) and serve a functional purpose. It is interesting to note that with both NaCl and $CaCl_2$, relatively high ionic strengths in the NMR tube (300 and 495 mM, respectively, at the final titration point) were required to observe the spectral changes in both the N- and C-termini, and this is consistent with the high ionic strength of 165 mM measured in the enamel fluid from developing pig teeth (*37*).

Overall Rotational Correlation Time (τ_c) Estimation. To further characterize the [NaCl] and [$CaCl_2$] dependent intensity perturbations observed in the $^1H-^{15}N$ spectra, an overall τ_c for rp(H)M180 was estimated from backbone amide ^{15}N $T_{1\rho}/T_1$ ratios (*32*) in the absence and presence of salt. For similarly shaped proteins, τ_c is linearly related to the molecular weight (*38*), and consequently, estimations of τ_c have been used to identify protein dimerization (*39*). With rp(H)M180, measuring these values would also help determine if the disappearing cross-peak in the $^1H-^{15}N$ spectra were due to protein−protein or protein−ion interactions. We assume no major change in the structure of monomeric amelogenin because of the addition of salt, which is a reasonable assumption given that the chemical shift dispersion in the ^1H and ^{15}N dimension does not increase in the presence of salt. Ion binding ($Na^+ = 23$, $Ca^{2+} = 43$, and $Cl^- = 35$ Da) to amelogenin would not significantly alter the molecular weight of rp(H)M180 (21,600 Da) while protein−protein association would. For rp(H)M180 at pH 3.0, in the monomeric, salt-free condition, a τ_c of 4.6 ± 0.2 ns was observed. At a Na^+/rp(H)M180 molar ratio of 2330:1 the value of τ_c increased to 6.4 ± 0.2 ns. Upon the addition of calcium to a Ca^{2+}/protein molar ratio of 700:1, τ_c increased to 6.0 ± 0.4 ns and remained near that value up to a Ca^{2+}/protein molar ratio of 1500:1. Hence, at high concentrations both NaCl and $CaCl_2$ manifested a similar increase in τ_c suggesting that similar complexes had formed. The magnitude of the τ_c increase (50%), assuming no major change in the structure of rp(H)M180, is more likely due to protein−protein rather than protein−ion interactions. Note that our estimation of τ_c for monomeric rp(H)M180 (4.6 ± 0.2 ns) is smaller than the τ_c determined for similarly sized globular proteins (~15 ± 3 ns) (*39*). Because τ_c is related to the shape and packing configuration of the molecule (*38, 40*), rp(H)M180 must have a shape markedly distinct from a spherical, globular protein as suggested by others (*5, 14, 15*).

Dynamic Light Scattering. Because NMR-based τ_c estimations and size exclusion chromatography experiments (see Supporting Information) failed to conclusively identify protein−protein interactions, dynamic light scattering experiments were performed on rp(H)M180 in 2% acetic acid under various solution conditions at approximately the same protein concentration used for the NMR experiments (0.14 mM) to estimate the size of the amelogenin aggregates. Such experiments provide a direct measure of particle size and size distributions in solution and have been extensively used to study amelogenin in solution (*5, 17, 18, 41, 42*). The results of the DLS experiments are shown in Figure 3 and summarized in Table 1. In the absence of additional salt in 2% acetic acid at pH 3.0, a condition where NMR spectroscopy shows amelogenin is a monomer in solution, the particle diameter was measured to be 4.1 nm, a value that agrees well with the size determined by others for amelogenin monomers in solution (*14, 18, 19*). Under the same conditions

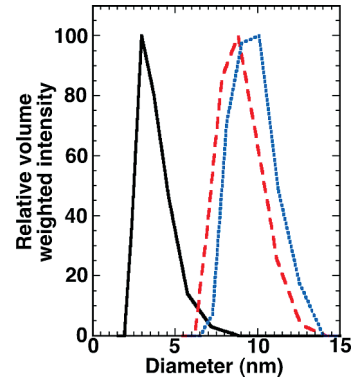

FIGURE 3: Dynamic light scattering data for rp(H)M180 (0.14 mM) prepared in 2% acetic acid (pH 3.0) with no salt (black solid line), 100 mM NaCl (red dashed line), and 50 mM $CaCl_2$ (blue dotted line). The increase in diameter with either salt is consistent with the self-association of amelogenin into larger complexes.

Table 1: Diameter of rp(H)M180 Determined by Dynamic Light Scattering as a Function of Buffer Condition[a]

sample conditions	smallest observed diameter (nm)
no salt, pH 3.0	4.1
50 mM CaCl2, pH 3.0	10.1
100 mM NaCl, pH 3.0	9.6
pH 4.5 adjusted with NaOH	11.0
pH 5.0 adjusted with NaOH	14.0
no salt, pH 4.5	4.5
no salt, pH 5.5	4.3

[a] All samples contained 2% acetic acid except those at pH 4.5 and 5.5 with no salt where rp(H)M180 was resuspended in water and the pH adjusted only with dilute acetic acid.

Table 2: Summary of Estimated τ_c Values Calculated for rp(H)M180 as a Function of Buffer Condition[a]

sample condition	salt/rp(H)M180	estimated τ_c (ns)
no salt, pH 3.0		4.6±0.2
no salt, pH 4.4		4.1±0.6
no salt, pH 5.5		4.6±0.4
pH 4.5, adjusted with NaOH		5.6±0.2
pH 5.0, adjusted with NaOH		6.2±0.1
$CaCl_2$, pH 3.0	700/1	6.0±0.4
$CaCl_2$, pH 3.0	1150/1	6.2±0.4
$CaCl_2$, pH 3.0	1500/1	6.2±0.5
NaCl, pH 3.0	2330/1	6.4±0.2

[a] All samples contained 2% acetic acid except the solutions at pH 4.4 and 5.5 with no salt, which just contained a small amount of acetic acid (≪2%).

in the presence of 50 mM $CaCl_2$ or 100 mM NaCl, conditions where NMR spectroscopy indicates amelogenin has associated into dimers, the diameter of the smallest components more than doubles to a diameter of ~10 nm. The large increase in physical size of amelogenin due to the presence of salt clearly cannot be due to protein−ion associations alone and strongly suggests the formation of dimers or larger complexes in solution.

NMR Spectroscopy: pH Titrations. Solution pH is a well-documented initiator of amelogenin nanosphere formation (*5*) with optimal values between 6 and 8. By adjusting the pH of the 2% acetic acid solutions with 1 M NaOH, the NMR data suggested that we were witnessing nanosphere self-assembly affected by an increase in pH as the value of τ_c increased from 4.6 ± 0.2 ns to 6.2 ± 0.1 ns in going from pH 3.0 to 5.0 (Table 2). Furthermore, the $^1H-^{15}N$ HSQC spectra displayed features observed by increasing the

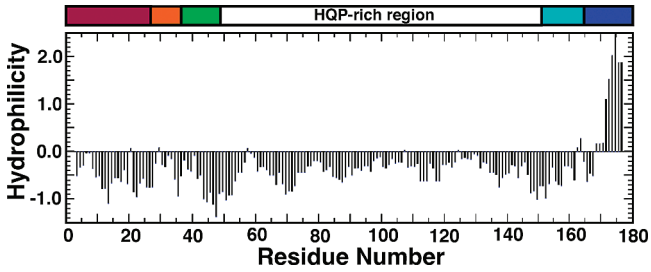

FIGURE 4: Plot of the hydrophilicity score for rp(H)M180 determined by the method of Hopp and Woods (*52*) using ProtScale software of ExPASy (Expert Protein Analysis System). Positive values are hydrophilic and negative values hydrophobic. On top is a schematic illustration of murine amelogenin with various regions highlighted according to the description in Figure 2.

salt concentration, intensity reduction, and disappearance of cross-peaks (data not shown). The DLS data on these samples corroborated amelogenin self-assembly as the smallest observed diameter rose from 4.1 nm at pH 3.0 to 14.0 nm at pH 5.0 (Table 1). However, it soon became apparent that the changes observed in the NMR data with increasing pH were due to an increase in the Na^+ concentration that occurred when adjusting the pH with 1 M NaOH. When the experiments were repeated with solutions prepared by adjusting the pH only with acetic acid (in the absence of salt), there was no experimentally significant change in τ_c nor was there an experimentally significant change in the smallest observed diameter detected by DLS (Table 1). Furthermore, it was possible to keep enough rp(H)M180 in solution ($>\sim$0.05 mM) to record an $^1H-^{15}N$ HSQC spectrum up to a pH of 5.5 when adjusting the pH with acetic acid, while pH 5.0 was the highest the pH could be raised and still record a $^1H-^{15}N$ HSQC spectrum using 1 M NaOH to adjust the pH. These observations corroborate with those made by Tan et al. showing the solubility at pH 5.0 decreased as the ionic strength was increased (*43, 44*). Hence, solution conditions must be rigorously taken into account when studying amelogenin self-assembly *in vitro* using NMR concentrations (greater than 0.1 mM). The solution parameters to allow us to study details of the initial events of nanosphere self-assembly due to changes in pH using NMR spectroscopy are still under investigation.

Biological Implications. Intra- and intermolecular association of hydrophophic regions in proteins is believed to be one of the driving forces responsible for protein folding and protein association/aggregation (*45*), and it has been suggested to play a functional role in the self-assembly of amelogenin as well (*18, 19*). This is because there is significant sequence conservation of amelogenin among different species and the amelogenin sequence is rich in hydrophobic residues (*10*). Indeed, the hydrophobic environment within nanospheres is believed to play a role in the latter stages of calcium hydroxyapatite crystal growth (*26, 46*). Increasing the ionic strength of the solution generally increases the tendency for the hydrophobic regions to interact (*47*). The hydrophilicity plot for the primary sequence of rp(H)M180, shown in Figure 4, indicates that the majority of the protein is hydrophobic except for the hydrophilic C-terminus. However, while there is some periodicity to the degree of hydrophobicity, the N-terminal region clearly does not stand out as being the most hydrophobic region of the protein, and hence, other factors must be involved in the

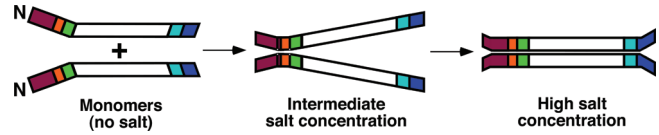

FIGURE 5: Schematic showing the proposed stepwise interaction mechanism of amelogenin self-assembly into dimers. The monomers (left) first interact with each other near the N-terminus at intermediate salt concentrations (middle). At high salt concentrations (right), a larger region near the N-terminus and a region near the C-terminus interact. The monomers are drawn as elongated rods because the general consensus is that in the monomeric form much of the protein has the predisposition to form an extended β-sheet/β-strand and β-spiral/polyproline II (PPII) structure (*5, 14, 15*).

N-terminal interactions that initiate self-assembly with increasing ionic strength. It is perhaps not surprising that interactions involving the C-terminus requires higher ionic strength, and occur following the events at the N-terminus, given its hydrophilic nature.

Ionic strength has been observed to play a role in nanosphere size and heterogeneity (*17*), as well as the solubility (*44*) of amelogenin. The NMR and DLS data presented here directly address the role of ionic strength in nanosphere assembly and suggest that high salt affects a transition from a monomer to a dimer (and larger) state. Such an interpretation is in agreement with the current amelogenin nanosphere assembly paradigm where dimers and trimers form first followed by the assembly of hexamers (*19*). In turn, these three building blocks congregate in a progressive manner to form nanospheres containing 20−100 amelogenin molecules. If our interpretation of the NMR and DLS data is correct, then our data suggests that dimer assembly is first initiated near the N-terminal region (E18-Y34) followed by the concurrent involvement of residues near both the N- (G8-Q57) and the C-terminal (L141-T171) regions as shown schematically in Figure 5. It is unlikely that rp(H)M180 forms a dimer via antiparallel alignment as such an event should result in the disappearance of amide resonances near both the N- and C-termini simultaneously, and this is clearly not the case as illustrated in Figure 2. Instead, dimerization must be occurring first with head-to-head associations near the N-terminus followed by tail-to-tail associations near the C-terminus. In this model, the extreme N- and C-termini, e.g., those residues which have been identified not to be involved in the initial self-assembly process, hang out from the complex allowing increased motional freedom. The C-terminus is believed to be important for initially anchoring the nanosphere to apatite, and there is recent solid state NMR evidence supporting such a species (*48*−*50*). Note that the residues in regions S8-Q57 and L141-T121 that appear to be involved with the salt induced changes in molecular dynamics reported here are in good agreement with previous studies demonstrating that residues within regions M1-M42 and S157−K173 participate in protein−protein interactions in nanosphere self-assembly (*20, 23*). Furthermore, chromosomal-linked amelogenesis imperfecta is a genetic disease where point mutations of T21→I and P41→T result in the development of defective enamel (*51*), suggesting that a region near the N-terminus is important for proper enamel formation. Given the suspected importance of the N-terminus in nanosphere self-assembly and enamel formation, the initial events we observe at the N-terminus as a function of increasing NaCl and $CaCl_2$ concentrations may correspond to the initial events involved in the process of nanosphere self-assembly.

SUPPORTING INFORMATION AVAILABLE

Two figures showing the tentative assignments of the $^{1}H-^{15}N$ HSQC spectra of rp(H)M180 at a salt:rp(H)M180 molar ratio of 1500:1 ($CaCl_2$) and 3000:1 (NaCl); a description of the experimental methods and results of size exclusion chromatography experiments with rp(H)M180 using elution buffers with and without various $CaCl_2$ or NaCl. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Ten Cate, A. R. (1994) *Oral histology: Developement, Structure, and Function*, 4th ed., Mosby, St. Louis, MO.
2. Landis, W. J. (1995) The strength of calcified tissue depends in part on the molecular structure and organization of its constituent mineral crystals in their organix matrix. *Bone 16*, 533–544.
3. White, S. N., Luo, W., Paine, M. L., Fong, H., Sarikaya, M., and Snead, M. L. (2001) Biological organization of hydroxyapatite crystallites into a fibrous continuum toughens and controls anisotrophy in human enamel. *J. Dent. Res. 80*, 321–326.
4. Hunter, G. (1996) Interfacial aspects of biomineralization. *Curr. Opin. Mat. Sci. 1*, 430–435.
5. Margolis, H. C., Beniash, E., and Fowler, C. E. (2006) Role of macromoleculare assembly of enamel matrix proteins in enamel formation. *Crit. Rev. Oral Biol. Med. 85*, 775–793.
6. Termine, J. D., Belcourt, A. B., Christner, P. J., Conn, K. M., and Nylen, M. U. (1990) Properties of dissociatively extracted fetal tooth matrix proteins. *J. Biol. Chem. 255*, 9760–9768.
7. Fincham, A. G., Moradian-Oldak, J., Simmer, J. P., Sarte, P., Lau, E. C., Diekwisch, T., and Slavkin, H. C. (1994) Self-assembly of a recombinant amelogenin protein generates supramolecular structures. *J. Struct. Biol. 112*, 103–109.
8. Fincham, A. G., Moradian-Oldak, J., Diekwisch, T. G. H., Lyaruu, D. M., Wright, J. T., Jr., P. B., and Slavkin, H. C. (1995) Evidence for amelogenein "nanospheres" as functional components of secretory-stage enamel matrix. *J. Struct. Biol. 115*, 50–59.
9. Uchida, T., Tanabe, T., Fukae, M., Shimizu, M., Yamada, M., Miake, K., and Kobayashi, S. (1991) Immunochemical and immunohistochemical studies, using antisera against porcine 25 kDa amelogenin, 89 kDa enamelin and the 12−17 kDa nonamelogenins, on immature enamel of the pig and rat. *Histochemistry 96*, 129–138.
10. Toyosawa, S., O'hUigin, F., Figueroa, F., Tichy, H., and Klein, J. (1998) Identification and characterization of amelogenin genes in monotremes, reptiles, and amphibians. *Proc. Natl. Acad. Sci. U.S.A. 95*, 13056–13061.
11. Snead, M. L., Lau, E. C., Zeichner-David, M., Fincham, A. G., Woo, S. L., and Slavkin, H. C. (1986) DNA sequence for cloned cDNA for murine amelogenin reveals the amino acid sequence for enamel-specific proteins. *Biochem. Biophys. Res. Commum. 129*, 812–818.
12. Renugopalakrishnan, V., Strawich, E. S., Horowitz, P. M., and Glimcher, M. J. (1986) Studies of the secondary structure of amelogenin from bovin tooth enamel. *Biochemistry 25*, 4879–4887.
13. Goto, Y., Kogure, E., Takagi, T., Aimoto, S., and Aoba, T. (1993) Molecular conformation of porcine amelogenin in solution: three folding units at the N-terminal, central, and C-termininal regions. *J. Biochem. (Toyko) 113*, 55–60.
14. Lakshminarayanan, R., Fan, D., Du, C., and Moradian-Oldak, J. (2007) The role of secondary structure in the entropically driven amelogenin self-assembly. *Biophys. J. 93*, 3664–3674.
15. Matsushima, N., Izumi, Y., and Aoba, T. (1998) Small-angle X-ray scattering and computer-aided molecular modeling studies of 20 kDa fragment of porcine amelogenin: does amelogenin adopt an elongated bundle structure? *J. Biochem. 123*, 150–156.
16. Robinson, S., Fuchs, P., and Weatherell, J. A. (1981) The appearance of developing rat incisor enamel using a freeze fracturing technique. *J. Cryst. Growth 53*, 160–165.
17. Moradian-Oldak, J., Simmer, J. P., Lau, E. C., Sarte, P. E., Slavkin, H. C., and Fincham, A. G. (1994) Detection of monodisperse aggregates of a recombinant amelogenin by dynamic light scattering. *Biopolymers 34*, 1339–1347.
18. Moradian-Oldak, J., Leung, W., and Fincham, A. G. (1998) Temperature and pH-dependence of amelogin self-assembly: A particle size distribution study. *J. Struct. Biol. 122*, 320–327.
19. Du, C., Falini, G., Fermani, S., Abbott, C., and Moradian-Oldak, J. (2005) Supramolecular assembly of amelogenin nanospheres into birefringent microribbons. *Science 307*, 1450–1454.
20. Moradian-Oldak, J., Paine, M. L., Lei, Y. P., Fincham, A. G., and Snead, M. L. (2000) Self-assembly properties of recombinant engineered amelogenin proteins analyzed by dynamic light scattering and atomic force microscopy. *J. Struct. Biol. 131*, 27–37.
21. Paine, M. L., White, S. N., Luo, W., Fong, H., Sarikaya, M., and Snead, M. L. (2001) Regulated gene expression dictates enamel structure and tooth function. *Matrix Biol. 20*, 273–292.
22. Moradian-Oldak, J., Jiminez, I., Maltby, D., and Fincham, A. G. (2001) Controlled proteolysis of amelogenins reveals exposure of both carboxy- and amino terminal regions. *Biopolymers 58*, 606–616.
23. Paine, M. L., and Snead, M. L. (1997) Protein interactions during assembly of the enamel organic extracellular matrix. *J. Bone Min. Res. 12*, 221–227.
24. Wen, H. B., Fincham, A. G., and Moradian-Oldak, J. (2001) Progressive accretion of amelgenin molecules during nanosphere assembly revealed by atomic force microscopy. *Matrix Biol. 20*, 387–395.
25. Dempsey, C. E. (1990) The actions of melittin on membranes. *Biochim. Biophys. Acta 1031*, 143–161.
26. Moradian-Oldak, J. (2001) Amelogenins: assembly, processing and control of crystal morphology. *Matrix Biol. 20*, 293–305.
27. Buchko, G. W., Bekhazi, J., Cort, J. R., Valentine, N. B., Snead, M. L., and Shaw, W. J. (2008) ^{1}H, ^{13}C and ^{15}N Resonance assignments of murine amelogenin, an enamel biomineralization protein. *Biomol. NMR Ass. 2*, 89–91.
28. Buchko, G. W., McAteer, K., Wallace, S. S., and Kennedy, M. A. (2005) Solution-state NMR investigation of DNA binding interactions in *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg): a dynamic description of the DNA/protein interface. *DNA Repair 4*, 327–339.
29. Zuiderweg, E. R. P. (2002) Mapping protein-protein interactions in solution using NMR spectroscopy. *Biochemistry 41*, 1–7.
30. Ishima, R., and Torchia, D. A. (2000) Protein dynamics from NMR. *Nat. Struct. Biol. 7*, 740–743.
31. Simmer, J. P., Lau, E. C., Hu, C. C., Bringas, P., Santos, V., Aoba, T., Lacey, M., Nelson, D., Zeichner-David, M., Snead, M. L., Slavkin, H. C., and Fincham, A. G. (1994) Isolation and characterization of a mouse amelogenin expressed in *Escherichia coli*. *Calcif. Tissue Int. 54*, 312–319.
32. Szyperski, T., Yeh, D., Sukumaran, D., Moseley, H., and Montelione, G. (2002) Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment. *Proc. Natl. Acad. Sci. U.S.A. 99*, 8009–8014.
33. 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.
34. Lee, D., Hilty, C., Wider, G., and Wuthrich, K. (2006) Effective rotational correlation times of proteins from NMR relaxation interference. *J. Magn. Reson. 178*, 72–76.
35. Mulder, F. A. A., Hon, G., Muhandiram, D. R., Dahlquist, F. W., and Kay, L. E. (2000) Flexibility and ligand exchange in a buried cavity mutant of T4 lysozyme studied by multinuclear NMR. *Biochemistry 39*, 12614–12622.
36. Sogah, D. Y., Perle-Treves, D., Voyer, N., and DeGrado, W. F. (1994) Design and synthesis of polytripeptide (LeuGlnPro) based upon the matrix protein amelogenin. *Macromol. Symp. 88*, 149–163.
37. Aoba, T., and Moreno, E. C. (1987) The enamel fluid in the early secretory stage of porcine amelogenesis: chemical composition and saturation with respect to enamel mineral. *Calcif. Tissue Int. 41*, 86–94.
38. Krishnan, V. V., and Cosman, M. (1998) An emperical relationship between rotational correlation time and solvent accesible area. *J. Biomol. NMR 12*, 177–182.
39. Bhattacharjya, S., Ping, X., Gingras, R., Shaykhutdinov, R., Wu, C., Whiteway, M., and Ni, F. (2004) Solution structure of the dimeric SAM domain of MAPKKK Ste11 and its interactions with the adaptor protein STE50 from the budding yeast: implications for Ste11 activitation and signal transmission through the Ste50-Ste11 complex. *J. Mol. Biol. 344*, 1071–1087.
40. Squire, P. G., and Himmel, M. E. (1979) Hydrodynamics and protein hydration. *Arch. Biochem. Biophys. 196*, 165–177.

41. Fincham, A. G., Leung, W., Tan, J., and Moradian-Oldak, J. (1998) Does amelogenin nanosphere assembly proceed through intermediary-sized structures? *Conn. Tis. Res. 38*, 237–240.
42. Aichmayer, B., Margolis, H. C., Sigel, R., Yamakoshi, Y., Simmer, J. P., and Fratzl, P. (2005) The onset of amelogenin nanosphere aggregation studied by small-angle x-ray scattering and dynamic light scattering. *J. Struct. Biol. 151*, 239–249.
43. Tan, J., Leung, W., Moradian-Oldak, J., Zeichner-David, M., and Fincham, A. G. (1998) Quantitative analysis of amelogenin solubility. *J. Dent. Res. 77*, 1388–1396.
44. Tan, J., Leung, W., Moradian-Oldak, J., Zeichner-David, M., and Fincham, A. G. (1998) The pH dependent amelogenin solubility and its biological signficance. *Conn. Tis. Res. 38*, 215–221.
45. Miller, S., Janin, J., Lesk, A. M., and Chothia, C. (1987) Interior and surface of monomeric proteins. *J. Mol. Biol. 196*, 641–656.
46. Habelitz, S., Kullar, A., Marshall, S. J., DenBesten, P. K., Balooch, M., Marshall, G. W., and Li, W. (2004) Amelogenin-guided crystal growth on fluoroapatite glass-ceramics. *J. Dent. Res. 83*, 698–702.
47. Kohn, W. D., Kay, C. M., and Hodges, R. S. (1997) Salt effects on protein stability: two-stranded α-helical coiled-coils containing inter- or intrahelical ion pairs. *J. Mol. Biol. 267*, 1039–1052.
48. Shaw, W. J., Campbell, A. A., Paine, M. L., and Snead, M. L. (2004) The COOH terminus of the amelogenin, LRAP, is oriented next to the hydroxyapatite surface. *J. Biol. Chem. 279*, 40263–40266.
49. Shaw, W. J., Ferris, K., Tarasevich, B. J., and Larson, J. L. (2008) The structure and orientation of the C-terminus of LRAP. *Biophys. J. 94*, 3247–3257.
50. Shaw, W. J., and Ferris, K. (2008) Structure, orientation and dynamics of the C-terminal hexapeptide of LRAP determined using solid state NMR, *J. Phys. Chem. B*, in press.
51. Ravassipour, D., Hart, P. S., Hart, T. C., Ritter, A. V., Yamauchi, M., Gibson, C., and Wright, J. T. (2000) Unique enamel phenotype associated with amelogenin gene (AMELX) codon 41 point mutation. *J. Dent. Res. 79*, 1476–1481.
52. Hopp, T. P., and Woods, K. R. (1981) Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U.S.A. 75*, 3824–3828.

BI8018288