**Supporting Online Material**

# Cementomimetics – Constructing a cementum-like biomineralized microlayer *via* amelogenin-derived peptides

Mustafa Gungormus, Ersin E. Oren, Candan Tamerler, Hanson Fong, Jeremy A. Horst, Marketa Hnilova, Martha J. Somerman, Ram Samudrala, Malcolm L. Snead, and Mehmet Sarikaya

## Combinatorial selection of peptides

Commercially available Ph.D. Phage Display Peptide Library Kits (c7c and 12) were used as the biocombinatorial library (New England BioLabs Inc., USA) and synthetic HA powder was used as the target substrate (a generous gift from Ivan Transevica Institute, Department of Materials Science, Ukraine). Prior to exposing to the peptide - phage library, the HA powder was cleaned by ultrasonication in a 1/1 methanol/acetone mixture followed by ultrasonication in isopropanol. Cleaned HA powder was incubated with the phage-peptide library overnight in a potassium phosphate/sodium carbonate (PC) buffer (pH 7.4), containing 0.1% detergent (Tween 20 and Tween 80, Merck, USA) at room temperature with constant rotation. The detergent prevents non-specific interactions between phage particles and enables individual phage particle to powder surface interaction to occur. After incubating the HAp powder substrate with the phage library, the powder was washed several times with PC buffer containing 0.1% detergent to remove any non-specifically or weakly bound phage particles from the HA surface. The phages remaining on the HA surface were then eluted by 0.2 M Glycine HCl (pH 2.2) solution, followed by ultrasonication. Eluted phage were transferred to an early-log phase *E.coli* ER2738 culture and amplified for 4 hours. The amplified phage are isolated by polyethylene glycol (PEG) precipitation. After the third round of panning, the phage obtained from each round of selection are plated on LB-Agar media containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Xgal) and isopropyl-β-D-thiogalactopyranosid (IPTG) in serial dilutions to obtain single phage plaques. Since ER2738 strain lacks the *lacZ*$\alpha$ gene, only the cells infected by M13mp 19 bacteriophage, carrying the *lacZ*$\alpha$ gene can produce β– galactosidase and provide a chromogenic change in color when exposed to Xgal. Therefore, infected cells can hydrolyze X-Gal and form blue phage plaques. After growing the phages on LB-Agar plates, single blue plaques containing clonal expanded phage particles were picked, amplified and the amino acid sequence of the corresponding polypeptide segment was identified by DNA sequencing. These steps are illustrated in figure S1.

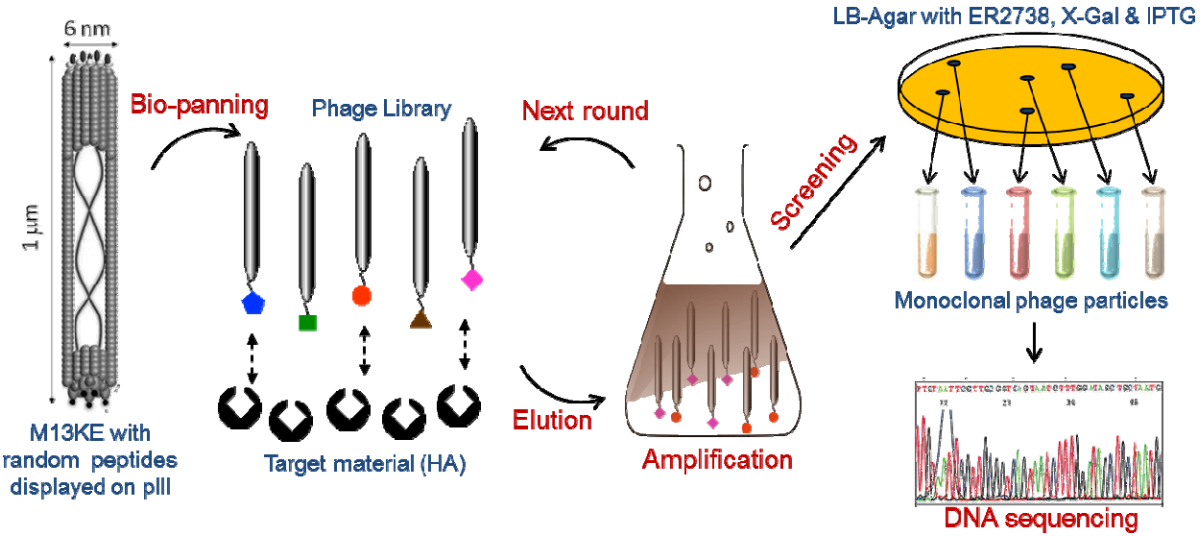


**Figure S1**: Schematic illustration of the selection process of the HA-binding peptides using phage display method.

**Semi-quantitative analysis of HA binding**

The binding affinity of each clone was assessed by immunofluorescence fluorescence microscopy and enzyme-linked immunosorbent assay (ELISA). For the fluorescence microscopy, cleaned HAp powder was incubated overnight with selected monoclonal phages at $10^{11}$ PFU/mL concentration in PC buffer containing 0.1% detergent at room temperature and with constant rotation. Wild type M13mp19 bacteriophage with no random peptides displayed on the pIII protein was used as negative control. After incubation, the powder was washed three times with the same buffer. After the washing, the phage remaining bound to the surface are labeled with mouse anti-M13 IgG (Amersham Biosciences, USA) and Alexa Fluor 488-anti-mouse IgG2a conjugate (Invitrogen Corporation, USA). After labeling, the powder suspension was transferred to a microscope slide and observed using a TE 300L microscope (Nikon, Japan). The approximate surface coverage of the phage particles was calculated using MetaMorph Imaging System Ver. 6.2 (Photometrics UK Ltd., UK, formerly Universal Imaging Co., USA) by comparing the approximate surface area of the powder in the bright field image to the approximate phage coverage in the fluorescence image.

For ELISA analysis, cleaned HAp powder was incubated overnight with selected monoclonal phages at $10^{11}$ PFU/ml concentration in PBS buffer containing 2% BSA at room temperature with constant rotation. Wild type M13mp 19 bacteriophage, displaying no peptides on the p(III) protein, was used as negative control. The powder was washed 3 times with PBS buffer containing 0.05% detergent. After the washing, the bound phages on the surface were labeled with horseradish peroxidase (HRP)-anti-M13 IgG conjugates (Amersham Biosciences, USA) in PBS buffer containing 2% BSA for 30 minutes with constant rotation and washed 4 times with PBS buffer. The reaction ready HRP substrate 3, 3´, 5 , 5´ tetramethylbenzidine (TMB)-ELISA (Pierce, USA) was added into the solution. The oxidation of the TMB by HRP yields a diimine form of TMB. The reaction was stopped after 2 minutes by adding 1 M $H_2SO_4$, which results in a yellow color with maximum absorbance at 450nm. The optical density (OD) of the reaction solution was measured at 450 nm wave length using a microplate reader (Tecan Trading AG, Switzerland).

**Table S1:** Number of the characterized peptides from the 7 amino acids cyclic (c7c) and 12 amino acids linear phage-peptide libraries.

| | **Strong** | **Moderate** | **Weak** | **Total** |
|---|---|---|---|---|
| **PhD c7c** | 12 | 27 | 18 | 57 |
| **PhD 12** | 16 | 20 | 13 | 49 |

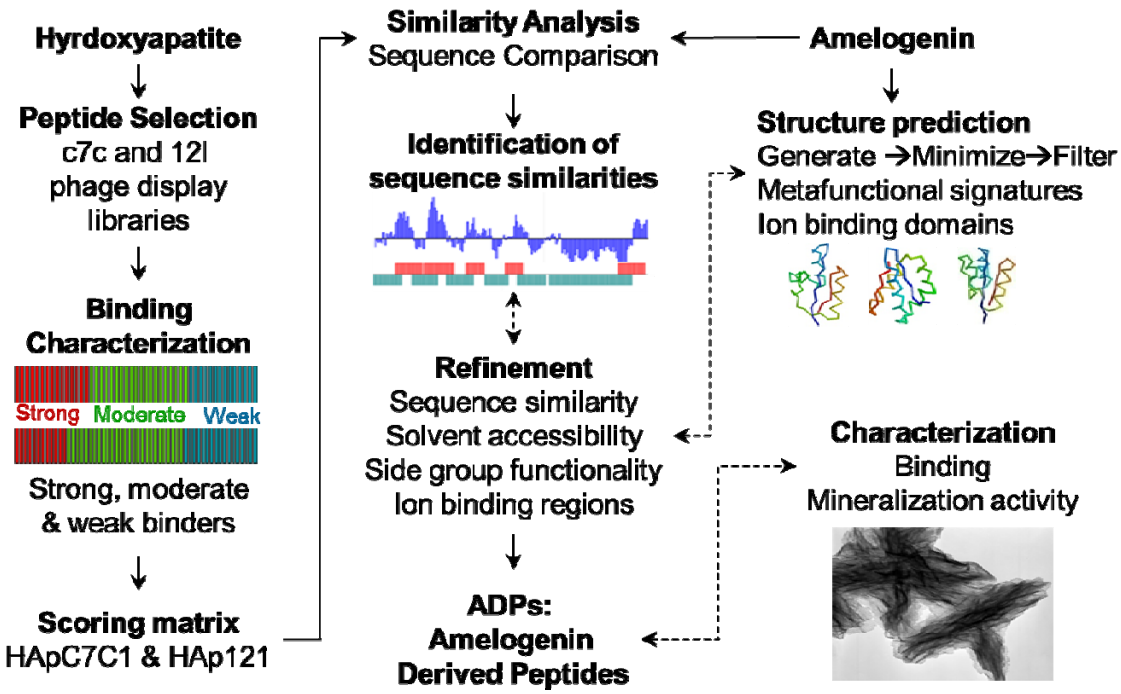


**Figure S2:** Procedure for designing amelogenin derived peptides

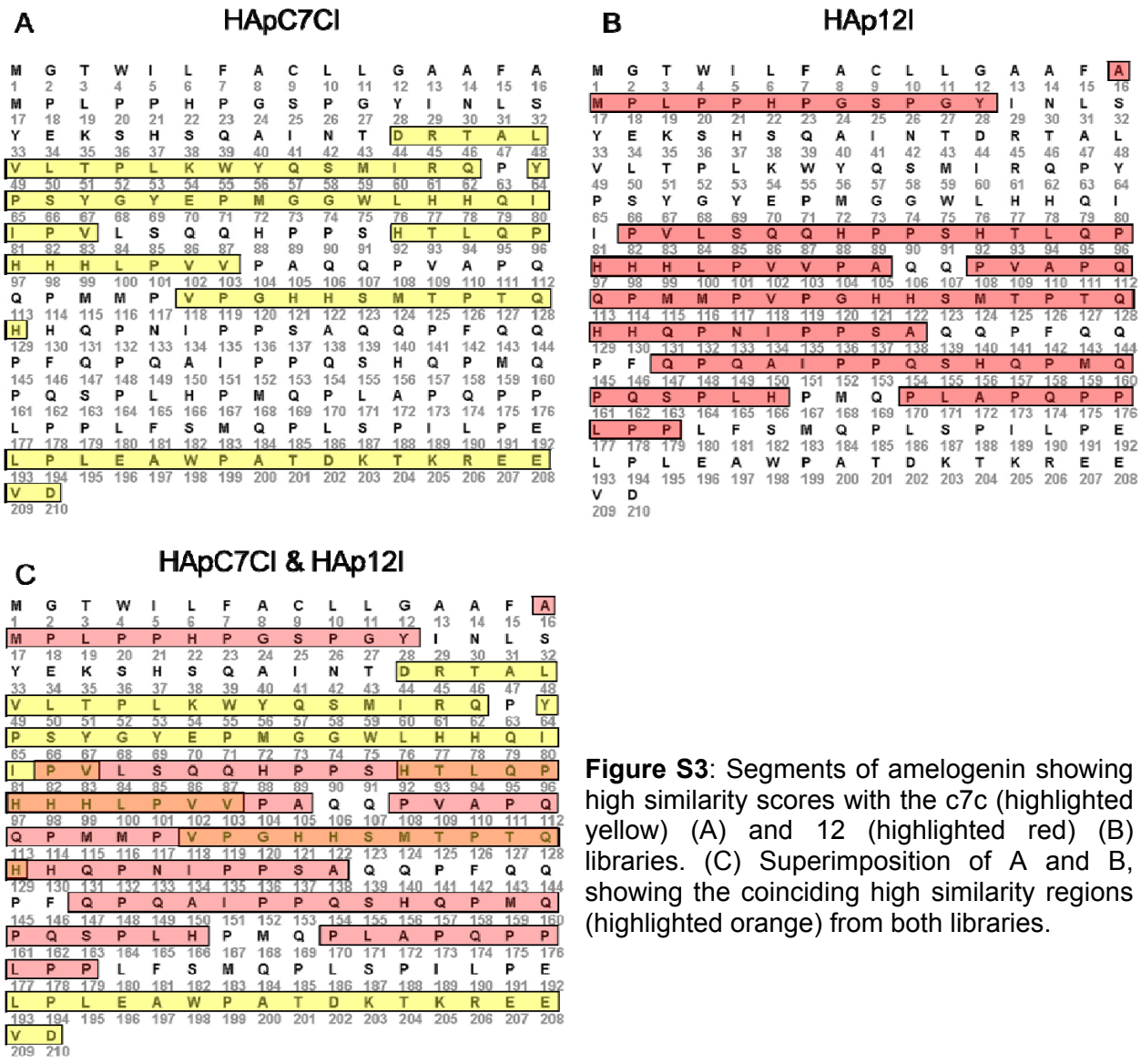


**Figure S3**: Segments of amelogenin showing high similarity scores with the c7c (highlighted yellow) (A) and 12 (highlighted red) (B) libraries. (C) Superimposition of A and B, showing the coinciding high similarity regions (highlighted orange) from both libraries.

**Protein Structure Prediction**

Contemporary protein structure prediction methods use known structures as variable length templates for construction assembly. We have developed an initial set of incomplete structure alignments in the form of meta-predictions using 3D-Jury (<http://bioinfo.pl/meta>) (1). and incorporated these as templates for the development of many full length models using the Protinfo protein structure prediction protocol <http://protinfo.compbio.washington.edu/protinfo_abcmfr> (2) (RAMP: A suite of programs to aid in the modeling of protein structure and function <http://compbio.- washington.edu/ramp/>) (3). In addition we have retrieved the full-length models from the I-TASSER server <http://zhang.bioinformatics.ku.edu/I-TASSER> (4). Next, we have used an all-atom conditional probability discriminatory function, RAPDF (5), to select the best models. These models were

iteratively refined using template-, geometry- and physics-based methods included within the RAMP suite (3).

We used RAPDF to select the best five models among all of the decoys created for the full-length M180 amelogenin. We then have picked the best combination of regions from the top 5 models, using a graph-theoretic clique finding approach, in the creation of an optimized conformation model (6), the results are shown in figure S4.

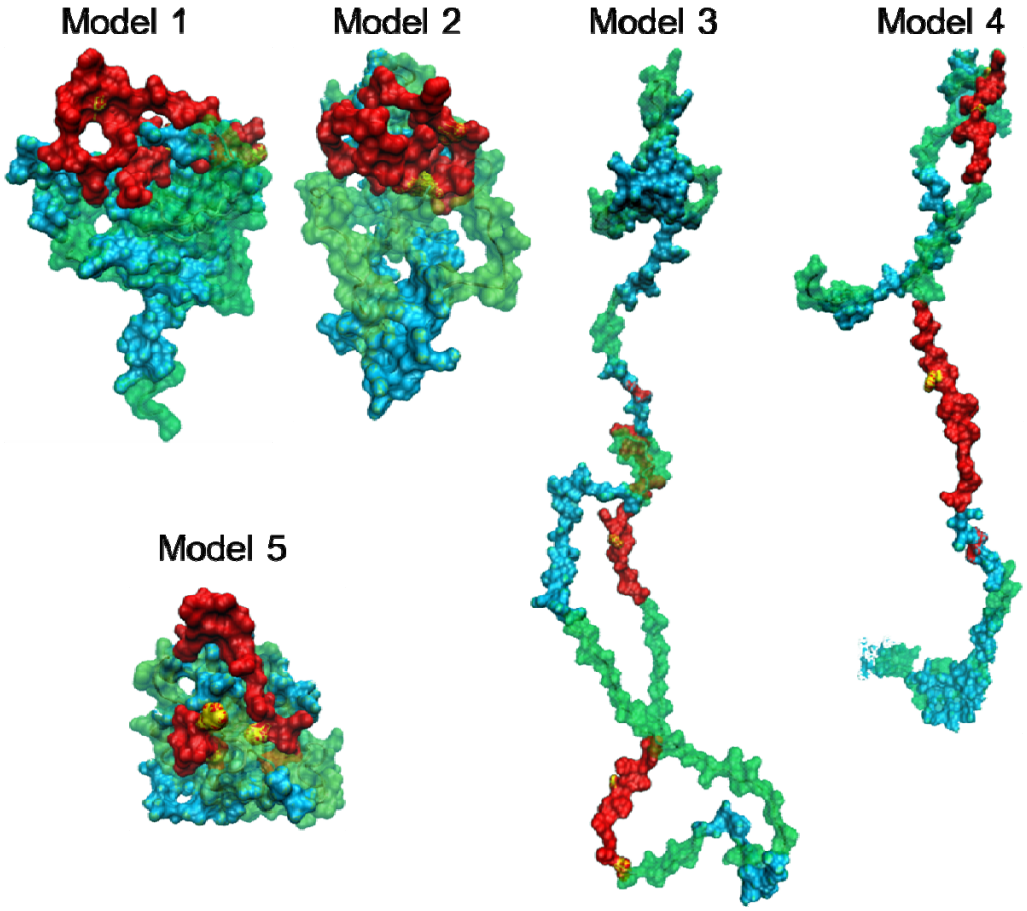


**Figure S4:** Top 5 optimized conformation models obtained from RAMP.

### $Ca^{2+}$ Ion Binding Domain Predictions

We have performed an exhaustive grid-based search placing ions into all solvent exposed sites in the 3D structure. The sites were scored with a grid based exhaustive sampling algorithm, clustered using a variant of an unsupervised hierarchy clustering algorithm, and refined with Monte Carlo minimization. The scoring function was developed from geometrical parameters observed for inorganic single atom ions coordinating small molecule organics found in many crystallography structures publicly available through the PDB (7). This scoring function was verified by accurately ranking the binding affinities between proteins and ions, predicting the naturally occurring ion binding sites of proteins and recovering the native conformation of protein-substrate binding <http://protinfo.compbio.washington.edu/soak> (8)

**Metafunctional Signatures**

The Meta-Functional Signature (MFS) method comprises five sub-scores. The compilation of these scores was trained on two databases of functionally important residues (9) using a simple logit model <http://protinfo.compbio.washington.edu/mfs>. (10-12) The MFS method can also be applied without structural analysis, using only scores based on sequence information, hidden Markov model (HMM), State-to-Step Ratio (SSR), and amino acid type (AA_type), to produce highly accurate identification of amino acids known to be important to function. (10)

*SSR* - The evolutionary context for a region is modeled by creating a phylogenetic tree for the sequence surrounding each amino acid in a protein, using multiple sequence alignments of PSI-BLAST results (13). The root of the tree represents the theoretical ancestral sequence, while each leaf of the tree represents a PSI-BLAST result sequence. We quantify the evolutionary divergence of each amino acid from its evolutionary context through the phylogenetic tree. SSR describes the ratio of the number of different amino acids presenting at a position across the multiple sequence alignments, to the total number of changes between the input protein and the base sequence within the phylogenetic tree. (10) *HMM* - We have compiled a HMM from the multiple sequence alignments described above, using the HMMER package (14). We then have compared emission frequency estimates from the HMM model with the amino acid background frequency given by karlin.c of the BLAST program package (13), to produce the HMM relative entropy score for each amino acid position (15).

*AA-type* – We have assigned a simple score to each amino acid type, developed from the frequency of each amino acid type appearing within active sites in two large databases.

*Structural Stability scores – virtual mutations* – We have evaluated the importance of each residue to the structural stability of the protein by performing a series of virtual mutations to the 19 alternate amino acids, modeling induced changes in tertiary structure, and comparing simulated energy function scores (RAPDF; (5)) for the mutation set for each residue. RAPDF_spread, the variation of RAPDF scores across all possible mutants for the residue, indicates the structural volatility of the position, and therefore the contribution of the position to stability. The RAPDF_dif structural stability score represents the sum of the simple differences between the RAPDF score for the input structure and all 19 alternatives, which indicates the effect of the naturally occurring amino acid identity on stabilization of the structure (11).

The structural stability score was implemented both within MFS and used separately for identifying fold bearing regions not likely to contribute to function but may create specific conformations important to function for the regions they span.

***In vitro* Solution Biomineralization**

To investigate the effect of peptides on in vitro calcium phosphate nucleation, an alkaline phosphatase (AP) based mineralization model was used. (16) This model mimics the biological matrix vesicle mediated mineralization, where inorganic phosphate is cleaved from an organic phosphate compound by AP and reacts with calcium inside the matrix vesicle. Prior to the mineralization experiments the activity of AP was measured in the presence of the peptides, to assess the effects of either peptide on the enzyme activity. *β*-Glycerophosphate (*β*-GP) solutions at 14.4 mM containing no peptide and 0.4 mM peptides were prepared in 25 mM Tris-HCl buffer (pH 7.4). The reaction was started by adding $1.4 \times 10^{-6}$ g/mL bacterial AP (Invitrogen, U.S.A.) to the solution. The reaction was carried at 37°C, and samples were collected at times 0.25-, 1-, 3-, 16- and 24-hr. The phosphate released in each solution was measured using PiPer Phosphate Assay Kit (Invitrogen Co., U.S.A.) and plotted against time.

For the mineralization experiments, mineralization solutions containing 24 mM $Ca^{2+}$ and 14.4mM *β*-GP and 0.4 mM peptide was prepared in 25 mM Tris-HCl buffer (pH 7.4). A mineralization solution containing no peptide was prepared as the negative control. Mineralization reactions were commenced by adding bacterial AP to the solution at a final concentration of $1.4 \times 10^{-6}$ g/mL. Subsequent mineral formation kinetics was monitored by continuous measuring of absorbance at 820 nm wave length using a Tecan Safire microplate reader (Tecan Trading AG, Switzerland) as well as by periodic assays of calcium ions throughout the reaction. For periodic assays of calcium and phosphate, 10 *μ*L of the reaction solution was collected at 0.25-, 0.5-, 1- and 1.5-hr. The calcium concentration in the solutions was determined using QuantiChrome Calcium Assay Kit (Bioassays, U.S.A.). All mineralization reactions were carried out at 37°C. Samples for electron microscopy analysis were prepared by placing a 10 *μ*L aliquot from the reaction solution onto a carbon coated TEM grid. After 1 min, liquid on the TEM grid was carefully wicked away and the grids were vacuum-dried for 5 min. The dried grids were stored in a desiccated container until analyzed. SEM imaging was performed at 10 kV using a JSM 7000F (JEOL, Japan) SEM in secondary electron imaging mode. TEM imaging and electron diffraction analysis were performed using an EMS 420T TEM (FEI, Inc., USA; formerly Philips, The Netherlands) operated at 120 kV.

We noted that there were three distinct trends of mineralization among the tested ADPs. (Figure 2) The majority of the peptides (ADP1, ADP2, ADP3, ADP4, ADP6 and ADP8) exhibited similar kinetics to the negative control, where no peptide was present. The phage display selected HABP1 and the ADP7 exhibited a slow mineralization trend. Interestingly, full-length rM180 and the ADP5 exhibited a fast mineralization trend where more than half of the available free $Ca^{2+}$ was consumed at the end of the 90 minutes. (Figure S5) The microstructural and crystallographic analysis of the synthesized minerals via SEM, TEM and XRD revealed another interesting consequence regarding the relationship between the mineral binding and mineralization activity. Similar to the mineralization kinetics, three distinct trends of mineral morphologies were observed. The majority of the peptides (ADP1, ADP2, ADP3, ADP4, ADP6, ADP8 and HABP1) produced spherulitic particles consistent with the formation of spherical amorphous calcium phosphate (ACP) and transformation into crystalline phases. (17-19) The amount of the radiating crystalline blade like particles from the spherulites were slightly higher for ADP1, ADP2, ADP4 and HABP1 (strong binders) compared to ADP3, ADP6, ADP8 (weak binders) and no peptide, an order indicating that the amorphous to crystalline transformation rates were slightly different. In the case of rM180 and ADP7, however, a completely different morphology was observed. Needle like nano-crystals were organized into bundle-like assemblies observed for rM180 and ADP7. The bundle-like assemblies appeared to be better organized in the case of rM180, which is consistent with the self-assembly properties of amelogenin and the *in vitro* mineralization behavior of recombinant amelogenin (20-22). The particles formed in the presence of ADP5 were much smaller spherulites with a less electron-dense core and smaller radiating crystals. The smaller particles sizes in the presence of ADP5 may be due to its nucleation dominated regime, as observed from in data describing the mineralization kinetics (Figure S3). The crystallographic analysis via XRD spectroscopy also confirmed the morphological differences were due to different crystal structure. In all cases except for rM180 and ADP7, the minerals yielded a broad peak around $2\theta$ 25-30$^{o}$ indicating a poorly crystalline phase. In the presence of rM180 and ADP7, however, the XRD patterns were consisted of numerous sharp peaks indicative of crystalline HAp. The major peaks were observed at $2\theta = 31.8^{o}$ (d = 2.81Å) and 32.1$^{o}$ (d = 2.78Å) corresponding to (211) and (300) planes of HAp, respectively. (Figure 2d)

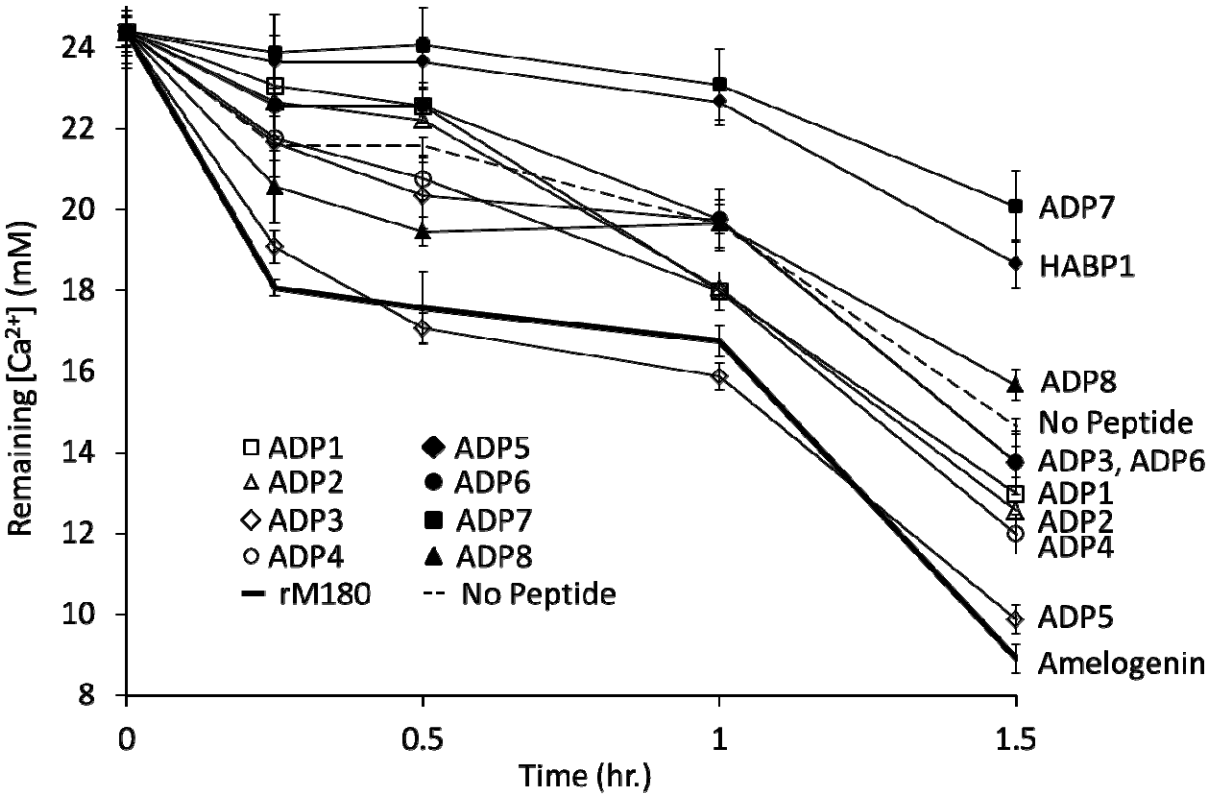


**Figure S5:** $Ca^{+2}$ consumption rates of the ADPs for the initial 90 minutes

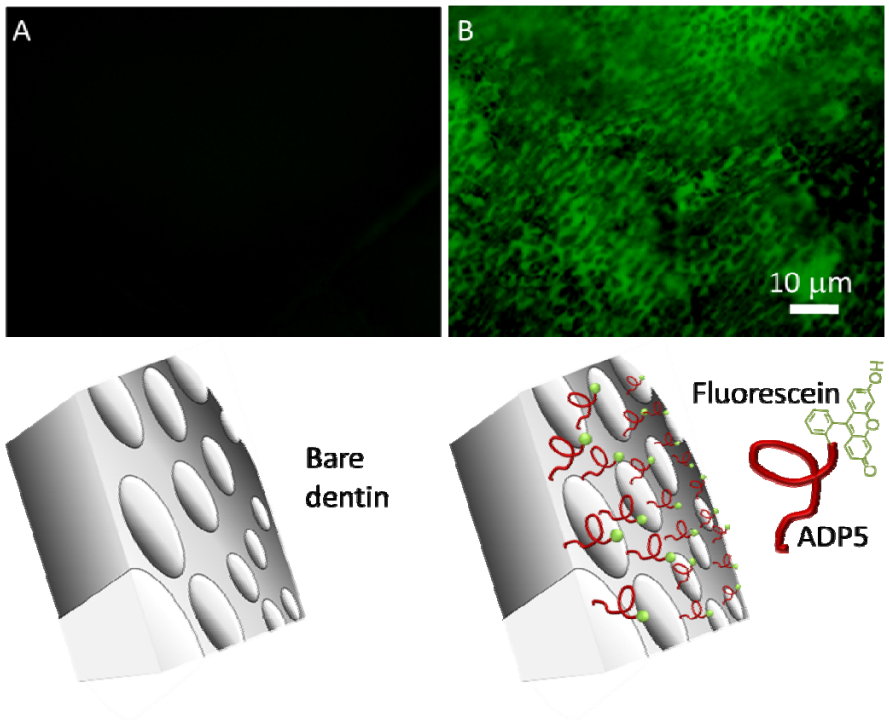


**Figure S6:** Fluorescence microscopy images of the root surfaces coated with no peptide (a) and f-ADP5 (b)

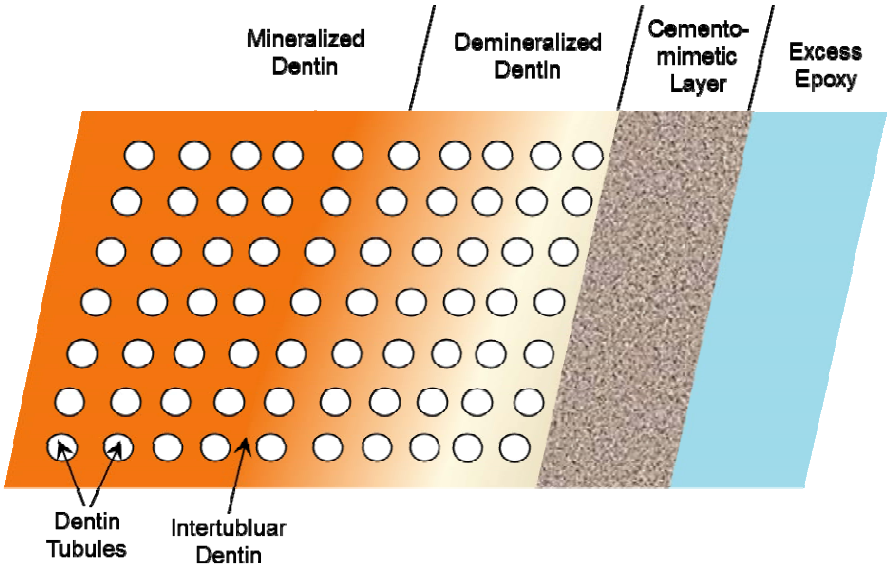


**Figure S7:** Schematic of the sample which nanoindentation characterization was performed. 20 indentations were made on each region: mineralized dentin, demineralized dentin (~20 um), and cementomimetic layer (~10 um). Indentations in dentin were made strictly on intertubular dentin.

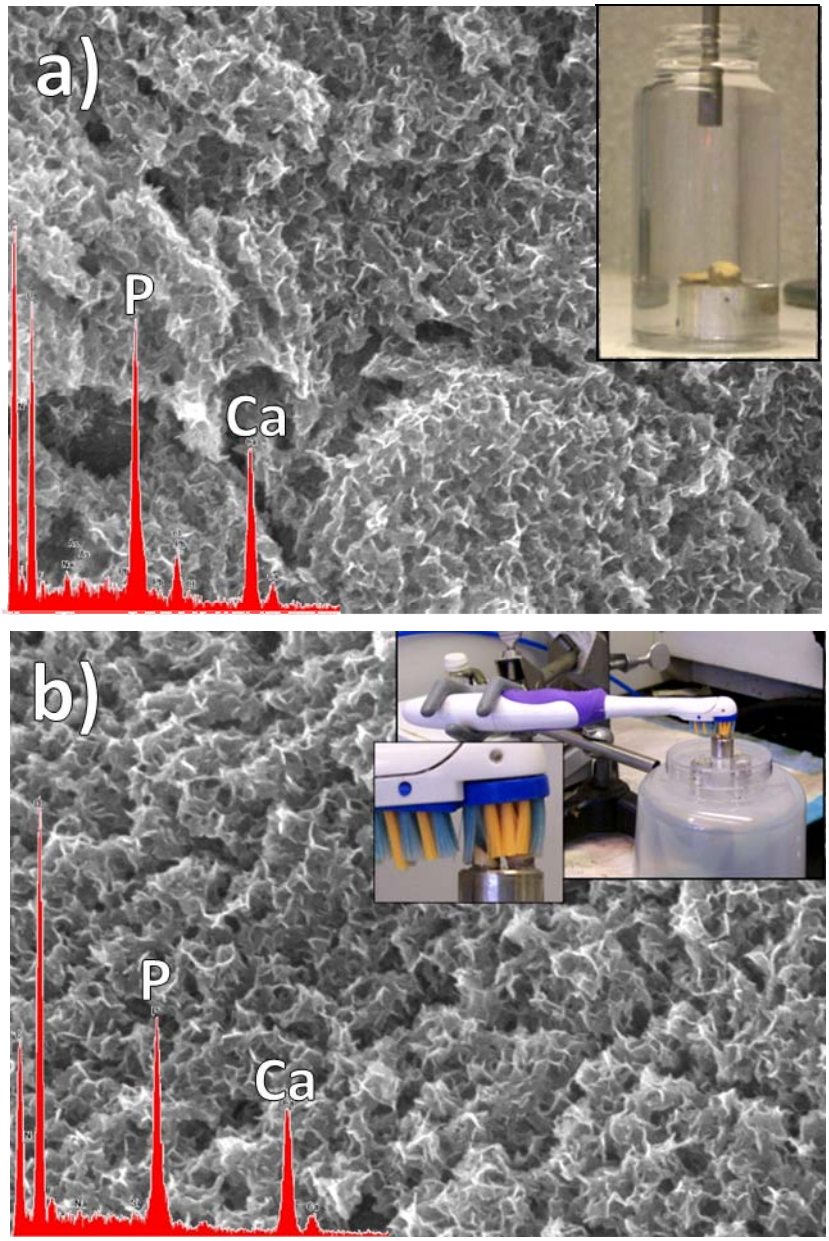


**Figure S8:** SEM micrographs and the corresponding EDXS spectra of the cementomimetic layer after qualitative mechanical analysis. (a) Ultrasonication and (b) Brushing. (Insets: experimental setups)

**References**


1. Ginalski K, Elofsson A, Fischer D, Rychlewski L. 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics* 2003; **19**(8)**:** 1015-1018.

2. Hung LH, Ngan SC, Liu T, Samudrala R. PROTINFO: new algorithms for enhanced protein structure predictions. *Nucleic Acids Res* 2005; **33:** W77-W80.

3. Samudrala R, Xia Y, Huang E, Levitt M. Ab initio protein structure prediction using a combined hierarchical approach. *Proteins-Structure Function and Genetics* 1999**:** 194-198.

4. Zhang Y. Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins-Structure Function and Bioinformatics* 2007; **69:** 108-117.

5. Samudrala R, Moult J. An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction. *Journal of Molecular Biology* 1998; **275**(5)**:** 895-916.

6. Liu T, Guerquin M, Samudrala R. Improving the accuracy of template-based predictions by mixing and matching between initial models. *Bmc Structural Biology* 2008; **8**.

7. Berman GP, James DFV, Hughes RJ, Gulley MS, Holzscheiter MH, Lopez GV. Dynamical stability and quantum chaos of ions in a linear trap. *Physical Review A* 2000; **61**(2): A#: 042307.

8. Cheng G, Samudrala R. An all-atom geometrical knowledge-based scoringfunction to predict protein metal ion binding sites, affinities and specificities. in preparation (2012).

9. Chelliah V, Chen L, Blundell TL, Lovell SC. Distinguishing structural and functional restraints in evolution in order to identify interaction sites. *Journal of Molecular Biology* 2004; **342**(5)**:** 1487-1504.

10. Wang K, Horst JA, Cheng G, Nickle DC, Samudrala R. Protein Meta-Functional Signatures from Combining Sequence, Structure, Evolution, and Amino Acid Property Information. *Plos Computational Biology* 2008; **4**(9): A#e1000181.

11. Cheng G, Qian B, Samudrala R, Baker D. Improvement in protein functional site prediction by distinguishing structural and functional constraints on protein family evolution using computational design. *Nucleic Acids Res* 2005; **33**(18)**:** 5861-5867.

12. Wang K, Samudrala R. FSSA: a novel method for identifying functional signatures from structural alignments. *Bioinformatics* 2005; **21**(13)**:** 2969-2977.

13. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W*, et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**(17)**:** 3389-3402.

14. Eddy SR. Hidden Markov models and genome sequence analysis. *Faseb Journal* 1998; **12**(8)**:** Abs.103, A1327-1327.

15. Wang K, Samudrala R. Incorporating background frequency improves entropy-based residue conservation measures. *Bmc Bioinformatics* 2006; **7**: A#385, DOI: 10.1186/1471-2105-7-385.

16. Gungormus M, Fong H, Kim IW, Evans JS, Tamerler C, Sarikaya M. Regulation of in vitro calcium phosphate mineralization by combinatorially selected hydroxyapatite-binding peptides. *Biomacromolecules* 2008; **9**(3)**:** 966-973.

17. Arys A, Jedwab J, Pireaux JJ, Philippart C, Dourov N. Brushite in the Pulp of Primary Molars. *Journal of Oral Pathology & Medicine* 1989; **18**(7)**:** 371-376.

18. Kodaka T, Hirayama A, Mori R, Sano T. Spherulitic brushite stones in the dental pulp of a cow. *Journal of Electron Microscopy* 1998; **47**(1)**:** 57-65.

19. Achilles W, Jockel U, Schaper A, Burk M, Riedmiller H. In-vitro Formation of Urinary Stones - Generation of Spherulites of Calcium-Phosphate in Gel and Overgrowth with Calcium-Oxalate Using a New Flow Model of Crystallization. *Scanning Microscopy* 1995; **9**(2)**:** 577-586.

20. Beniash E, Simmer JP, Margolis HC. The effect of recombinant mouse amelogenins on the formation and organization of hydroxyapatite crystals in vitro. *J Struct Biol* 2005; **149**(2)**:** 182-190.

21. Fan Y, Sun Z, Moradian-Oldak J. Controlled remineralization of enamel in the presence of amelogenin and fluoride. *Biomaterials* 2009; **30**(4)**:** 478-483.

22. Moradian-Oldak J, Paine ML, Lei YP, Fincham AG, Snead ML. Self-assembly properties of recombinant engineered amelogenin proteins analyzed by dynamic light scattering and atomic force microscopy. *J Struct Biol* 2000; **131**(1)**:** 27-37.