Control of calcium phosphate nucleation and transformation through interactions of enamelin and amelogenin exhibits the "goldilocks effect"

Citation for published version (APA):

DOI:
10.1021/acs.cgd.8b01066

Document status and date:
Published: 01/12/2018

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Control of calcium phosphate nucleation and transformation through interactions of enamelin and amelogenin exhibits the “Goldilocks effect”

Jinhui Tao1, Andreas Fijneman,1,2 Jiaqi Wan1#, Saumya Prajapati3, Kaushik Mukherjee3,
Alejandro Fernandez-Martinez4, Janet. Moradian-Oldak3* and James.J. De Yoreo1,5*

1Physical Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA 99352
2Laboratory of Materials and Interface Chemistry and Center of Multiscale Electron Microscopy,
Department of Chemical Engineering and Chemistry, Eindhoven University of Technology,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
3University of Southern California, Center of Craniofacial and Molecular Biology, Los Angeles,
CA, USA 90033
4Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTerre, Grenoble
38000, France
5Department of Materials Science and Engineering, University of Washington, Seattle, WA,
98195.

Correspondence:
joldak@usc.edu
james.deyoreo@pnnl.gov

# Current address: College of Materials Science and Engineering, Qingdao University of Science and
Technology, Qingdao 266042, China
Abstract

Although amelogenin comprises the vast majority of the matrix that templates calcium phosphate nucleation during enamel formation, other proteins, particularly enamelin, are also known to play an important role in the formation of enamel’s intricate architecture. However, there is little understanding of the interplay between amelogenin and enamelin in controlling processes of mineral nucleation and growth. Here we used an in vitro model to investigate the impact of enamelin interaction with amelogenin on calcium phosphate nucleation for a range of enamelin-to-amelogenin ratios. We found that amelogenin alone is a weak promoter of nucleation, but addition of enamelin enhanced nucleation rates in a highly non-linear, non-monotonic manner reaching a sharp maximum at a ratio of 1:50 enamelin:amelogenin. We provide a phenomenological model to explain this effect that assumes only isolated enamelin proteins can act as sites of enhanced nucleation, while enamelin oligomers cannot. Even when interaction is random, the model reproduces the observed behavior, suggesting a simple means to tightly control the timing and extent of nucleation and phase transformation by amelogenin and enamelin.

Introduction

During mineralization of hard tissues, soluble proteins act in a conjunction with an insoluble macromolecular matrix to guide biomineral formation.\textsuperscript{1} For example, recent work demonstrated that, in the case of calcite forming coccolithophorids, this cooperative effect is so important that, in the absence of the soluble protein, no mineralization of the matrix occurs.\textsuperscript{2} In some cases, such as tooth enamel, the matrix, which is comprised primarily of amelogenin, is processed by enzymes after secretion into the extracellular environment so that the end product contains only trace amounts of the original matrix.\textsuperscript{3} In other tissues, like bone and dentin, the collagen-based insoluble matrix becomes an essential part of the resulting structure and contributes to tissue function.\textsuperscript{4} Bone, dentin and enamel all contain numerous matrix-mineral interfaces during their development. For example, non-amelogenins in enamel or non-collagenous proteins in dentin and bone commonly have specific motifs that co-assemble with the matrix proteins amelogenin and collagen and additional charged functional groups that interact with the mineral surfaces\textsuperscript{5-10}. One of the biggest challenges in understanding biomineralization is to reveal the structural and functional relationship between the mineralizing constituents, the insoluble matrix and the
soluble proteins during the different stages of mineral nucleation, phase transformation and growth.

Here we focus on amelogenin, which is considered necessary for proper formation of dental enamel, as it comprises more than 90% of the organic component in the early stage of enamel formation.\textsuperscript{11} Amelogenin protein is mostly hydrophobic and contains a hydrophilic C terminus. It self-assembles \textit{in vitro} into nanospheres or nanoribbons depending upon the pH and the calcium or phosphate species concentration in solution.\textsuperscript{12-18} Amelogenin-null mouse models demonstrate the dramatic importance of amelogenin, as they exhibit a type of \textit{amelogenesis imperfecta (AI)} with disrupted patterns of enamel rods that compromise the integrity of the enamel.\textsuperscript{19,20} Moreover, amelogenin has been found to mediate calcium phosphate nucleation and growth \textit{in vitro} in different experimental settings.\textsuperscript{18,21-23}

Despite the fact that amelogenin comprises the vast majority of the enamel-forming matrix, it does not act alone. Non-amelogenin proteins like enamelin and ameloblastin play crucial roles in enamel formation, even though they are present in smaller quantities. For example, enamelin, a large glycoprotein (Mw=186 kDa), constitutes <5% of the extracellular matrix but is required for enamel mineral formation, as shown unequivocally in gene-targeted mouse models and human mutations resulting in hypoplastic \textit{AI}.\textsuperscript{24-28} No true enamel or apatite mineral ribbons are formed in \textit{Enam} \textsuperscript{-/-} mice.\textsuperscript{29} The abnormally thin and disorganized enamel in \textit{AI} patients suggests that enamelin is required to drive crystal formation, achieve structural organization of the apatite prisms and develop optimal enamel thickness.\textsuperscript{26,27} Despite its critical function in normal enamel formation, the mechanistic role of enamelin in the development of the mineral is not yet understood.

Like amelogenin, enamelin is processed by enamel proteinases immediately upon secretion. The major and the most stable cleavage product, 32 kDa enamelin, distributes throughout the developing enamel layer.\textsuperscript{30-33} Importantly, some mutations in the enamelin gene causing \textit{AI} have been reported to be within the 32-kDa enamelin fragment.\textsuperscript{34} \textit{In vitro} studies have demonstrated that the interaction between amelogenin and enamelin has a structural impact on both proteins, as well as a synergistic effect on apatite nucleation and octacalcium phosphate (OCP) crystal morphology.\textsuperscript{35-37} In the present study, to advance understanding of the interplay between amelogenin and enamelin in controlling enamel mineralization, we used \textit{in situ} AFM, \textit{in situ} Raman spectroscopy, SEM, TEM, and X-ray scattering to evaluate the effect of amelogenin and
the 32 kDa enamelin interactions on the nucleation, growth and phase transformation of calcium phosphate in vitro. We followed nucleation and growth events in bulk solution as well as on substrates with different degrees of hydrophobicity. Finally, we utilized Grand Canonical Monte Carlo (GCMC) simulations to elucidate how co-assembly impacts the status of enamelin monomers and aggregates, and to provide a rationale for the observed cooperative and dose-dependent effects.

Results and Discussion

First we used in situ AFM to observe the co-assembly of purified recombinant porcine full-length amelogenin rP172 (19.6 kDa) and native 32 kDa enamelin fragments on muscovite mica for a range of enamelin:rP172 ratios, following the sample preparation protocol shown schematically in Fig. 1. Phosphate solutions containing amelogenin and enamelin were injected into an AFM fluid cell and the protein coverage was determined as a function of time without flow (Figs. 2a,b and Fig. S1). The results show that increasing the enamelin:rP172 ratio led to slower adsorption kinetics (Figs. 2c and Fig. S1) with a smooth decrease in adsorption rate coefficient (Fig.2d). Previous in vitro biophysical studies performed in solution have demonstrated that 32kDa enamelin had the effect of dissociating amelogenin nanospheres that formed at pH 8, whereas it stabilized smaller amelogenin oligomers. Moreover, we have shown that amelogenin nanospheres dissociate and are not stable on charged surfaces. The addition of enamelin to amelogenin may affect the charge distribution on the protein particles formed in the solution, hence resulting in their different adsorption behaviors. The height distribution of particles did not change at different enamelin:rP172 ratios, which indicates that enamelin does not change the association between rP172 assemblies when adsorbed on the surface (Fig S1g). The slower adsorption kinetics of enamelin/rP172 complexes at higher enamelin:rP172 ratio may be therefore due to the stronger repulsive force between negatively charged mica and negative charged enamelin/rP172 complex.

After mixing amelogenin and enamelin in phosphate solution, the AFM fluid cell was flushed with water to remove the phosphate buffer. Mineralization of the protein-covered surfaces was then induced by flowing supersaturated calcium phosphate solution containing 1.5 mM CaCl₂ and 9.5 mM Na₂HPO₄ (pH 7.40) into the fluid cell (Fig. 3a-d). As previously shown, under these conditions, the initial mineral phase formed was amorphous calcium phosphate (ACP), which then transformed into octacalcium phosphate (OCP) and finally to HAP. The results show that
the inclusion of enamelin impacted the rate of ACP nucleation (Fig. 3e, f), but the effect was highly non-linear, exhibiting a sharp maximum at an enamelin:rP172 ratio of 1:50, when the rate was enhanced by 400% over the value of amelogenin-adsorbed films in the absence of enamelin. In contrast, when the enamelin:rP172 ratio was increased above 1:50, nucleation was inhibited (Fig. 3f). The sharp maximum at an enamelin:rP172 ratio of 1:50 is achieved in repeated experiments and exceeds the rate for pure amelogenin by five to fifteen times the error in the measurements. This dose-dependent effect on calcium phosphate nucleation in which a sharp maximum was reached at a low enamelin concentration and was then followed by inhibition was observed on the surfaces of highly ordered pyrolytic graphite (HOPG) (Fig. S2). Such a surface is both uncharged and highly hydrophobic, showing that neither the surface charge of mica nor hydrophobicity of HOPG affected the observed dose-dependent effect.

To determine whether the effect of enamelin on ACP nucleation when added with rP172 was influenced by some other property of the substrates, we monitored the separate and combined effects of these two proteins on calcium phosphate nucleation and phase transformation in bulk solution and followed the progress of nucleation by in situ Raman spectroscopy and X-ray diffraction (Fig. 4a-f). Based on the position of the Raman peak (Fig. 4c) and the short-range order within 0.5 nm in the X-ray pair distribution function (PDF) (Fig. 4f), we conclude that, in the absence of protein, the mineral formed was still amorphous even after 30 min. By about 40 minutes, long-range order up to 1.5 nm became evident in the X-ray PDF data and the Raman peak shifted to the position characteristic of crystal, which indicates the phase transformation started at 40-50 min. Short, rod-shaped crystals were observed in TEM images around the surface of ACP particles when the phase transformation began, and the number and length of these short crystals increased with prolonged incubation (data not shown). Finally, the crystals were in the shape of thin plates with less controlled crystal orientation, as shown by both morphology and electron diffraction (Figs. S3, S4). The presence of the proteins individually or in combination does not appear to have a strong effect on the crystal morphology after 200 min though the kinetics of transformation are altered (Figs. S4).

The time evolution of the optical images in the presence of protein shows that rP172 alone enhanced the rate of ACP nucleation (Figs. 4a, S5), while also increasing the lifetime of the ACP phase (Fig. 4d-e). However, when rP172 and enamelin were mixed together in solution, a more complicated dependence on enamelin:rP172 ratio was observed; the nucleation rate again
increased with enamelin:rP172 ratio, reaching a maximum at ~ 4:50 (Fig 4a, b). The lifetime of the ACP phase (Fig. 4d-e) also increased with increasing enamelin:rP172 ratio, reaching a maximum at a ratio of about 2:50, at which the lifetime exceeded that observed with rP172 alone. However, the lifetime then decreased to a minimum at about 8:50; at this ratio, the lifetime was below that obtained with either rP172 or enamelin alone (Fig 6S), before increasing again to a level equivalent to that of rP172 alone (Fig. 4e).

The results show that, over the range of enamelin:rP172 ratios in which nucleation of ACP is enhanced by the presence of enamelin, the degree to which nucleation of the more stable phase is inhibited in bulk solution is also increased and thus ACP can survive for longer times over this range. Moreover, these results show that the synergistic effect of enamelin and amelogenin occurred both on substrates and in bulk solution. This synergistic and dose-dependent effect has also been reported for the nucleation of apatite crystals inside a gelatin gel diffusion system. Similarly synergistic regulation by 32 kDa enamelin and a recombinant pig amelogenin affected the morphology of (OCP) crystals, with the interplay of enamelin and amelogenin enhancing the stability of the transient amorphous calcium phosphate (ACP) phase and increasing the length-to-width ratio (aspect ratio) of the OCP crystals.

These results demonstrate that enamelin is a promoter of ACP nucleation when co-assembled with amelogenin regardless of whether it is in bulk solution or deposited on a surface, provided the concentration is sufficiently low, with the degree of enhancement exhibiting a sharp maximum near an enamelin:rP172 ratio of 1:50 on a surface and 4:50 in bulk solution. It is noteworthy that both proteins; enamelin and amelogenin stabilize ACP phase, but their ability to stabilize ACP is maximized in a similar range of enamelin:amelogenin ratios (1:50-4:50) at which nucleation is most significantly enhanced, both in solution and when adsorbed on surfaces.

The fact that the degree of ACP nucleation enhancement decreases above a certain ratio implies that the observed dependence is not simply a result of saturating the available surface sites (i.e., above a certain concentration, no further gain is achieved). Rather, the dependence implies that, above a certain concentration of enamelin, further addition of enamelin eliminates the sites of enhancement created at lower concentration. To explain this effect we propose a simple statistical model that hypothesizes isolated enamelin monomers are sites of enhanced nucleation, but those that have other enamelin monomers as neighbors will form aggregates or oligomers of
enamelin whose conformation or exposed surface differs from that of the isolated proteins and renders them unable to promote nucleation. This assumption is based on the previous reports that support the calcium-binding potential of 32 kDa enamelin. Circular dichroism spectroscopy showed that enamelin exhibits a conformational change with a preference for the \( \beta \)-sheet upon addition of calcium ions. In addition, previous studies using DLS showed that the 32 kDa enamelin solution is monomeric up to 6.4 \( \mu \)g/mL in neutral pH bulk solutions with calcium ions, but above this concentration large aggregates were detected. This indicates that the 32 kDa enamelin has a strong self-binding nature and forms large aggregates when it reaches a critical concentration.

To determine whether this hypothesis leads to a concentration dependence of nucleation promoters that is consistent with the observed dependence, we used a Grand Canonical Monte Carlo algorithm to simulate the random deposition of a monomeric adsorbates onto a surface and analyzed the results to determine the number of isolated “monomers” on that surface as a function of the total fraction of surface sites containing an adsorbate (Fig. 5a,b). The results show that the percentage of surface sites containing an isolated monomer is a strong, non-linear function of adsorbate coverage (Fig. 5c) that rises linearly at low coverage and reaches a sharp peak at a fractional coverage of about 0.02, before falling rapidly. Equating the adsorbates (green squares in Fig. 5b) to enamelin and the open sites (white squares in Fig. 5b) to rP172, the maximum then lies near an enamelin:rP172 weight ratio of 0.022, which corresponds to an enamelin:rP172 ratio close to 1:50 in bulk solution (Fig. 5d). Above that ratio, the fraction of isolated monomers falls rapidly while the percentage that has one or more neighboring rises. Introducing an attractive interaction between nearest neighbors biases the adsorption maximum peak toward a lower enamelin:rP172 ratios as enamelin self-aggregates more readily whereas introducing a repulsive interaction has the opposite effect (Fig S7). However, the maximum peak still always exists in these cases.

With the results of this simple statistical, the dependence of nucleation rate on the enamelin:rP172 ratio can be understood, provided isolated enamelin proteins within the amelogenin matrix are preferred sites of nucleation while those with enamelin neighbors (i.e., enamelin oligomers or aggregates) are inactive. As the enamelin:rP172 ratio increases, at first the number of isolated enamelin sites increases rapidly and thus the rate of nucleation also increases rapidly. However, beyond the value where the percentage of enamelin sites that are isolated
reaches a maximum, which is expected to be in the vicinity of 1:50, the rate of nucleation begins to drop again. Because active nucleating sites that would have been covered with amelogenin are now covered with enamelin in the form of aggregates or oligomers, the rate of nucleation drops below that of a pure amelogenin film, leading to a regime of inhibition, as observed experimentally (Fig. 3f).

These findings raise the question of why enamelin monomers would promote nucleation, but enamelin aggregates or oligomers would not. Generally speaking, protein aggregation and misfolding can be expected to interrupt the structured assembly needed to promote ion binding and mineral nucleation. Structured assembly of proteins is the basis for directed mineral nucleation, both in natural biominerals and in biomimetic systems, and can easily be understood as a result of the local structural and chemical environment created through assembly. One clear example is the case of silica production by silicateins, during which distinct peptide residues are brought into juxtaposition. The loss of protein function in inducing mineral nucleation can also be explained if the protein-protein interactions leading to random aggregation or structured oligomerization involve the peptide segments or side chains that are required for promotion of mineral nucleation. Because we do not know which structural and/or chemical features of enamelin are responsible for promotion of ACP nucleation, we cannot identify which enamelin-enamelin interaction motifs are responsible for the loss of function. Nonetheless, given that 32 kDa enamelin is highly charged, as are the regions often implicated in promotion of mineral nucleation by proteins and peptides, blocking the exposure of charged sidechains is one obvious explanation for the observed effect. We speculate that during enamelin oligomerization or aggregation these highly charged anionic or cationic regions may undergo coordination with one another via Ca$^{2+}$ or HPO$_4^{2-}$ species and thus no longer be available to promote nucleation.

In summary, the in vitro experimental results presented here have important implications for the functional mechanisms of proteins in enamel mineral formation. While the bulk solution concentrations of proteins used in these experiments are below the in vivo protein concentrations in enamel fluid, we believe that the physiochemical events that occur at the mineral-protein interface may well be relevant. The in vitro models used here are designed to follow such events at high precision and at molecular or atomic levels and the surfaces employed act to enhance protein concentrations well beyond their bulk values. Thus they may well mimic the in vivo
situation at the protein-mineral interface. Moreover, whether the protein concentrations at the time of mineral genesis in these isolated environments are at the elevated levels seen in enamel fluid, or are substantially smaller, is unknown. Finally, the principle elucidated by these results, that a tendency towards co-localization by proteins can lead to complex dependencies of mineral formation kinetics on relative concentrations simply due to the statistics of interactions, is a general one that is applicable regardless of the specific proteins under consideration or their concentrations.

Our results support the concepts of; a) protein dose dependency b) cooperative effects between proteins, and c) the control of calcium phosphate phase transformation in enamel formation. The 32 kDa enamelin selected for this study is the most stable cleavage product of the full-length enamelin and the only one that can be purified and isolated. It is rich in glutamic and aspartic acid residues and is highly phosphorylated. It exhibits a remarkable level of control over nucleation and phase transformation kinetics, as well as final morphology. Alteration of enamelin phosphorylation sites leads to enamel malformation in AI patients. The present study highlights the concept of enamelin dose-dependency, which has also been demonstrated in transgenic mouse models in which expression levels of the enamelin gene were controlled. Hu et al. have shown that only an optimal quantity of enamelin will lead to normal enamel formation. Deficiency or excess of enamelin protein eradicates the formation of enamel crystals and affects prism formation. The concept of synergy between enamel-forming proteins was put forth in an in vivo study of double knockout mice in which the genes for amelogenin (Amel) and ameloblastin (Ambn) were deleted. The destructive effect on enamel was more severe in the double knockout compared to Amel or Ambn single knockouts.

The results of the present study provide additional in vitro evidence for cooperative effects between amelogenin and enamelin. We have previously reported interactions, and co-localization between amelogenin and enamelin in enamel. We have further demonstrated their cooperative function in controlling crystal morphology. Here we report the precise level of control over ACP nucleation and phase transformation by these two proteins. The fact that these two proteins interacts and colocalize during the early stage of enamel formation supports the notion that the stability of the ACP mineral phase at the early state of enamel formation is not merely controlled by amelogenin and its proteolytic products, but it is a cooperative function between amelogenin and non-amelogenins. Given the large diversity of enamel proteins including
enamelin, ameloblastin and amelotin, and considering the complexity in their pattern of expression, secretion, processing and assembly, the research presented here provides a general strategy to evaluate the synergetic role of these proteins through co-assembly in controlling the mineralization process. Understanding such molecular processes is a key step enabling the synthesis of biomimetic hybrid materials for medical applications.

**Experimental Procedures**

*Amelogenin Expression and Purification*

Recombinant porcine amelogenin (rP172) was expressed in *Escherichia coli* strain BL21-codon plus (DE3-RP, Agilent Technologies, Inc., Santa Clara, CA) as previously described.\(^{49-51}\) Protein purification was accomplished on a reverse phase C4 column (10 × 250 mm, 5µm) mounted on a Varian Prostar HPLC system (ProStar/Dynamics6, version 6.41 Varian, Palo Alto, CA), using a linear gradient of 60% acetonitrile at flow rate of 1.5 ml/min. Recombinant porcine amelogenin rP172 is an analog of full-length native porcine P173, which has 173 amino acids, but rP172 lacks the N-terminal Met and a phosphate group on Ser16.\(^{52}\)

*Enamelin isolation and purification:*

The 32kDa enamelin was extracted from unerupted 2\(^{nd}\) and 3\(^{rd}\) mandibular molars of six-month-old pig jaws, purified and characterized as described previously following the method described by Yamakoshi et al.\(^{5,30,31}\) Samples were scraped from 2\(^{nd}\) and 3\(^{rd}\) unerupted molars (Farmers John Clougherty Co., Los Angeles, CA, USA, through Sierra for Medical Sciences, Santa Fe Springs, CA, USA). The pooled enamel samples were homogenized in 50 mM Sørensen buffer (pH 7.4) with proteinase and phosphatase inhibitors. This extraction process was repeated three times. The combined supernatant was processed and purified by reverse-phase high performance liquid chromatography (RP-HPLC) as reported previously.\(^{35}\) The purified 32kDa enamelin was characterized by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), stains-all staining and Edman degradation.\(^{35}\)

*Amelogenin rP172 and enamelin 32 kDa co-assembly*

The rP172 and enamelin 32 kDa were dissolved separately in pure water to yield final concentrations of 1000 µg/mL and 300 µg/mL, respectively. The rP172 was diluted in a phosphate buffer containing 19 mM Na\(_2\)HPO\(_4\) at pH 7.40. The final rP172 concentration was fixed at 50 µg/mL in all cases. Enamelin 32kDa was introduced into rP172 solution at various final enamelin concentrations from 1 µg/mL (1:50), 2 µg/mL (2:50), 4 µg/mL (4:50) to 8 µg/mL
(8:50). Then the mixed protein solutions were injected into liquid cells with a freshly cleaved muscovite mica disc (diameter 9.9 mm, Ted Pella, Inc.) or highly oriented pyrolytic graphite (HOPG) on the bottom, the solution was incubated in the liquid cell and the adsorbed proteins were imaged by in situ AFM, with the nonadsorbed proteins flushed away by water after 50 min. The images were analyzed using the image processing software package SPIP 5.1.4 (Image Metrology A/S, Hørsholm, Denmark). The particle size distribution and areal coverage were calculated by the Particle & Pore Analysis module included in the SPIP 5.1.4 software.\(^5^3\)

**In situ Atomic Force Microscopy (AFM)**

All in situ AFM images were captured in tapping mode at room temperature (25°C) with a Nanoscope 8 atomic force microscope (Digital Instruments J scanner, Bruker) with hybrid probes consisting of silicon tips on silicon nitride cantilevers (HYDRA triangular lever, \(k=0.088 \text{ N/m, tip radius } <10 \text{ nm; resonance frequency 75 kHz in air; Applied Nanostructures, Inc, www.appnano.com}\)). The drive amplitude was about 20 nm (in fluid), and the signal-to-noise ratio was maintained above 10. The scanning rate was 1 Hz. The amplitude setpoint was carefully tuned to minimize the average loading force (~50 pN) loaded during in situ imaging. The supersaturated calcium phosphate solution was made by first mixing 3.0 mM of CaCl\(_2\) with 19.0 mM of Na\(_2\)HPO\(_4\) (pH 7.40) at equal flow rates by T junction before pumping the solution into the fluid cell at a constant rate of 3 mL/h at 25°C, which ensured kinetically limited nucleation conditions in our systems. The supersaturation \(\sigma\) is 3.47 for the above solution. Supersaturation is defined as \(\sigma = \ln \left[ \left( \alpha Ca^{2+} \right)^{5} \left( \alpha PO_{4}^{3-} \right)^{3} \left( \alpha OH^{-} \right) / Ksp \right] \), where \(\alpha\) denotes the species activity and \(Ksp\) denotes the equilibrium solubility constant at 25°C. Speciation calculations were performed using Davies’s extended Debye-Hückel equation\(^5^4\) from mass balance expressions for total calcium and total phosphate with appropriate equilibrium constants by successive approximation for the ionic strength. The solubility activity products of HAP are \((3.0610 \times 10^{-7} \text{M})^{9.55}\). Calcium and phosphate solutions were prepared immediately before use from reagent grade calcium chloride (CaCl\(_2\)) and di-sodium hydrogen phosphate (Na\(_2\)HPO\(_4\)) in water (pH 7.40). Each protein substrate was assembled immediately before calcium phosphate nucleation experiments and was briefly cleaned by flushing 1mL of deionized (\(\geq 18 \text{ M} \Omega\)) water to remove unbound proteins. Measurements were made over a range of enamelin:rP172 ratios and nucleation rates were determined from data collected shortly after the onset of nucleation, where the particle number density exhibited a linear dependence on the elapsed time. Both mica
and HOPG were chosen to study the substrate effect on enamelin and rP172 co-assembly, which results in different mineral nucleation rates with different enamelin:rP172 ratios. Data used in our analyses were typically collected within the first 1 h of each experiment.\textsuperscript{39} For the \textit{in situ} studies, we repeated each experiment twice. Due to the well-controlled conditions including the protein ratio, solution supersaturation and flow rate, we found that the nucleation rates were very reproducible. We specifically repeated the experiments to make sure two observables were reproducible: 1) the number of nucleation events in specific area, and 2) the time scale over which the nucleation events occur.

\textit{In situ Raman spectroscopy}

\textit{In situ} micro-Raman spectroscopy under an inverted microscope was used to investigate the separate and cooperative effects of rP172 and enamelin on the kinetics of phase transformation of amorphous calcium phosphate in solution in real time. The supersaturated calcium phosphate solutions mentioned above ($\sigma$=3.47) were made by injecting 100 µl of 3.0 mM CaCl\textsubscript{2} solution into 100 µl of 19.0 mM Na\textsubscript{2}HPO\textsubscript{4} (pH 7.40) with 50 µg/mL rP172 and different enamelin concentrations. Therefore the composition of final solution is 1.5 mM CaCl\textsubscript{2}, 9.5 mM Na\textsubscript{2}HPO\textsubscript{4} and 25 µg/mL rP172 with different concentrations of enamelin. The concentrations of enamelin in the solution with enamelin:rP172 ratio of 1:50, 2:50, 4:50, 8:50, 15:50 are 0.5 µg/mL, 1.0 µg/mL, 2.0 µg/mL, 4.0 µg/mL, 7.5 µg/mL with the 25 µg/mL rP172. 1:0, 2:0, 4:0, 8:0 and 15:0 means the solution contained only enamelin at the same enamelin concentration as of the solution designated 1:50, 2:50, 4:50, 8:50 and 15:50, respectively. These supersaturated solutions were then put into a plastic vial with a 50 µm×50 µm square pinhole on the bottom (Fig. S5). The pinhole was used for laser penetration from the inverted optical microscope. This design excludes the use of other solid substrates between the mineral and laser, which minimizes the substrate background signal. The Raman spectra were collected from 100 to 4000 cm\textsuperscript{-1} under backscattering geometry by a LabRAM ARAMIS confocal Raman Microscope (HORIBA scientific, Japan) operated at a resolution of 2 cm\textsuperscript{-1} with an excitation wavelength of 532 nm and laser power of 2.5 mW. A ×100 TU plan fluro objective with numerical aperture of 0.90 (Nikon, USA) was used to focus the solution/air interface and collect spectra for 20 seconds. For \textit{in situ} Raman, we double checked five different solution conditions and evaluated the reproducibility of the ACP lifetime. The relative error in the life time ranged from 3.5% to 7.3%. These errors did not change the observation that: enamelin only inhibits ACP transformation and enamelin
enhances ACP transformation in the presence of rP172 at enamelin:rP172 ratio of 8:50. The centers of peaks at different time points were identified by fitting with a Guassian function. The nuclei of ACP in solution were imaged by an inverted optical microscope, and the number of nuclei in each frame was counted manually with Image J.

**Synchrotron-based X-ray scattering experiments**

For monitoring the phases of calcium phosphate at different time points, supersaturated calcium phosphate solution was made by first mixing 3.5 mM of CaCl$_2$ with 19.0 mM of Na$_2$HPO$_4$ (pH 7.40) in equal volumes in an Eppendorf tube and left in a static state for a period of time between 3-90 minutes. The suspension was then centrifuged at 11800 rpm for 5 mins, after which the supernatant was carefully withdrawn and the solids on the bottom were washed with ethanol three times. The time points are calculated from the time when CaCl$_2$ and Na$_2$HPO$_4$ solutions were mixed until the ethanol wash. The solids were dried in a vacuum for 6 hrs before taking the X-ray scattering measurements.

High-energy X-ray total scattering experiments were performed at beamline 11-ID-B of the Advanced Photon Source at Argonne National Laboratory, USA. Scattering data were collected with a Perkin Elmer amorphous silicon detector using the rapid-acquisition pair distribution function technique.$^{56,57}$ Measurements of the samples, empty Kapton capillary tube, and background were made at ambient temperature in a q-range of 0-22 Å$^{-1}$ using an incident radiation of ~ 60 keV ($\lambda=0.2127$ Å). Corrections for sample to detector distance, tilt angle of the detector with respect to the direction of the incident radiation, and polarization were performed using Fit2D software.$^{58}$ Total scattering structure functions and pair distribution functions (PDF) were obtained using the PDF-GetX2 software.$^{59}$ A crystalline standard (CeO$_2$) was measured and used to calculate sample-detector distance and the instrumental resolution effect on the PDF.$^{60}$

**Transmission Electron Microscopy (TEM)**

For study of the morphology and crystallinity of calcium phosphate at different time points, supersaturated calcium phosphate solution was made as described above for X-ray scattering. Minerals were collected at different time points by dropping 50 µL aged supersaturated solution onto a lacey carbon-coated 400 mesh TEM copper grid (Ted Pella, Inc.) and kept static for 1 min, after which the non-adsorbed mineral was flushed away with 100 µL water three times. After that, the TEM copper grids were dried in a vacuum. TEM was performed with a field
emission FEI Tecnai G² F20 (Hitachi Ltd.) operated at 200kV. Images were collected by an
Eagle CCD (1024×1024 pixels), which was included in the Tecnai User Interface (TUI).

**Scanning Electron Microscopy (SEM)**

SEM was performed with a Helios 600 dual-beam focused ion beam/scanning electron
microscope at 5 kV. The mineral samples on HOPG substrate were taken after AFM study,
flushed with water and dried with N₂ immediately. No coating was used before the SEM
imaging.

**Grand Canonical Monte Carlo (GCMC) simulation**

In the Grand Canonical ensemble, one first chooses randomly whether a trial particle insertion or
deletion is attempted. If insertion is chosen, a particle is placed with uniform probability density
inside the system. If deletion is chosen, then one deletes one out of N particles randomly. The
trial move is then accepted or rejected according to the usual Monte Carlo algorithm. As usual, a
trial move from an original state (o) to a new state (n) is accepted with different probability
expressions for insertion, deletion and displacement.⁴¹,⁴² That is, in the insertion step, a particle is
placed at a random location inside the simulation box with a probability

\[
\text{acc}(N \rightarrow N + 1) = \min \left(1, \frac{V_z}{N+1} \exp \left(-\beta(U_{N+1} - U_N)\right)\right);
\]

in the deletion step, a randomly selected
particle species is removed with a probability

\[
\text{acc}(N \rightarrow N - 1) = \min \left(1, \frac{N}{V_z} \exp(-\beta(U_N - U_{N-1}))\right);
\]

and, in the displacement step, a particle is randomly selected and displaced with a probability

\[
\text{acc}(o \rightarrow n) = \min(1, \exp(-\beta(U_n - U_o)));
\]

Where activity z was used to measure the particles chemical
potential μ with the relationship of

\[
Z = e^{\beta\mu}/\Lambda^3 \quad \text{outside the simulation box, the } 1/V \text{ is a factor}
\]

results from placing the particle with uniform probability anywhere inside the simulation box, β
is 1/k_BT, and Λ is the de Broglie thermal wavelength.

The code was downloaded from [http://www.cchem.berkeley.edu/chem195/_monte_carlo.html](http://www.cchem.berkeley.edu/chem195/_monte_carlo.html)
and modified as necessary for our calculations. The simulation box is a cube of 50×50 lattices.
The simulation uses the above-mentioned three standard types of GCMC insertion, deletion and
displacement moves to sample the total and monomer distributions of enamelin (green dots) in
amelogenin (white squares) matrices.⁴¹,⁴² 50000 steps were used for the GCMC simulation and
frames were output every 250 steps. The coverage of total green particles and green particles
without neighbors were calculated and averaged for 200 frames in the ensemble for different
activities. The relationships of total green dots and green dot “monomer” coverages to activity z were also calculated in the presence of 1 $k_B T$ repulsive or attractive interactions between nearest neighbors of green dots.

**Acknowledgements**

The authors gratefully acknowledge funding from the National Institutes of Health, NIDCR R01-DE013414 to JMO. The AFM, Raman, and TEM studies and the kinetic Monte Carlo simulations were performed at the Pacific Northwest National Laboratory (PNNL), which is a multi-program national laboratory operated by Battelle for the U.S. Department of Energy (DOE) under Contract DE-AC05-76RL01830. Jinhui Tao acknowledges support from the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory. The SEM imaging was performed using EMSL, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at PNNL. Amelogenin and enamelin production and purification were carried out at the Center for Craniofacial Molecular Biology, University of Southern California. Synchrotron-based X-ray scattering experiments were done at Argonne National Laboratory and use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract DEAC02-06CH11357. The authors thank Karena Chapman for her assistance during synchrotron X-ray scattering experiments at sector 11-ID-B, Advanced Photon Source, Argonne National Laboratory, Dr. Bin Wang and Mr. Guomin Zhu for assistance with collection of optical video for nucleation studies, Dr. Dongsheng Li for assistance in the TEM instrument access and Dr. Elias Nakouzi for constructive discussion of GCMC code modification.

**Supporting Information**

AFM images of amelogenin-enamelin adsorption, AFM images of calcium phosphate nucleation, SEM images of mineral morphology with amelogenin rP172, TEM images of minerals formed at different stages of nucleation, schematic illustrating the liquid cell for Raman spectroscopy, curves representing the equilibrium coverage of hypothetical enamelin monomers as represented by the green dots in Fig. 4.
Figure 1. Schematic representation showing procedure for protein complexes and subsequent calcium phosphate nucleation experiments. a) enamelin and rP172 mixtures at different enamelin:rP172 ratios followed by water washing, b) exposure to flowing supersaturated calcium phosphate solution, c) induction of nucleation.
Figure 2. *In situ* AFM observations of protein assembly on muscovite mica and time sequences of *in situ* AFM images showing adsorption of proteins on muscovite mica for a) enamelin:rP172=0:50 (0 µg/mL enamelin and 50 µg/mL rP172), and b) enamelin:rP172=4:50 (4 µg/mL enamelin and 50 µg/mL rP172). The first time point showed more separated protein particles because of the low coverage, as coverage increases at longer time, the coalescence of protein results in the flattening in topography, then the contrast of image drops. c) Measured adsorption kinetics of proteins for different enamelin:rP172 ratios where θ is the fraction of the surface covered by protein. The straight lines are linear fits of the initial stage of the adsorption process; d) the relationship between the kinetic constant for adsorption and the range of enamelin:rP172 ratios from the fit in c). All scale bars are 200 nm. The time when protein solution is added into liquid cell is defined as Time = 0.
Figure 3. *In situ* AFM observations of calcium phosphate nucleation and growth on rP172 and enamelin:rP172 films co-assembled on mica for enamelin:rP172 ratios of: a) 0:50 (50 µg/mL rP172 only), b) 1:50 (1 µg/mL enamelin and 50 µg/mL rP172), c) 2:50 (2 µg/mL enamelin and 50 µg/mL rP172), and d) 4:50 (4 µg/mL enamelin and 50 µg/mL rP172). Time series of *in situ* AFM images showing ACP (bright particles) nucleation and crystallization on the protein films shown in (a-d) in calcium phosphate solution. No free proteins are included in calcium phosphate solution. ACP particles (bright particles) dynamically nucleate and increase in numbers in the first 30 min. e) Dependence of the number density of nuclei on time t at five different enamelin:rP172 ratios obtained from the AFM in (a-d) plus that for a film with an enamelin:rP172 ratio of 8:50. f) Dependence of the steady-state nucleation rate $J_n$ on time t at five different enamelin:rP172 ratios derived from slopes linearly fitted by straight lines in e, the mean value is averaged between two experiments and error bars are added by using the larger error values from the experimental data and linear fitting. The scale bars for (a-d) are 200 nm. The time when CaCl$_2$ solution is added into Na$_2$HPO$_4$ solution is defined as t = 0.
Figure 4. Nucleation rates and ACP lifetimes in bulk solutions containing rP172 and enamelin as monitored by in situ Raman spectra and X-ray scattering. a) Number density of nuclei vs time and b) nucleation rate derived from (a) by linear fitting for a range of enamelin:rP172 ratios. c) Raman spectra of mineral at a series of time points in the absence of either rP172 or enamelin. Symmetric stretch ν₁ (PO₄) peak is used as a fingerprint of different calcium phosphate phases; the broad peak of ACP centered at 954 cm⁻¹ shifts to sharp peaks of OCP and HAP centered at 960 cm⁻¹. Spectra of four common calcium phosphate minerals, dicalcium phosphate dihydrate (DCPD), dicalcium phosphate anhydrous (DCPA), OCP and HAP, are shown for reference. d) The center of the symmetric stretch ν₁ (PO₄) peak at different times for a range of enamelin:rP172 ratios (protein concentration: 0 µg/mL enamelin and 25 µg/mL rP172 in 0:50, 0.5 µg/mL enamelin and 25 µg/mL rP172 in 1:50, 1 µg/mL enamelin and 25 µg/mL rP172 in 2:50, 2 µg/mL enamelin and 25 µg/mL rP172 in 4:50, 4 µg/mL enamelin and 25 µg/mL rP172 in 8:50, 7.5 µg/mL enamelin and 25 µg/mL rP172 in 15:50). Spectra dominated by ACP and HAP are colored blue and purple, respectively. e) The life time of ACP at different enamelin:rP172 ratios derived from d). f) Pair distribution functions (PDF) of mineral phase collected from solution at different times in the absence of rP172 or enamelin. The time when CaCl₂ solution is added into Na₂HPO₄ and protein mixed solution is defined as t = 0.
Figure 5. Simulation of random deposition of protein particles onto a surface in the absence of calcium phosphate. The number of green dots (representing monomers) out of the total number of green dots in a Grand Canonical Monte Carlo (GCMC) simulation. 

a) Scheme of GCMC simulation in which the simulated box is in equilibrium of environment with controlled chemical potential. The displacement, insertion and deletion of green dots is judged by different energy criteria (see main text).

b) Equilibrium snapshots of configurations of green dots during four typical activities ($z = e^{\beta \mu / N}$), which is a measurement of the chemical potential $\mu$ (See experimental procedures for details).

c) The relationship between the total green dot coverage and green dot “monomer” coverage with respect to the green dot activity. 

d) The relationship between green dot “monomer” coverage and the derived green dot/white square weight ratio based on the properties of enamelin monomer (green dots) and rP172 oligomer (white squares) (The molecular weights of enamelin monomer and rP172 monomer are 32 kDa and 19.6 kDa,
respectively; the average aggregation number of rP172 monomer is 17.8 inside the oligomers with an average height of 2.4 nm for the weight ratio calculation)\textsuperscript{48}. Both (c) and (d) are derived with a lattice number of 50×50. Each point is the average of 200 frames in the ensemble.

References:

17. Wiedemann-Bidlack, F. B.; Benias, E.; Yamakoshi, Y.; Simmer, J. P.; Margolis, H. C.,


33. Uchida, T.; Tanabe, T.; Fukae, M.; Shimizu, M. Immunocytochemical and
immunochemical detection of a 32kDa nonamelogenin and related proteins in porcine

34. Gutierrez, S. J.; Chaves, M.; Torres, D. M.; Briceno, I. Identification of a novel mutation
in the enamalin gene in a family with autosomal-dominant amelogenesis imperfecta.

35. Fan, D.; Lakshminarayanan, R.; Moradian-Oldak, J. The 32 kDa enamelin undergoes

36. Bouropoulos, N.; Moradian-Oldak, J. Induction of apatite by the cooperative effect of

enamelin and amelogenin cooperate to regulate the growth morphology of octacalcium

38. Chen, C. L.; Bromley, K. M.; Moradian-Oldak, J.; DeYoreo, J. J., In Situ AFM Study of
Amelogenin Assembly and Disassembly Dynamics on Charged Surfaces Provides

39. Habraken, W. J.; Tao, J.; Brylka, L. J.; Friedrich, H.; Bertinetti, L.; Schenk, A. S.; Verch,
A.; Dmitrović, V.; Bomans, P. H.; Frederik, P. M.; Laven, J.; van der Schoot, P.;
Aichmayer, B.; de With, G.; DeYoreo, J. J.; Sommerdijk, N. A. Ion-association
complexes unite classical and non-classical theories for the biomimetic nucleation of

40. Yang, H; Fung S. Y.; Pritzker, M.; Chen, P. Modification of hydrophilic and hydrophobic


42. Malasics, A.; Gillespie, D.; Boda, D. Simulating prescribed particle densities in the

43. Nudelman, F.; Pieterse, K.; George, A.; Bomans, P. H.; Friedrich, H.; Brylka, L. J.;
Hilbers, P. A.; de With, G.; Sommerdijk, N. A. The role of collagen in bone apatite
formation in the presence of hydroxyapatite nucleation inhibitors. *Nat. Mater.* 2010, 9
1004.

Morse, D. E. Silicatein filaments and subunits from a marine sponge direct the

45. Maity, B.; Abe, S.; Ueno, T. Observation of gold sub-nanocluster nucleation within a

Simmer, J. P.; Hu, J. C. Altered enamelin phosphorylation site causes amelogenesis

47. Beniash, E.; Metzler, R. A.; Lam, R. S. K.; Gilbert, P., Transient Amorphous Calcium
Phosphate in Forming Enamel. *Journal of Structural Biology* 2009, 166, 133-143.


Supplementary Materials for

Control of calcium phosphate nucleation and transformation through interactions of enamelin and amelogenin exhibits the “Goldilocks effect”

Jinhui Tao¹, Andreas Fijneman¹,² Jiaqi Wan¹#, Saumya Prajapati³, Kaushik Mukherjee³, Alejandro Fernandez-Martinez⁴, Janet. Moradian-Oldak³* and James.J. De Yoreo¹,⁵*

¹Physical Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA 99352
²Laboratory of Materials and Interface Chemistry and Center of Multiscale Electron Microscopy, Department of Chemical Engineering and Chemistry, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
³University of Southern California, Center of Craniofacial and Molecular Biology, Los Angeles, CA, USA 90033
⁴Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTERRE, Grenoble 38000, France
⁵Department of Materials Science and Engineering, University of Washington, Seattle, WA, 98195

# Current address: College of Materials Science and Engineering, Qingdao University of Science and Technology, Qingdao 266042, China

*Correspondence:
joldak@usc.edu
james.deyoreo@pnnl.gov
Figure S1. *In situ* AFM images of protein adsorption on mica (001) for a range of enamelin:rP172 ratios: a) 0:50; b) 1:50; c) 2:50; d) 4:50; e) 8:50; f) Protein coverage vs time for the same range of enamelin:rP172 ratios; g) distribution of particle heights from first image in each condition: a) 5.1 min; b) 3.6 min; c) 3.6 min; d) 2.5 min; e) 4.5 min, which has largest number of monodisperse particles for data analysis. Scale bars are a-e: 200 nm; g: 20%.
Figure S2. *In situ* AFM observations of calcium phosphate nucleation and growth on rP172 and enamelin:rP172 films co-assembled on HOPG for enamelin:rP172 ratios of: a) 0:50, b) 1:50, c) 2:50, d) 4:50. e) The number density of nuclei for the same range of enamelin:rP172 ratios from the first image of each condition. All scale bars are 400 nm.

Figure S3. SEM images of mineral morphology after mineralization on rP172 surface for a) 35 min; b) 124 min.
Figure S4. Morphology and crystallinity of final stage mineral formed in bulk solutions with different enamelin and rP172 concentrations, as monitored by *ex situ* TEM and SAED a) in the absence of enamelin or rP172; b) 25 µg/mL rP172; c) 0.5 µg/mL enamelin; d) 0.5 µg/mL enamelin and 25 µg/mL rP172 with enamelin:rP172 ratio of 1:50; e) 4 µg/mL enamelin; f) 4 µg/mL enamelin and 25 µg/mL rP172 with enamelin:rP172 ratio of 8:50. The time when CaCl$_2$ solution is added into Na$_2$HPO$_4$ and protein mixed solution is defined as t = 0.
Figure S5. a) Schematic illustrating the liquid cell for Raman spectroscopy on an inverted optical microscope. The bottom of the liquid cell has a square pinhole (50 µm × 50 µm) for laser penetration, the dark region in the right inset is the water-air interface, and the bright particles are the nucleated ACP. b-g) Snapshots of ACP nucleation at in the absence of enamelin and rP172, which give an example of how to obtain the nucleation rate. The nucleation rates of ACP in the cases with different enamelin:rP172 ratios were measured in the same way. The particles are counted and labelled with small crosses. All scale bars are 4 µm.

Figure S6. a) The center of the symmetric stretch ν₁ (PO₄) peak at different times for a range of enamelin concentrations (no protein in 0:0, 0.5 µg/mL in 1:0, 1 µg/mL in 2:0, 2 µg/mL in 4:0, 4 µg/mL in 8:0, and 7.5 µg/mL in 15:0). Spectra dominated by ACP and HAP are colored blue and purple, respectively. b) The life time of ACP at different enamelin concentrations derived from a).
Figure S7. The equilibrium coverage of hypothetical enamelin monomers as represented by the green dots in Fig. 4a and b and the total coverage of hypothetical enamelin proteins (total green dots) vs hypothetical enamelin activity in bulk solution for two different monomer-monomer interactions. a) Attractive interaction of $1\ k_B T$ between nearest neighbors; b) repulsive interaction of $1\ k_B T$ between nearest neighbors.