Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 45 (2006) 43–53
Protein Expression & Purification
www.elsevier.com/locate/yprep

High yield of biologically active recombinant human amelogenin using the baculovirus expression system

Angela L. Taylor, Amir Haze-Filderman, Anat Blumenfeld, Boaz Shay, Leah Dafni, Eli Rosenfeld, Yoav Leiser, Eran Fermon, Yael Gruenbaum-Cohen, Dan Deutsch *

Dental Research Laboratory, Institute of Dental Sciences, Hebrew University—Hadassah Faculty of Dental Medicine, Hebrew University, Jerusalem, Israel

Received 29 March 2005, and in revised form 19 May 2005
Available online 21 June 2005

Abstract

The amelogenins are secreted by the ameloblast cells of developing teeth; they constitute about 90% of the enamel matrix proteins and play an important role in enamel biomineralization. Recent evidence suggests that amelogenin may also be involved in the regeneration of the periodontal tissues and that different isoforms may have cell-signalling effects. During enamel development and mineralization, the amelogenins are lost from the tissue due to sequential degradation by specific proteases, making isolation of substantial purified quantities of full-length amelogenin challenging. The aim of the present study was to express and characterize a recombinant human amelogenin protein in the eukaryotic baculovirus system in quantities sufficient for structural and functional studies. Human cDNA coding for a 175 amino acid amelogenin protein was subcloned into the pFastBac HTb vector (Invitrogen), this system adds a hexa-histidine tag and an rTEV protease cleavage site to the amino terminus of the expressed protein, enabling effective one-step purification by Ni^{2+}–NTA affinity chromatography. The recombinant protein was expressed in *Spodoptera frugiperda* (Sf9) insect cells and the yield of purified his-tagged human amelogenin ($rHAM^{+}$) was up to 10 mg/L culture. Recombinant human amelogenin ($rHAM^{+}$) was characterized by SDS–PAGE, Western blot, ESI-TOF spectrometry, peptide mapping, and MS/MS sequencing. Production of significant amounts of pure, full-length amelogenin opened up the possibility to investigate novel functions of amelogenin. Our recent in vivo regeneration studies reveal that the $rHAM^{+}$ alone could bring about regeneration of the periodontal tissues; cementum, periodontal ligament, and bone.

Keywords: Recombinant human amelogenin; Baculovirus; Insect cell culture; ESI-TOF; Mass spectrometry; MS/MS; Peptide mapping

The amelogenins play a major role in the biomineralization and structural organization of enamel [1–3]. The developing extracellular enamel matrix is comprised of a number of proteins of which the amelogenins constitute about 90% [4]. They are hydrophobic molecules that self-assemble into nanosphere structures, which are thought to be involved in regulating the ultrastructural organization of the developing enamel crystallite [5–8]. During enamel development and mineralization, the secreted amelogenins are lost from the tissue, along with most of the other enamel matrix proteins, due to sequential degradation by specific proteases. They are replaced by mineral ions, calcium, and phosphorus, which eventually results in fully mineralized, hard, and mature enamel [2,3,9]. The protein undergoes post-translational modifications and post-secretory processing [1,10–12]. These factors, as well as alternative mRNA splicing, give rise to a heterogeneous mixture of polypeptides in the enamel matrix [3].

The human amelogenin gene has been mapped to Xp22.1–p22.3 and Yp11.2 with 90% of the transcripts expressed from the X- and 10% from the Y-chromosome

* Corresponding author. Fax: +972 2 675 7307.
E-mail address: dddan@cc.huji.ac.il (D. Deutsch).

1046-5928/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2005.05.010

[13–15]. The gene contains 7 exons, which undergo alternative mRNA splicing. The most abundant isoform of the native protein secreted into the enamel matrix lacks the internal region encoded by exon 4 [15]. More recently, two additional exons downstream of exon 7 have been identified in a rare alternatively spliced RNA transcript of amelogenin [16,17]. Mutations in the X-chromosomal copy of the amelogenin gene have been associated with the hereditary disease amelogenesis imperfecta, which illustrates the importance of amelogenin in developing enamel [18,19]. Defective enamel formation has also been demonstrated by knock-out of amelogenin expression using antisense oligonucleotides [20], ribozymes [21], and a recent amelogenin null mouse which resulted in enamel characteristic of hypoplastic amelogenesis imperfecta [22].

The amelogenins were thought to be tissue-specific and exclusively expressed by the enamel producing ameloblast cells [6,23], however, various isoforms have since been found in the dentin matrix [24] and the associated odontoblast cells [25,26]. Recent reports suggest amelogenin expression in the periodontal ligament and Hertwig's epithelial root sheath of the tooth attachment apparatus [27–29]. The expression of amelogenin in different tissues suggests that amelogenin may be multifunctional. A major discovery that highlights a new role for enamel matrix proteins was the finding that the application of an enamel matrix protein extract to tooth root surfaces, in sites of diseased periodontium, promotes the regeneration of all the periodontal tissues [30,31]. Since the developing enamel matrix is a mélange of various isoforms of amelogenin and other enamel matrix proteins and their degradation products, proteases and other minor components ([6,32] and Deutsch et al. 2004, personal communication), the production of recombinant proteins free of other matrix elements is a valuable tool to investigate which components are responsible for the regeneration. Amelogenin was suggested to be such a candidate. In addition, low molecular mass amelogenin polypeptides have since been associated with cell signaling and have been suggested to have osteogenic potential [33,34].

The described heterogeneity of the matrix makes isolation and purification of sufficiently large quantities of any one of the amelogenin isoforms from the developing tissue, extremely challenging. This problem has been previously overcome by high expression of mouse amelogenin [35], porcine amelogenin [36], and human amelogenin [37,38] in prokaryotic bacterial (*Escherichia coli*) systems. Here, we describe the expression of a recombinant human amelogenin protein (rHAM$^+$)[1] in the eukaryotic baculovirus system for the first time, which has made possible the investigation of novel functions of amelogenin by providing relatively high amounts of pure amelogenin, free from contamination of other enamel matrix proteins.

Our current in vivo regeneration studies revealed that this recombinant amelogenin protein alone could bring about substantial regeneration of the periodontal tissues: cementum, periodontal ligament, and alveolar bone after induced periodontitis [29,39,40,58].

Materials and methods

Construction of the recombinant amelogenin baculovirus

The recombinant human amelogenin vector was constructed and purified according to standard recombinant DNA techniques [41] as follows: human amelogenin cDNA was amplified by PCR from a recombinant plasmid [37] containing human amelogenin X cDNA (GenBank Accession No. M86932), representing the most abundant amelogenin mRNA transcript in the developing enamel, which lacks exon 4 and codes for a 175 amino acid protein [15]. Specific oligonucleotide primers were designed to include restriction enzyme sites (underlined) *Sfo*I on the sense primer (5′…CTG AAG GGC GCC ATG CCT CTA CCA CCT CAT CCT G…3′) and *Eco*RI on the antisense primer (5′…TCG CCG GAA TTC TTA ATC CAC TTC CTC CCG CTT GGT…3′) for subcloning into the pFastBac HTb donor vector of the baculovirus expression system (Invitrogen). This system adds a hexa-histidine tag and an rTEV protease cleavage site to the amino terminus of the expressed protein. The PCR product was cloned into the donor vector and the recombinant vector was sequenced to confirm correct cDNA insertion using an automated DNA sequencer (ABI prizm 377, Perkin-Elmer, USA). The recombinant vector was then electrotransformed into DH10BAC bacterial cells (Invitrogen) for transposition of the human amelogenin cDNA into the genetically modified baculovirus (*Autographa californica*) genome (bacmid). Positive recombinant bacmids were used to transfect *Spodoptera frugiperda* (Sf9) insect cells for viral particle formation. All procedures for the production of viral particles were performed according to the manufacturer's manual (Bac-to-Bac, Invitrogen). Recombinant amelogenin baculovirus underwent one round of plaque purification and subsequently three rounds of amplification (48 h each) by infecting Sf9 monolayers in 25 cm^2 and then 75 cm^2 T-flasks to generate 100 ml of high titre virus (3×10^7 to 3×10^8 pfu/ml as determined by plaque assay) for rHAM$^+$ expression.

Sf9 cell maintenance

Sf9 cells adapted to serum-free Bioinsect-1 medium containing 100 U/ml penicillin, 0.1 mg/ml streptomycin,

[1] *Abbreviations used:* rHAM$^+$, recombinant human amelogenin protein; ESI-TOF/MS, electrospray ionization time-of-flight mass spectrophotometry; CID, collision-induced dissociation; MOI, multiplicity of infection.

and 0.25 μg/ml amphotericin were incubated at 27 °C (±1 °C) in either monolayer or suspension cultures. All media components were obtained from Biological Industries (Israel). Suspension cultures were grown at a constant orbital shaking speed of 97 rpm. Stock cells grown in suspension were seeded at 0.8–1.5 × 10^6 cells/ml and the cells were passaged every 2–3 days while still in log phase (4–6 × 10^6 cells/ml). Viable cell counts were determined by Trypan blue exclusion (Sigma) and cultures with viabilities below 97% were discarded.

Optimization of recombinant amelogenin expression

For optimization assays, multiplicity of infection (MOI) between 0.1 and 15 was used. Cells were harvested at regular intervals (24, 48, 72, and 96 h post-infection), and the total and viable cell counts were monitored. Crude extracts were prepared by resuspension of the cell pellets in lysis buffer (50 mM Tris, pH 7.5, 1% (v/v) Triton X-100, and 1 mM PMSF), sonication on ice (3× 10 s for 1 ml lysate, at 40% of maximum energy output (Vibra-Cell; Sonics, USA)), and centrifugation at 15,000 rpm for 15 min. The supernatants (crude extracts) were analysed by SDS–PAGE and Western blot for recombinant amelogenin expression.

Optimized expression conditions were as follows: 400 ml of log phase Sf9 cells (3–4 × 10^6 cells/ml) with viability above 98% and final concentration of 3 × 10^6 cells/ml were infected at MOI of 0.5 pfu/cell with recombinant human amelogenin baculovirus and 0.5% (v/v) fetal calf serum was also added. The percentage cell viability was monitored every 24 h and the infected cells were harvested at 48–72 h post-infection when cells were still mostly viable (between 60 and 90%). Harvested cells were washed twice in 1× PBS containing 1 mM PMSF, at 1000 rpm for 5 min at 4 °C, and the pellet was stored at −70 °C until protein purification.

Purification of the recombinant human amelogenin protein

Cell pellets were resuspended in 1:10 volumes lysis buffer (6 M GuHCl; 0.1 M NaH_2PO_4 pH 7.5; 10 mM imidazole; and 0.5% (v/v) protease inhibitor cocktail (P8849; Sigma)). Sonication was performed on ice for 5× 10 s per 10 ml lysate at 40% of maximum energy output (Vibra-Cell; Sonics, USA). The lysate was cleared by centrifugation at 11,000 rpm for 30 min at 4 °C. The supernatant (crude extract) was incubated batchwise with pre-equilibrated Ni–NTA^{2+} resin (Qiagen) at a ratio of 1 ml beads/8 ml lysate, rotating constantly (∼15 rpm) at 4 °C for 90 min. The unbound proteins were removed by spinning the resin for 5 min at 4000 rpm, and the resin was packed into a column.

Chromatography was performed on an ÄKTA Explorer FPLC system (Amersham Pharmacia), and the absorption (280 nm) and conductivity of the eluent was monitored. The column was washed at a flow rate of 2 ml/min with 5–10 column volumes (CV) of lysis buffer, 20CV wash buffer 1 (6 M urea; 0.1 M NaH_2PO_4; 0.01 M Tris–HCl; and 10 mM imidazole, pH 7.0); 10CV wash buffer 2 (wash buffer 1 at pH 6.3); 5CV wash buffer 2 adding 5% (v/v) elution buffer (6 M urea; 0.1 M NaH_2PO_4; 0.01 M Tris–HCl; and 500 mM imidazole; pH 4.8), and eluted on a 35–500 mM imidazole gradient using the elution buffer at a flow rate of 0.5 ml/min. Alternatively, 6 M urea was replaced with 4 M GuHCl in the wash and elution buffers. The eluted protein was analysed by SDS–PAGE and visualized by Coomassie brilliant blue R250 (Sigma). Purified protein was dialysed (MWCO 3500, SnakeSkin; Pierce, USA) against 50 mM acetic acid at 4 °C and the dialysis buffer was exchanged every 4 h for 2–3 days, after which the $rHAM^+$ was aliquoted and either lyophilized or fast-freezed in liquid nitrogen and stored at −70 °C. Protein concentrations were determined by the Bradford protein assay using Bio-Rad Protein Assay Reagent [42]. Histidine tag cleavage was performed using a histidine tagged rTEV protease (Invitrogen) according to the manufacturer's protocol.

SDS–PAGE and Western blot analysis

The recombinant amelogenin was analysed by SDS–PAGE (2–40% Tris–glycine precast gels, Invitrogen, USA) and visualized with Coomassie blue. Proteins were transferred onto nitrocellulose and the membranes were blocked in Superblotto (10% (v/v) PBS; 1 M glucose; 3% (w/v) BSA; 1% (w/v) milk powder; 10% (v/v) glycerol; and 0.5% (v/v) Tween 20) shaking at 37 °C for 1 h. The membranes were then incubated shaking overnight at room temperature with the following antibodies diluted in 1% (w/v) gelatine–TBS: rabbit polyclonal (1:750–1:1000) or mouse monoclonal (1:100–1:300) antibodies against human [43] and bovine amelogenin [44]; rabbit polyclonal antibodies (1:1000, LF109-CSA or LF108-LPH) against synthetic peptides to 10 C-terminal amino acids of human amelogenin, kindly provided by Dr. Larry Fischer (NIDCR, NIH, USA). In addition, mouse monoclonal antibody to hexa-histidine (1:1000, MMS-156P, Babco-Convance, USA) was used. Alkaline phosphatase-conjugated secondary antibodies against polyclonal and monoclonal antibodies were anti-rabbit IgG (diluted 1:10,000) and anti-mouse IgG (diluted 1:7500), respectively (Promega, USA). Visualization was performed with BCIP-NBT reagents (Promega, USA).

Electrospray ionization time-of-flight mass spectrometry (ESI-TOF/MS)

Mass spectra were conducted in positive ion mode on a Q-TOF II mass spectrometer (Micromass, UK) equipped with a nano-electrospray ion source. The

rHAM$^+$ samples (10–50 pmol in 50 mM acetic acid) were loaded onto reverse phase ZipTip$_{C4}$ microcolumns (Millipore, USA) and eluted with 50% (v/v) acetonitrile and 1% (v/v) formic acid. The mass estimation of the proteins was performed by deconvolution of raw multicharged peaks using the MaxEnt1 algorithm (MassLynx package, MaxEnt Solutions, UK).

Enzymatic peptide mapping and MS/MS sequencing

rHAM$^+$ was cleaved in solution using Asp-N endoproteinase, at pH 8.0 in 25 mM ammonium bicarbonate, at 37 °C overnight. The samples were then desalted on reverse phase ZipTip$_{C18}$ microcolumns. The peptides were injected into the ESI mass spectrometer for mass determination and for MS/MS sequencing by collision-induced dissociation (CID). The processing of the MS/MS data was performed using MaxEnt3 (MassLynx package, UK) supplied by Micromass. MS/MS data were searched against the NCBI non-redundant database using MASCOT (Matrix Science, London). The PepSeq application (BioLynx, MassLynx package, UK) was used to interpret the data and for assignment of sequence tags. The amino acid sequence tag data were compared and mapped to the recombinant human amelogenin sequence.

Results and discussion

Expression strategy for recombinant human amelogenin

Amelogenin is a hydrophobic enamel matrix protein that possesses the unique physiological characteristic of self-aggregation into nanospheres, an important structural feature related to its function in the developing enamel; but one that makes it challenging to express, isolate, and analyse. Expression of a recombinant amelogenin protein has only been reported in prokaryotic systems so far [35–38]. These recombinant amelogenin proteins have been used in various structural and functional studies [5,45]. We chose to express a recombinant human amelogenin protein in eukaryotic Sf9 insect cells using a recombinant baculovirus, containing the human amelogenin cDNA fused to a sequence for a cleavable hexa-histidine tag for effective, one-step purification (see Materials and methods). The baculovirus expression system has a number of potential benefits, including high levels of expression, correct folding and post-translational modification, and production of biologically active proteins for analysis. In addition, since one of our main reasons for producing the recombinant amelogenin protein was for in vivo studies, we preferred to use the baculovirus system, to avoid the necessity for testing and removal of endotoxin that may contaminate *E. coli* products. One of the potential disadvantages of the baculovirus system is that it is a lytic system and various proteases are released into the medium by the insect host cells, and therefore proteolysis could be a problem for a secreted recombinant amelogenin protein already susceptible to proteolysis. Due to the potential problems of recombinant amelogenin aggregation and proteolysis we preferred to use a non-secreting, baculovirus system.

The rHAM$^+$ corresponds to the most abundant human amelogenin mRNA transcript in the developing enamel, coding for the 175 amino acid human amelogenin described by Salido et al. [15]. This isoform lacks the 14 amino acids encoded by exon 4 (in most species almost all amelogenin mRNA transcripts expressed in developing enamel lack exon 4) [34] and has an additional 25 amino acids at the N-terminal which contain a hexa-histidine tag, a spacer region, and an rTEV protease cleavage site. Fig. 1 illustrates the N-terminal region of the corresponding recombinant construct after DNA sequencing, which shows correct in-frame positioning of the human amelogenin cDNA with respects to the start codon (ATG) of the vector. The correct in-frame protein translation was later confirmed by MS/MS sequencing (Figs. 6 and 7).

Optimization of recombinant protein expression

Recombinant amelogenin protein expression was optimized to find reproducible conditions and to curb protein degradation; the batch size, MOI, and time of harvest all proved to be critical factors. At high MOI (between 15 and 30), the purified yield of rHAM$^+$ was relatively high, however, significant protein degradation was evident (Fig. 2A). Purified rHAM$^+$ analysed by SDS–PAGE, contained a protein band with an apparent molecular weight of ~32 kDa and an additional array of lower molecular weight bands in the 16–6 kDa region.

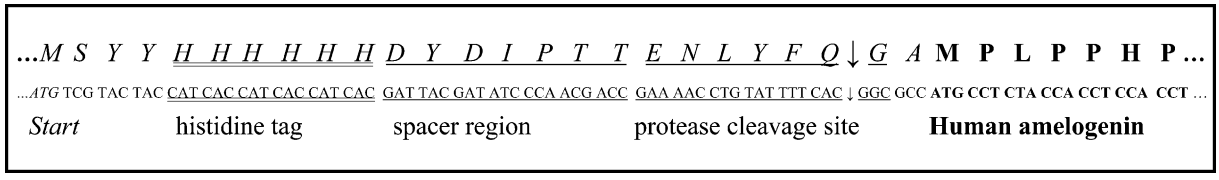

Fig. 1. Structure of the recombinant vector coding for the recombinant human amelogenin protein. It contains a hexa-histidine tag, spacer region, and an rTEV protease cleavage site (underlined and italics) followed by the human amelogenin sequence (bold). The arrow indicates the rTEV cleavage site between glutamine (Q) and glycine (G).

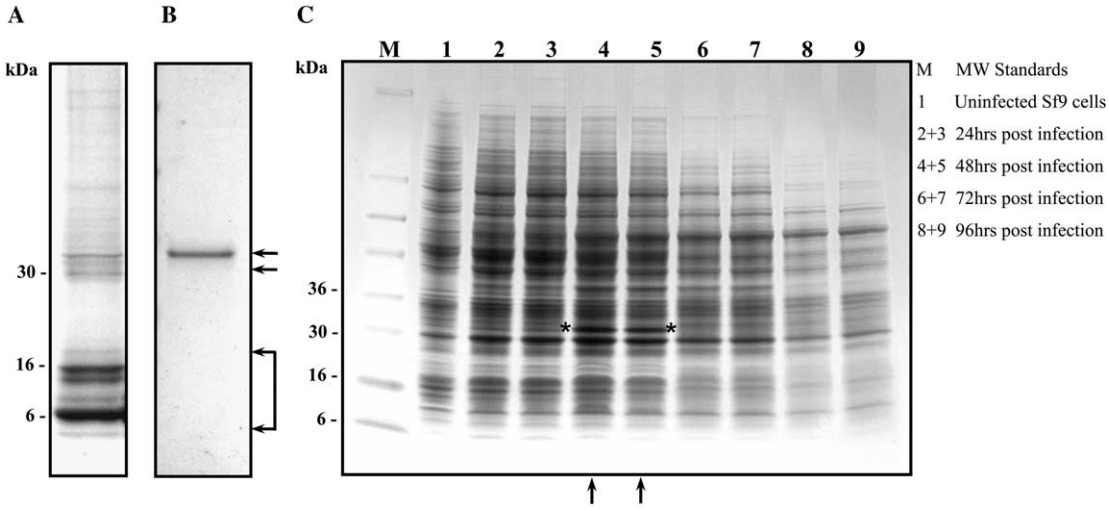

Fig. 2. Optimization of recombinant human amelogenin (rHAM$^+$) expression in Sf9 cells. (A) Purified rHAM$^+$ before optimization of infection. After 72 h of infection at MOI >15, the cell viability was <50% and most of the full-length rHAM$^+$ protein was degraded. (B) Purified rHAM$^+$ after optimization of infection. To increase the yield of intact protein, a lower MOI (0.5) was used and cells were harvested at 48 h when cells were 95% viable. Arrows to the right indicate the position of full-length rHAM$^+$ (~32 kDa) and its degradation products mainly in the 16–6 kDa region. (C) Time-course for rHAM$^+$ expression. Forty millilitre Sf9 cell suspension cultures were infected at MOI 0.5 and cells were harvested every 24 h. A band around 32 kDa appeared (*) at 48 h post-infection, which was absent from the uninfected cells. Arrows below indicate the preferred time of harvest (lanes 4 and 5, 48 h). Analyses were performed on 4–20% SDS–PAGE and visualized with Coomassie blue.

Protein degradation was evident at high MOI, in serum-free conditions and when cell viability was low (<50%). For optimization of full-length rHAM$^+$ expression, MOI between 0.1 and 15 was tested in time-course experiments. Cells were harvested at regular intervals (24, 48, 72, and 96 h) and the total and viable cell counts were monitored. Analysis of these crude extracts by SDS–PAGE revealed a narrow window of maximum full-length amelogenin expression, which was most evident at MOIs below 5 (Fig. 2C) (only the results for MOI 0.5 are shown). The optimized conditions for full-length rHAM$^+$ were as follows: MOI 0.5, 0.5% (v/v) FCS (which provides a competitive substrate and contains protease inhibitors) and harvesting between 48 and 72 h post-infection when cells were mostly viable. Under these conditions expression was 4–10 mg/L and the protein remained mostly intact (Fig. 2B). The baculovirus expression system used does not require multiple rounds of plaque purification due to the absence of parental virus DNA during site-specific transposition to generate recombinant bacmid DNA [Bac-to-Bac, BEVS Manual, Invitrogen]. However, initially we obtained variable results even when using identical infection conditions, e.g., initial batch volume and concentration, time-of-infection, MOI, and harvesting time. Accordingly, one round of plaque purification was performed as suggested in the manufacturer's troubleshooting, which amplifies recombinant virus from a single viral clone in order to eliminate any non-recombinant virus that may be present. One round of plaque purification was found to improve the reproducibility of the results.

Purification and identification of the recombinant human amelogenin protein

A high purity of rHAM$^+$ was achieved by one-step, metal affinity chromatography (NTA–Ni^{2+}), as visualized by Coomassie blue-stained SDS–PAGE gels (Figs. 2B, 3A, and 4B) and later confirmed by MS (Fig. 5). In this study, various obstacles to protein purification were encountered. Under native conditions rHAM$^+$ aggregated and was eluted in the unbound fraction (results not shown). Purification was therefore performed batch-wise under denative conditions to prevent amelogenin aggregation and to expose the histidine tag. Sonication proved to be an essential step to reduce the viscosity and ensure clarification of the lysates before addition to the nickel resin. Purification was also improved by the addition of up to 30 mM imidazole in the wash buffers and reduction of the wash and elution buffer pH. Usually the addition of 500 mM imidazole would be sufficient for elution of a histidine tagged protein but lowering the pH improved the elution profile, probably because amelogenin is more soluble at lower pH [35,46]. Purified rHAM$^+$ was dialysed against 50 mM acetic acid to maintain solubility [47].

Various monoclonal and polyclonal antibodies against human and bovine amelogenin, and against hexa-histidine were used to further characterize the rHAM$^+$. All these antibodies reacted strongly with the full-length rHAM$^+$ at about 32 kDa (for examples—see Fig. 3). The presence of the C-terminal portion of amelogenin was confirmed by Western blot using antibodies

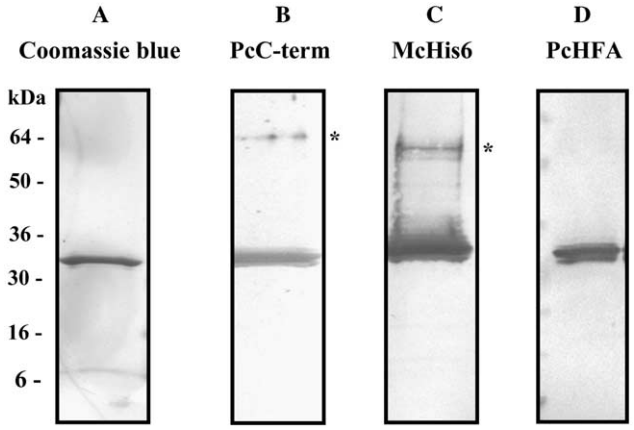

Fig. 3. Western blot detection by several antibodies of the affinity purified full-length rHAM$^+$ with minimal degradation. (A) Coomassie blue-stained gel. (B) Polyclonal antibody against 10 C-terminal amino acids of human amelogenin (PcC-term, LF109). (C) Monoclonal antibody against hexa-histidine (McHis6, Babco-Convance). (D) Polyclonal antibody against human amelogenin (PcHFA). Aggregates in the higher molecular weight region are marked by an asterisk (∗). Analyses were performed on 4–20% SDS–PAGE.

directed against 10 C-terminal amino acids of human amelogenin (Figs. 3B and 7). Antibody to polyclonal human amelogenin (PcHFA) also reacted with the degradation products in the 16–6 kDa region (Fig. 4). The amelogenin protein is degraded by specific proteases from the C-terminal during enamel development as observed in a number of species [6,48–50]. Occasionally

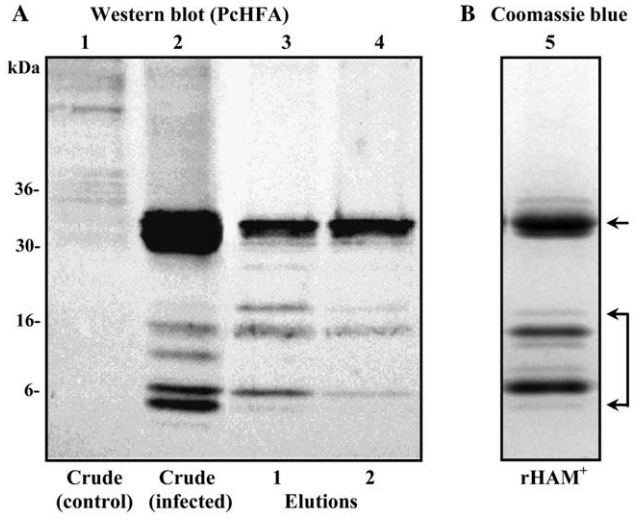

Fig. 4. Purification of rHAM$^+$ from recombinant baculovirus-infected Sf9 cells. (A) Western blot using polyclonal antibody to human amelogenin (pcHFA) showed that the antibodies reacted with full-length rHAM$^+$ (~32 kDa) and its degradation products (in the 16–6 kDa region); (1) crude extract of uninfected (control) cells, (2) crude extract of recombinant baculovirus infected cells, (3, 4) elutions of rHAM$^+$ purified on Ni^{2+}–NTA. (B) (5) Coomassie blue-stained gel. Purified recombinant human amelogenin (rHAM$^+$). Arrows indicate the position of full-length amelogenin and its degradation products. Analyses were performed on 4–20% SDS–PAGE.

aggregation of the rHAM$^+$ could be seen in the higher molecular weight region of the gels (Figs. 3B and C, marked by an asterisk). Aggregation of the rHAM$^+$ is most probably a consequence of amelogenin self-assembly [5,7,8,51].

Histidine tag removal

The histidine tag enables effective, one-step purification of the recombinant human amelogenin on NTA–Ni^{2+} resin. This hydrophilic tag is situated at the hydrophobic N-terminus of amelogenin. Removal of the histidine "arm" by protease cleavage (rTEV) was possible but not efficient; a number of studies using other recombinant histidine-tagged proteins have also reported a low cleavage efficiency and have shown that the histidine tag did not affect protein activity [52–55]. Our current functional studies, using the recombinant protein with the histidine-tag arm, showed that our rHAM$^+$ was biologically active ([29,39,40,56] and Haze-Filderman et al. (2005), in submission).

Recombinant human amelogenin characterization by mass spectrometry

Peptide mapping and MS/MS sequencing using ESI-TOF/MS

To identify and characterize the rHAM$^+$, we combined ESI mass spectral data of proteolytic peptides with their MS/MS fragmentation patterns (Table 1). Recombinant amelogenin was cleaved with Asp-N and the resulting MS peptide data were compared to masses predicted from the rHAM$^+$ sequence. Fig. 5 shows the multiply charged ESI spectrum of the Asp-N cleaved rHAM$^+$ peptides. Ten main peaks were matched to eight predicted (calculated) rHAM$^+$ peptide masses (labelled pepides 1–8). The rHAM$^+$ peptide sequences were determined by analysis of their fragmentation masses using tandem MS/MS scans on the most abundant ions in the MS spectrum (e.g., Fig. 6). The identified 20 amino acid sequence ^{13}DIPTTENLYFQGAMPLPPHP^{32} of peptide 1 ($[M+3H]^{+3}$ *m/z* 1125.27) is unique to the rHAM$^+$ protein spanning part of the spacer region of the his-tag arm, the protease cleavage site and the first seven amino acids of the human amelogenin protein (Figs. 6 and 7). It confirmed that the recombinant protein is in-frame with the first methionine of the human amelogenin sequence, as illustrated in Fig. 1. MPLPPHP corresponds to the beginning of the mature amelogenin protein and is conserved throughout most species. A summary of the position of these peptides in the rHAM$^+$ sequence can be seen in Fig. 7 and Table 1.

Upon examination of the MS data to verify if any post-translational or chemical modifications had occurred, by searching for deviations of the measured mass from the calculated peptide masses, no unambiguous indication of

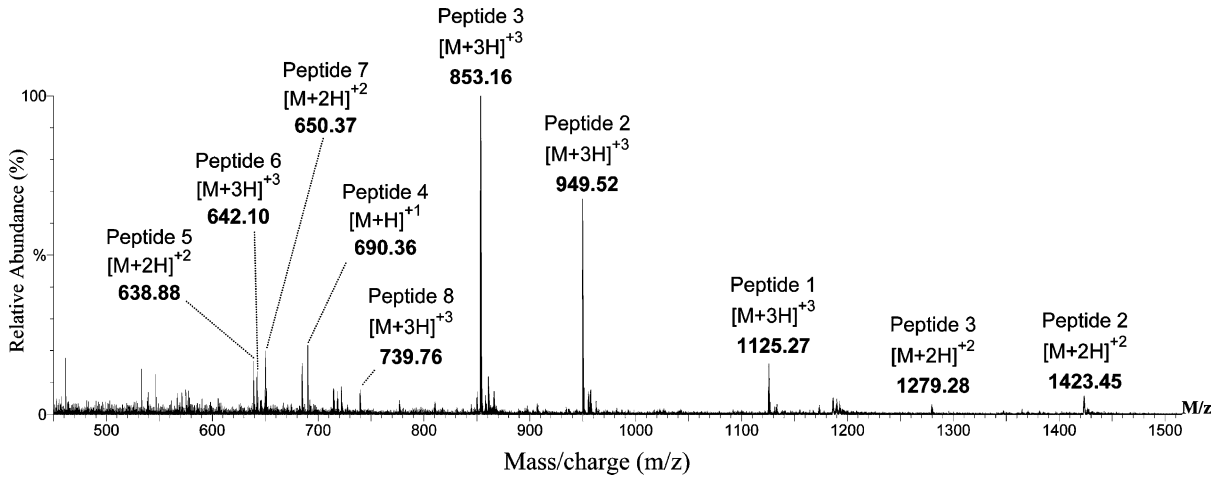

Fig. 5. ESI/MS peptide map of recombinant amelogenin (rHAM$^+$) peptides generated by Asp-N digestion. Ten main peaks corresponding to eight peptides (labelled peptides 1–8) were matched to predicted rHAM$^+$ peptide masses (see Table 1). The peptide ions were then subjected to MS/MS to obtain amino acid sequence information (see Fig. 6). Multiply charged ESI mass spectrum with signals for the protonated molecule ions $[M+H]^{+1}$, $[M+2H]^{+2}$, and $[M+3H]^{+3}$ at *m/z* range 400–1500. Peptides 2 and 3 appear as triply $[M+3H]^{+3}$ and doubly $[M+2H]^{+2}$ charged peptide ions.

phosphorylation was observed. All the detected peptides were within <1 Da of their predicted masses (Fig. 5 and Table 1). Analysis of phosphorylation by mass spectrometry is often challenging because of the low abundance, and low stoichiometry of phosphopeptides as compared to non-phosphorylated peptides; and other technical issues related to mass spectrometry [57]. A very small percentage of the rHAM$^+$ may have been phosphorylated by the baculovirus system and the purification and MS procedures may have stripped off the phosphates and are therefore not easily detectable. The identification of specific sequences and peptide masses for recombinant human amelogenin implies correct synthesis of the protein.

Full-length mass determination by ESI-TOF/MS

The full-length mass of rHAM$^+$ was 22942.0 Da (±1.5 Da, $n=3$) as measured by ESI-TOF/MS, which is about 90 Da less than the calculated mass (23032.46 Da) (Fig. 8). Examination of the observed proteolytic peptide masses did not account for the observed mass shift, as all of the eight detected peptides were within <1 Da of their predicted masses (Fig. 5 and Table 1). An array of peaks in the higher mass region to the main peak are indicative of water (18.015 Da) or sodium (22 Da) adducts. No significant difference in mass was noted between rHAM$^+$ purified in GuHCl alone or in GuHCl/urea (see Materials and methods). The apparent mass (~32 kDa) of rHAM$^+$ on SDS–PAGE (Figs. 2–4) is higher than the actual mass (~23 kDa) as measured by MS. This anomalous migration is characteristic of amelogenin and has been well documented [4,49].

Conclusion

This article describes the production, purification, and characterization of a rHAM$^+$ expressed in the

Table 1
Recombinant human amelogenin (rHAM$^+$) peptide masses and sequences detected by ESI/MS and MS/MS

Peptide number	Observed mass (MS^{ESI})	Calculated mass[a]	Charge	Δ mass (Da)	Peptide sequence[b]	Position in rHAM$^+$	Missed cleavage
Matching peptides for specific cleavage of rHAM$^+$ by Asp-N							
1	1125.27	1125.20	$[M+3H]^{+3}$	0.07	(Y)/**DIPTTENLYFQGAMPLPPHP**GHPGYINFSY/(E)	13–42	1
2	1423.45	1423.67	$[M+2H]^{+2}$	−0.22	(T)/**ENLYFQGAMPLPPHP**GHPGYINFSY/(E)	18–42	0
	949.52	949.45	$[M+3H]^{+3}$	0.07			
3	1279.28	1279.16	$[M+2H]^{+2}$	0.12	(Y)/**EVLTPLKWYQSIRPPYPSYGY**/(E)	43–63	0
	853.16	853.11	$[M+3H]^{+3}$	0.05			
4	690.36	690.31	$[M+H]^{+1}$	0.05	(L)/**EAWPST**/(D)	186–191	0
Matching peptides for non-specific enzyme cleavage							
5	638.88	638.85	$[M+2H]^{+2}$	0.03	(Y)/**EVLTPLKWYQ**/(S)	43–52	0
6	642.01	641.98	$[M+3H]^{+3}$	0.03	(A)/**MPLPPHPGHPGYINFSY**/(E)	26–42	0
7	650.37	650.32	$[M+2H]^{+2}$	0.05	(Q)/**SIRPPYPSYGY**/(E)	53–63	0
8	739.76	738.84	$[M+3H]^{+3}$	0.92	(L)/TPLKWYQS**IRPPYPSYGY**/(E)	46–63	0

[a] Peptide masses were calculated using the monoisotopic masses of the amino acid residues and are presented as mass/charge (*m/z*).
[b] Amino acids sequenced by tandem MS/MS are in bold. Asp-N cleaves before aspartic acid (D) and glutamic acid (E). Cleavage sites are indicated by "/."

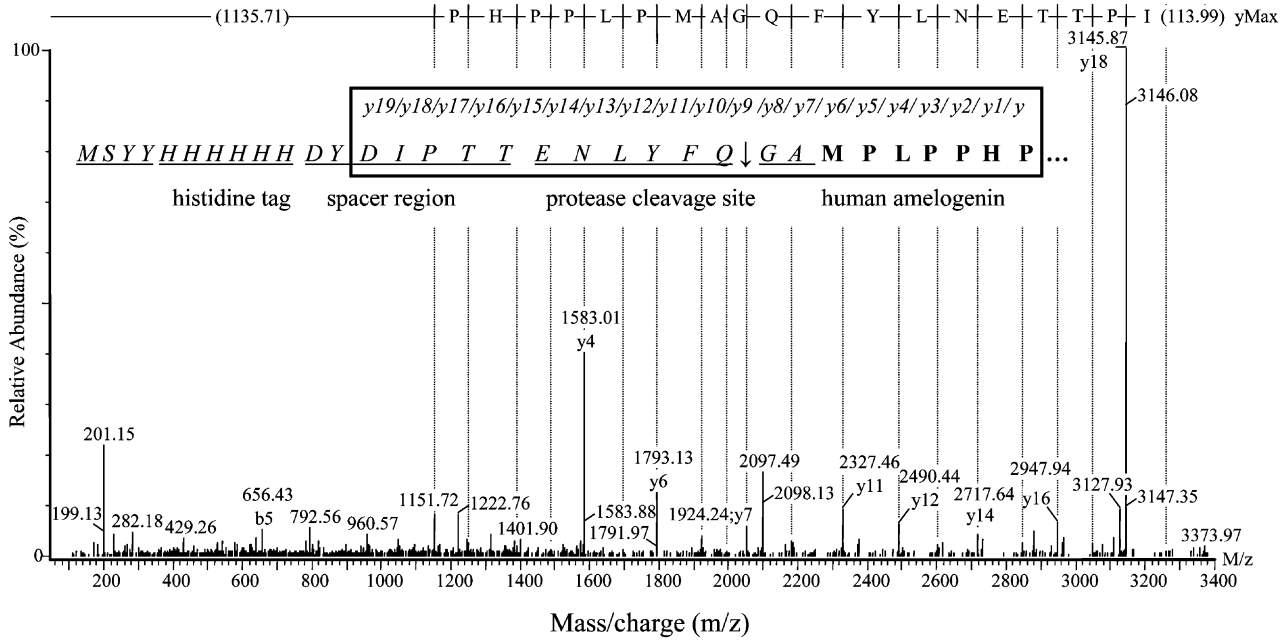

Fig. 6. MS/MS product spectrum of the precursor ion *m/z* 1125 (peptide 1 ($[M+3H]^{+3}$) in Fig. 5, equivalent to 3373.60 Da) at *m/z* range of 100–3400. (Inset) The identified amino acid sequence *y19 to y* is boxed and represented in reference to Fig. 1. The underlined amino acids at the N-terminus represent the additional recombinant sequence including the hexa-histidine.

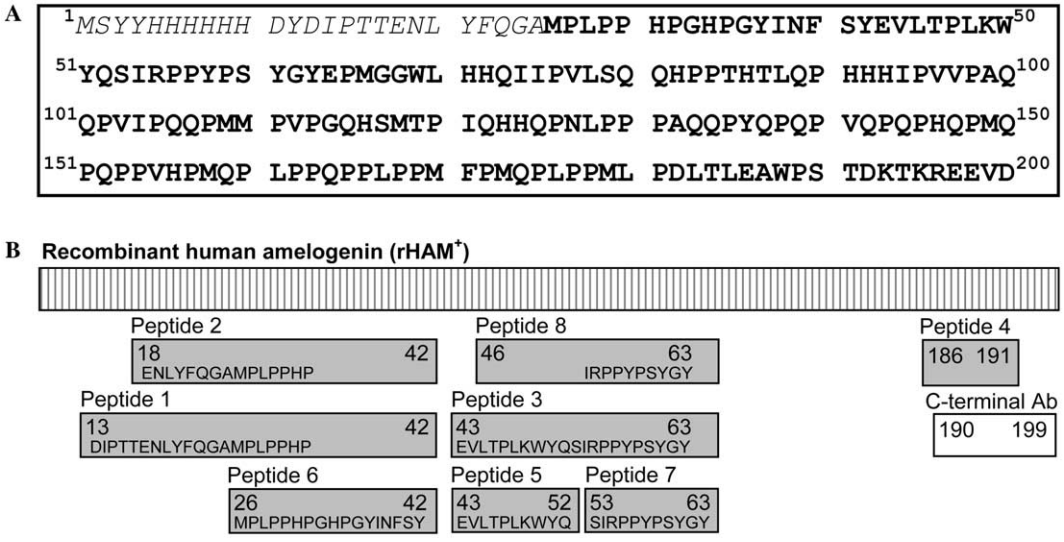

Fig. 7. (A) Amino acid sequence of the recombinant human amelogenin (rHAM$^+$) protein. The recombinant hexa-histidine arm at the N-terminus is in italics and the human amelogenin sequence is in bold. (B) Summary of rHAM$^+$ characterization by ESI/MS and MS/MS. rHAM$^+$ was cleaved with endoproteinase Asp-N (peptides 1–8) and the molecular masses of the peptides were measured by ESI/MS (see Fig. 5). Bars represent the locations of overlapping peptides with numbers indicating their position in the rHAM$^+$ sequence (the peptide bars are not to scale). Sequences obtained by MS/MS are printed inside the corresponding peptide bars. The presence of the C-terminus was confirmed using antibodies directed against 10 C-terminal amino acids of human amelogenin (represented by a clear bar).

eukaryotic baculovirus expression system. This histidine-tagged system enabled high purification through a single-step isolation. The yield of purified protein was 4–10 mg/L culture of mainly full-length protein, or full-length protein with some degradation products. Optimization of infection and purification conditions was essential to obtain a substantial, reproducible yield of full-length recombinant amelogenin. ESI mass spectrometry was successfully employed to identify and characterize the recombinant amelogenin and to assess the homogeneity and purity of the protein. As we described in the introduction, the expression of amelogenin in different tissues other than enamel, such as the Hertwig's epithelial root sheath, periodontal ligament, cementum,

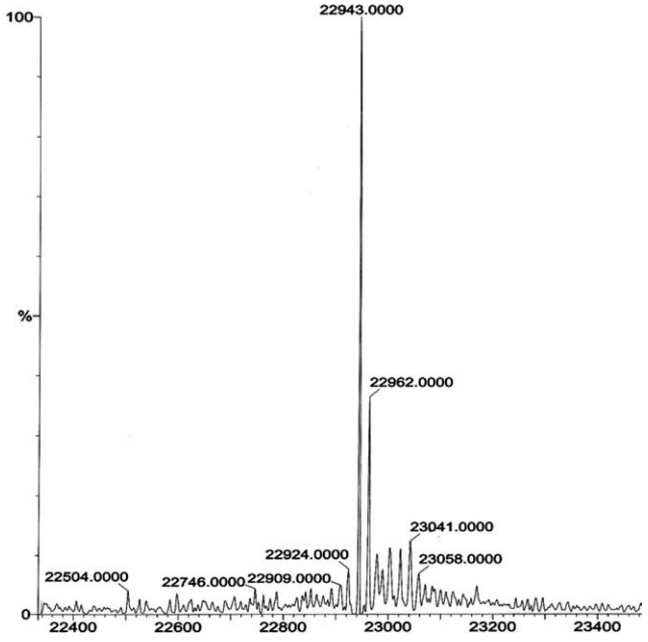

Fig. 8. ESI-TOF/MS deconvoluted mass spectrum of full-length rHAM$^+$. There is a main peak of 22943 Da and an array of smaller peaks of higher mass that differ by amounts indicative of water (18.015 Da) or sodium adducts (+22 Da).

and dentin suggests that amelogenin may have additional roles other than being a structural enamel matrix protein. Application of an enamel matrix protein extract, a mélange of many enamel matrix proteins and their degradation products including the predominant amelogenin protein, to tooth root surfaces in sites of diseased periodontium has been shown to promote the regeneration of all the periodontal tissues. However, it was not clear which of the proteins comprising the extracellular matrix is responsible for this regeneration. Our recent results showed, for the first time, that a single rHAM$^+$ and its degradation products (e.g., Fig. 4), which was free of other enamel matrix components, was able to induce the regeneration of all the periodontal supporting tissues (alveolar bone, cementum, and periodontal ligament) that were destroyed after induced periodontitis in the dog model [29,39,40,58]. The availability of substantial amounts of purified recombinant human amelogenin expressed in eukaryotic cells provides a valuable and reproducible source of biologically active human amelogenin protein. This recombinant protein is currently being used to study the underlying mechanisms associated with the function(s) of the human amelogenin protein in normal, diseased, and regenerating periodontal tissues and other mineralizing tissues.

Acknowledgments

We are grateful to Dr. Tsafi Danieli and Dr. Mario Lebendiker (Protein Expression Facility, Wolfson Center for Applied Structural Biology, Life Sciences Institute, Hebrew University of Jerusalem) for their valuable technical assistance and helpful advice; to Dr. Larry Fisher (NIDCR, NIH, USA) for providing polyclonal antibodies against the C-terminal of the human amelogenin protein (LF109-CSA and LF108-LPH); and to Dr. Ofra Moshel (Interdepartmental Equipment Services, Hebrew University of Jerusalem, Ein Kerem) for her advice and assistance with the mass spectrometric analyses.

References

[1] A.G. Fincham, J. Moradian-Oldak, Recent advances in amelogenin biochemistry, Connect. Tiss. Res. 32 (1995) 119–124.
[2] C. Robinson, S.J. Brookes, R.C. Shore, J. Kirkham, The developing enamel matrix: nature and function, Eur. J. Oral Sci. 1 (1998) 282–291.
[3] D. Deutsch, J. Catalano-Sherman, L. Dafni, S. David, A. Palmon, Enamel matrix proteins and ameloblast biology, Connect. Tissue Res. 32 (1995) 97–107.
[4] J.D. Termine, A.B. Belcourt, P.J. Christner, K.M. Conn, M.U. Nylen, Properties of dissociatively extracted fetal tooth matrix proteins. I. Principal molecular species in developing bovine enamel, J. Biol. Chem. 255 (1980) 9760–9768.
[5] A.G. Fincham, J. Moradian-Oldak, J.P. Simmer, P. Sarte, E.C. Lau, T. Diekwisch, H.C. Slavkin, Self-assembly of a recombinant amelogenin protein generates supramolecular structures, J. Struct. Biol. 112 (1994) 103–109.
[6] A.G. Fincham, J.P. Simmer, Amelogenin proteins of developing dental enamel, Ciba Found Symp. 205 (1997) 118–134.
[7] J. Moradian-Oldak, Amelogenins: assembly, processing and control of crystal morphology, Matrix Biol. 20 (2001) 293–305.
[8] M.L. Snead, Amelogenin protein exhibits a modular design: implications for form and function, Connect. Tissue Res. 44 (2003) 47–51.
[9] C.E. Smith, Cellular and chemical events during enamel maturation, Crit. Rev. Oral Biol. Med. 9 (1998) 128–161.
[10] A. Fincham, J. Moradian-Oldak, Amelogenin post-translational modifications: carboxy-terminal processing and the phosphorylation of bovine and porcine "TRAP" and "LRAP" amelogenins, Biochem. Biophys. Res. Commun. 197 (1993) 248–255.
[11] A.G. Fincham, J. Moradian-Oldak, P.E. Sarte, Mass-spectrographic analysis of a porcine amelogenin identifies a single phosphorylated locus, Calcif. Tissue Int. 55 (1994) 398–400.
[12] E. Salih, J.C. Huang, E. Strawich, M. Gouverneur, M.J. Glimcher, Enamel specific protein kinases and state of phosphorylation of purified amelogenins, Connect. Tissue Res. 38 (1998) 225–235 , discussion 241–246.
[13] E.C. Lau, T.K. Mohandas, L.J. Shapiro, H.C. Slavkin, M.L. Snead, Human and mouse amelogenin gene loci are on the sex chromosomes, Genomics 4 (1989) 162–168.
[14] A.G. Fincham, C.C. Bessem, E.C. Lau, Z. Pavlova, C. Shuler, H.C. Slavkin, M.L. Snead, Human developing enamel proteins exhibit a sex-linked dimorphism, Calcif. Tissue Int. 48 (1991) 288–290.
[15] E.C. Salido, H.Y. Paulie, K. Koprivnikar, L.C. Yu, L. Shapiro, The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes, Am. J. Hum. Genet. 50 (1992) 303–316.
[16] W. Li, C. Mathews, C. Gao, P.K. DenBesten, Identification of two additional exons at the 3′ end of the amelogenin gene, Arch. Oral Biol. 43 (1998) 497–504.
[17] O. Baba, N. Takahashi, T. Terashima, W. Li, P.K. DenBesten, Y. Takano, Expression of alternatively spliced RNA transcripts of

amelogenin gene exons 8 and 9 and its end products in the rat incisor, J. Histochem. Cytochem. 50 (2002) 1229–1236.
[18] P.S. Hart, T.C. Hart, J.P. Simmer, J.T. Wright, A nomenclature for X-linked amelogenesis imperfecta, Arch. Oral Biol. 47 (2002) 255–260.
[19] J.T. Wright, P.S. Hart, M.J. Aldred, K. Seow, P.J. Crawford, S.P. Hong, C.W. Gibson, T.C. Hart, Relationship of phenotype and genotype in X-linked amelogenesis imperfecta, Connect. Tissue Res. 44 (2003) 72–78.
[20] T. Diekwisch, S. David, P. Bringas Jr., V. Santos, H.C. Slavkin, Antisense inhibition of AMEL translation demonstrates supramolecular controls for enamel HAP crystal growth during embryonic mouse molar development, Development 117 (1993) 471–482.
[21] S.P. Lyngstadaas, S. Risnes, B.S. Sproat, P.S. Thrane, H.P. Prydz, A synthetic, chemically modified ribozyme eliminates amelogenin, the major translation product in developing mouse enamel in vivo, EMBO J. 14 (1995) 5224–5229.
[22] C.W. Gibson, Z.A. Yuan, B. Hal, G. Longenecker, E. Chen, T. Thyagarajan, T. Sreenath, J.T. Wright, S. Decker, R. Piddington, G. Harrison, A.B. Kulkarni, Amelogenin-deficient mice display an amelogenesis imperfecta phenotype, J. Biol. Chem. 276 (2001) 31871–31875.
[23] P. Adeleke-Stainback, E. Chen, P. Collier, Z.A. Yuan, R. Piddington, S. Decker, J. Rosenbloom, C.W. Gibson, Analysis of the regulatory region of the bovine X-chromosomal amelogenin gene, Connect. Tissue Res. 32 (1995) 115–118.
[24] D.R. Nebgen, H. Inoue, B. Sabsay, K. Wei, C.S. Ho, A. Veis, Identification of the chondrogenic-inducing activity from bovine dentin (bCIA) as a low-molecular-mass amelogenin polypeptide, J. Dent. Res. 78 (1999) 1484–1494.
[25] P. Papagerakis, M. MacDougall, D. Hotton, I. Bailleul-Forestier, M. Oboeuf, A. Berdal, Expression of amelogenin in odontoblasts, Bone 32 (2003) 228–240.
[26] S. Oida, T. Nagano, Y. Yamakoshi, H. Ando, M. Yamada, M. Fukae, Amelogenin gene expression in porcine odontoblasts, J. Dent. Res. 81 (2002) 103–108.
[27] C.D. Fong, L. Hammarstrom, Expression of amelin and amelogenin in epithelial root sheath remnants of fully formed rat molars, Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod. 90 (2000) 218–223.
[28] D.D. Bosshardt, A. Nanci, Hertwig's epithelial root sheath, enamel matrix proteins, and initiation of cementogenesis in porcine teeth, J. Clin. Periodontol. 31 (2004) 184–192.
[29] D. Deutsch, A. Haze-Filderman, A.L. Taylor, A. Blumenfeld, L. Dafni, B. Shay, Z. Mao, Y. Gruenbaum-Cohen, E. Rosenfeld, Y. Leiser, E. Fermon, E. Shai, P. Jawor, B. Zimmermann, S. Hagewald, J.P. Bernimoulin, Amelogenin in cementum, periodontal ligament and bone cells in normal and regenerating tissues, in: Abstract, Eighth International Conference on Tooth Morphogenesis and Differentiation, 2004, p. 32.
[30] L. Hammarstrom, L. Heijl, S. Gestrelius, Periodontal regeneration in a buccal dehiscence model in monkeys after application of enamel matrix proteins, J. Clin. Periodontol. 24 (1997) 669–677.
[31] L. Heijl, Periodontal regeneration with enamel matrix derivative in one human experimental defect. A case report, J. Clin. Periodontol. 24 (1997) 693–696.
[32] M.L. Paine, S.N. White, W. Luo, H. Fong, M. Sarikaya, M.L. Snead, Regulated gene expression dictates enamel structure and tooth function, Matrix Biol. 20 (2001) 273–292.
[33] A. Veis, K. Tompkins, K. Alvares, K. Wei, L. Wang, X.S. Wang, A.G. Brownell, S.M. Jengh, K.E. Healy, Specific amelogenin gene splice products have signaling effects on cells in culture and in implants in vivo, J. Biol. Chem. 275 (2000) 41263–41272.
[34] A. Veis, Amelogenin gene splice products: potential signaling molecules, Cell Mol. Life Sci. 60 (2003) 38–55.
[35] J.P. Simmer, E.C. Lau, C.C. Hu, T. Aoba, M. Lacy, D. Nelson, M. Zeichner-David, M.L. Snead, H.C. Slavkin, A.G. Fincham, Isolation and characterization of a mouse amelogenin expressed in *Escherichia coli*, Calcif. Tissue Int. 54 (1994) 312–319.
[36] O.H. Ryu, A.G. Fincham, C.C. Hu, C. Zhang, Q. Qian, J.D. Bartlett, J.P. Simmer, Characterization of recombinant pig enamelysin activity and cleavage of recombinant pig and mouse amelogenins, J. Dent. Res. 78 (1999) 743–750.
[37] D. Deutsch, E. Chityat, M. Hekmati, A. Palmon, Y. Farkash, L. Dafni, High Expression of Human Amelogenin in *E. coli.*, Adv. Dent. Res. 10 (1996) 187–194.
[38] W. Li, C. Gao, Y. Yan, P. DenBesten, X-linked amelogenesis imperfecta may result from decreased formation of tyrosine fich amelogenin peptide (TRAP), Arch. Oral Biol. 48 (2003) 177–183.
[39] D. Deutsch, A. Haze, A. Taylor, L. Dafni, A. Blumenfeld, B. Shay, E. Rosenfeld, Y. Leiser, E. Fermon, Y. Cohen, S. Hagewald, P. Jawor, B. Zimmermann, J.P. Bernimoulin, Regeneration of the periodontal tissues in vivo, using the recombinant human amelogenin protein, J. Dent. Res. 82 (2003) 292 (Abstract, Special issue B).
[40] S. Hagewald, J.P. Bernimoulin, N. Pischon, B. Zimmermann, L. Dafni, A. Taylor, A. Haze, D. Deutsch, Periodontal healing following application of different enamel matrix protein fractions, J. Dent. Res. 82 (2003) 292 (Abstract, Special issue B).
[41] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular cloning–A laboratory manual, second ed., Cold Spring Harbor Laboratory Press, New York, 1989.
[42] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye inding, Anal. Biochem. 72 (1976) 248–254.
[43] J. Catalano-Sherman, R. Laskov, A. Palmon, S. David, D. Deutsch, Production of a monoclonal antibody against human amelogenin, Calcif. Tissue Int. 54 (1994) 76–80.
[44] D. Deutsch, A. Palmon, J. Catalano-Sherman, R. Laskov, Production of monoclonal antibodies against enamelin and against amelogenin proteins of developing enamel matrix, Adv. Dent. Res. (1987) 282–288.
[45] S. Habelitz, A. Kullar, S.J. Marshall, P.K. DeBesten, M. Balooch, G.W. Marshall, W. Li, Amelogenin-guided crystal growth on fluoroapatite glass-ceramics, J. Dent. Res. 84 (2004) 698–702.
[46] J. Tan, W. Leung, J. Moradian-Oldak, M. Zeichner-David, A.G. Fincham, The pH dependent amelogenin solubility and its biological significance, Connect. Tissue Res. 38 (1998) 215–221 , discussion 241–246.
[47] T. Aoba, E.C. Moreno, T. Tanabe, M. Fukae, Effects of fluoride on matrix proteins and their properties in rat secretory enamel, J. Dent. Res. 69 (1990) 1248–1255.
[48] S.J. Brookes, C. Robinson, J. Kirkham, W.A. Bonass, Biochemistry and molecular biology of amelogenin proteins of developing dental enamel, Arch. Oral. Biol. 40 (1995) 1–14.
[49] W.Y. Chen, A.W. BelL, J.P. Simmer, C.E. Smith, Mass spectrometry of native rat amelogenins: primary transcripts, secretory isoforms and C-terminal degradation, J. Dent. Res. 79 (2000) 840–849.
[50] J. Moradian-Oldak, N. Gharakhanian, I. Jimenez, Limited proteolysis of amelogenin: toward understanding the proteolytic processes in enamel extracellular matrix, Connect. Tissue Res. 43 (2002) 450–455.
[51] J. Moradian-Oldak, J.P. Simmer, E.C. Lau, T. Diekwisch, H.C. Slavkin, A.G. Fincham, A review of the aggregation properties of a recombinant amelogenin, Connect. Tissue Res. 32 (1995) 125–130.
[52] A.K. Mohanty, C.R. Simmons, M.C. Wiener, Inhitition of tobacco etch virus protease activity by detergents, Protein Expr. Purif. 27 (2003) 109–114.
[53] T. Pawelczyk, R. Kowara, F. Golebiowski, A. Matecki, Expression in *Escherichia coli* and simple purification of human Fhit protein, Protein Expr. Purif. 18 (2000) 320–326.

[54] H. Wunderli-Ye, M. Solioz, Purification and functional analysis of the copper ATPase CopA of *Enterococcus hirae*, Biochem. Biophys. Res. Commun. 280 (2001) 713–719.
[55] J. Moradian-Oldak, M.L. Paine, Y.P. Lei, A.G. Fincham, M.L. Snead, Self-assembly properties of recombinant engineered amelogenin proteins analyzed by dynamic light scattering and atomic force microscopy, J. Struct. Biol. 131 (2000) 27–37.
[56] P. Jawor, S. Hagewald, J.P. Bernimoulin, L. Dafni, A. Taylor, A. Haze, D. Deutsch, Stimulation of PDL cells by enamel matrix proteins in vitro, J. Dent. Res. 82 (2003) 289 , Abstract, Special issue B).
[57] M. Adamczyk, J.C. Gebler, J. Wu, Selective analysis of phosphopeptides within a protein mixture by chemical modification, reversible biotinylation and mass spectrometry, Rapid. Commun. Mass Spectrom. 15 (2001) 1481–1488.
[58] A. Haze-Filderman, A.L. Taylor, A. Blumenfeld, Y. Leiser, B. Shay, E. Rosenfeld, E. Fermon, Y. Gruenbaum-Cohen, L. Dafni, K. Heikinheimo, C.W. Gibson, B. Zimmermann, S. Haegewald, J.-P. Bernimoulin, D. Deutsch, Amelgogenin induces regeneration of mesenchymal tissues by stem cell recruitment (2005, in submission).