Published in final edited form as:

Eur J Oral Sci. 2008 April ; 116(2): 113–123. doi:10.1111/j.1600-0722.2007.00519.x.

Distal *cis*-regulatory elements are required for tissue-specific expression of enamelin (*Enam*)

Yuanyuan Hu¹, Petros Papagerakis¹, Ling Ye³, Jerry Q. Feng³, James P. Simmer², and Jan C-C. Hu¹

¹Department of Orthodontics and Pediatric Dentistry, University of Michigan School of Dentistry, Ann Arbor, MI, USA

²Department of Biologic and Materials Sciences, University of Michigan School of Dentistry, Ann Arbor, MI, USA

³Department of Biomedical Sciences, Baylor College of Dentistry, Dallas, TX, USA

Abstract

Enamel formation is orchestrated by the sequential expression of genes encoding enamel matrix proteins; however, the mechanisms sustaining the spatio-temporal order of gene transcription during amelogenesis are poorly understood. The aim of this study was to characterize the *cis*-regulatory sequences necessary for normal expression of enamelin (*Enam*). Several enamelin transcription regulatory regions, showing high sequence homology among species, were identified. DNA constructs containing 5.2 or 3.9 kb regions upstream of the enamelin translation initiation site were linked to a LacZ reporter and used to generate transgenic mice. Only the 5.2-*Enam*-LacZ construct was sufficient to recapitulate the endogenous pattern of enamelin tooth-specific expression. The 3.9-*Enam*-LacZ transgenic lines showed no expression in dental cells, but ectopic β -galactosidase activity was detected in osteoblasts. Potential transcription factor-binding sites were identified that may be important in controlling enamelin basal promoter activity and in conferring enamelin tissue-specific expression. Our study provides new insights into regulatory mechanisms governing enamelin expression.

Keywords

ameloblasts; enamel proteins; enamelin; osteoblasts; tissue-specific expression

Dental enamel is a mineralized tissue produced by highly specialized epithelial cells, the ameloblasts. Enamel is composed of a thick, rod-containing (prismatic) layer sandwiched between thin initial and final aprismatic enamel layers. Ameloblasts have a complex life cycle (1), consisting of three morphologically distinct stages known as the pre-ameloblasts (cell differentiation stage), secretory ameloblasts (secretion of enamel proteins, elongation of crystals, establishing the enamel layer), and maturation stage ameloblasts (degradation of enamel proteins and hardening of enamel). Ameloblasts secrete matrix proteins that are responsible for creating and maintaining an extracellular milieu favorable for mineral deposition (2). Three proline-rich proteins (amelogenin, enamelin, and ameloblastin) are the major secretory products of ameloblasts and have been isolated from the enamel matrix of

developing teeth (3). Severe enamel phenotypes have been demonstrated in mice with defective amelogenin (*Amel*) (4), enamelin (*Enam*) (5), and ameloblastin (*Ambn*) (6) genes. Defects in the human amelogenin gene on the X chromosome (*AMELX*) cause X-linked amelogenesis imperfecta, while *ENAM* mutations cause autosomal-dominant amelogenesis imperfecta (7, 8). Significantly, the phenotypes of patients with *AMELX* and *ENAM* mutations are confined to the enamel layer of the dentition, suggesting that these proteins are specialized for amelogenesis.

Although there has been significant progress in our understanding of the proteomics and genetics of dental enamel formation (9), the mechanism of dental enamel formation has yet to be discovered. Increasing evidence suggests that each of the enamel proteins has novel properties that may correlate with unique functional roles. Amelogenin, for instance, is relatively absent from the mineralization front where enamel crystal elongate (10), but forms spherical macromolecular assemblies designated ‘nanospheres’ (11) that may space enamel crystals and organize them into parallel arrays (12,13). Ameloblastin (also called amelin or sheathlin) is thought to maintain rod and inter-rod boundaries (14,15) and attach ameloblasts to the enamel surface (6,16). Enamelin localizes to the mineralization front where it may be involved in crystal elongation and regulating crystal habit (17). In this report we focus on the biology of enamelin.

Enamelin is the largest enamel protein, having an apparent molecular mass on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) of 186 kDa (18). Enamelin cleavage products were first isolated and characterized from the developing enamel layer of unerupted pig teeth (19,20). Porcine enamelin cleavage product (from the N-terminal half of the protein) having apparent molecular mass values of 32 kDa (17,21-23) and 89 kDa (24) have been extensively characterized. Immunohistochemistry of developing porcine and bovine enamel using antibodies raised against the 89 kDa enamelin showed a ‘reverse honeycomb’ pattern caused by the specific localization of this part of enamelin within the enamel rods and because of its relative absence in the ‘sheath space’ that partially surrounds each enamel rod (25). Because of the preferential intraprismatic localization of enamelin and its ability to bind to hydroxyapatite, it is believed that enamelin cleavage products might regulate crystal habit by inhibiting mineral deposition on the sides of the crystallites (17).

Enamelin expression initiates in pre-ameloblasts and continues throughout the secretory stage, as demonstrated by *in situ* hybridization (26). Besides its expression by ameloblasts, enamelin is marginally expressed by odontoblasts forming the dentino–enamel junction (DEJ) (27). Enamelin expression is not detected in any other tissues during development. Execution of this spatially and temporally restricted program is of critical importance for the completion of amelogenesis, and perturbation of this program results in abnormal enamel in mice and humans. However, despite the importance of enamelin expression for proper enamel formation, it remains unclear how enamelin expression is initiated and maintained during development. While much attention has been given to the transcriptional regulation of amelogenin (28-34), much less is known about the regulation of enamelin. In fact, not a single study has investigated the transcriptional control of enamelin expression. Characterizing the genetic regulation of enamelin expression will improve our knowledge on the molecular control of enamelin, enhance our understanding of amelogenesis, and provide insights into the pathophysiology of amelogenesis imperfecta.

Material and methods

DNA constructs

Two mouse enamelin DNA fragments were generated by amplifying regions of a previously isolated P1 clone (35). The designs of the amplification fragments were based upon preliminary

sequence analysis of the human and mouse enamel genes, assuming that the major regulatory elements for enamel expression are located in one or both of these DNA regions. The DNA fragments used were: (i) a 3.9 kb fragment that consists of -3,209 bp of the 5' flanking region, and (ii) a 5.2 kb fragment consisting of -4,485 bp of the 5' flanking region, both followed by the 5' untranslated region consisting of 214 bp of non-coding exon 1, the entire 299 bp of intron 1, 61 bp of non-coding exon 2, the entire 143 bp of intron 2, and the first 24 bp of exon 3, and ending immediately upstream of the translation initiation codon.

The polymerase chain reaction (PCR) primer K-in3F 5'-GTCGACGGATCCAAAACTTCTGCTCCCAG-3' and EN1AR 5'-TTCCCGGGCACCAAACTTTCATAAGCC-3' were used to amplify the mEnam DNA in the P1 clone and generate the 3.9 kb construct. (The underlined sequences at the 3' ends of the primers hybridize to the enamel cDNA sequence, while the 5' ends introduce restriction sites for engineering purposes.) The amplification product was ligated into pPCRScrip for DNA sequencing. The correct clone was restricted with *SalI* and *XmaI* to produce a fragment of approximately 3,948 bp, which was inserted into pLacZ (36) that was modified by introducing a *XhoI* site in the 3' multiple cloning site. This clone, when double digested with *BamHI* and *XhoI*, produced 7,560- and 2,515-bp bands.

The 5.2 kb enamel reporter construct was made by Ozgene (Murdoch, WA, Australia). The *Enam*-specific PCR primers (p286_05) 5'-CAAACAGTCGACGTAACCTACTACCTTTGAGGG CGGTC-3' and (P286_06) 5'-CAAACAGTCGACGTTTTACCAAACTTTCATAAGCC-3' were used to amplify 129Sv/J genomic DNA to generate a product of 5,244 bp, which was cloned and characterized by DNA sequencing. Digestion of this clone with *SalI* released a 5,226 bp fragment, which was subcloned immediately 5' to the NLS- β -galactosidase sequence. Following confirmation by DNA sequencing, this Enam-reporter construct was double digested with *KpnI* and *SphI* and then the 8,838 bp product was used to produce transgenic mice.

Analysis of enamel-LacZ activity *in vitro*

To test if the two enamel-LacZ constructs (the 3.9 kb and the 5.2 kb fragments) contain the necessary regulatory elements to drive β -galactosidase expression, transient transfection experiments were performed using the ameloblast cell line LS8 (kindly provided by Dr M. Snead, Center for Craniofacial Molecular Biology, USC, Los Angeles, CA, USA). The LS8 cell line has been established by immortalizing primary cultures of enamel organ epithelium with SV40 large T antigen and has been shown to express amelogenin, ameloblastin, and enamel (37).

LS8 cells were plated, 24 h prior to transfection, in Dulbecco's modified Eagle's medium (Sigma, St Louis, MO, USA) supplemented with fetal bovine serum (10%), penicillin (100 units ml⁻¹), and streptomycin (100 μ g ml⁻¹). Cells were transfected at 70% confluence with the 3.9 kb and 5.2 kb enamel-LacZ reporter constructs, or with LacZ empty vector, using the Lipofectamine 2000 transfection kit (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's protocol. Forty-eight hours after transfection, the medium was removed, the cells were rinsed three times with phosphate-buffered saline (PBS), fixed with 0.5% glutaraldehyde, and assayed for β -galactosidase activity, as previously described (38). Staining was evaluated and photographed using a dissecting microscope (Nikon SMZ1000, Mager Scientific, Dexter, MI, USA) and a digital camera, Nikon DXM1200.

Generation of the 3.9 and 5.2 kb enamel-LacZ hemizygote transgenic mice

The 3.9 and 5.2 kb enamel-LacZ DNA constructs were released from the vector by restriction endonuclease digestion and purified using a Qiaquick gel extraction kit (Qiagen, Valencia, CA,

USA). Each purified construct was microinjected into (C57BL/6X SJL)F2 eggs to generate transgenic mice. In total, eight independent lines (from three productive founders harboring the 3.9 kb construct and from five productive founders harboring the 5.2 kb construct) were generated at the Nathan Shock Aging Center at the University of Texas Health Science Center at San Antonio (UTHSCSA) and at the Transgenic Animal Model Core at the University of Michigan, respectively. Founders were bred with C57BL/6 male and female mice obtained from the Jackson Laboratory (Bar Harbor, ME, USA). All mouse lines for subsequent studies were maintained as hemizygotes. Tail DNA was isolated for genotyping by using PCR and/or Southern blot analysis. The UTHSCSA and University of Michigan Committees on the Use and Care of Animals approved all experimental procedures involving mice. The *in situ* activity of the enamel-LacZ constructs was determined using a standard protocol (39) by staining tissue sections of newborn to postnatal (PN) day 21 mice for β -galactosidase activity. The results were compared with the endogenous enamel mRNA expression pattern obtained by *in situ* hybridization.

Genotyping by PCR

Genotype of the *Enam-LacZ* transgenic mice was determined by PCR analysis of genomic DNA from tail biopsy using primers F: 5'-AAGTTTTGGGATTTGGCTCA-3' and R: 5'-GTTGCACCACAGATGAAACG-3'. The amplification product was a 600 bp band containing part of the enamel promoter sequence, the nuclear localization sequence, and part of the β -galactosidase gene sequence. The positive amplification reaction control was the plasmid construct. β -globin was the control for DNA isolation. Amplification was performed at 94°C for 30 s, 58°C for 30 s, and 72°C for 1 min, followed by incubation for 10 min at 72°C. The reaction was amplified for 40 cycles and the PCR products were run on a 1.5% agarose gel and visualized after staining with ethidium bromide.

Genotyping by Southern blot

One-hundred micrograms of mouse-tail DNA was digested to completion with the restriction enzyme *Bgl*III (Stratagene, La Jolla, CA, USA), fractionated on 0.7% agarose gels, and transferred to positively charged nylon membranes (#1-209-299; Roche, Indianapolis, IN, USA). Membranes were pre-hybridized in $\times 6$ SSC (sodium chloride/sodium citrate buffer), $\times 5$ Denhardt's, 0.5% SDS, 100 $\mu\text{g ml}^{-1}$ of sheared and denatured salmon sperm DNA, and 50% formamide at 42°C for > 60 min. Random-primed DNA probes were labeled with [^{32}P]dCTP [αP] (Amersham, Piscataway, NJ, USA; redivue deoxycytidine 5'-[α - ^{32}P]triphosphate, triethylammonium salt, AA0005, 10 $\mu\text{Ci } \mu\text{l}^{-1}$, specific activity 3,000 Ci mmol $^{-1}$) using a Random Primed DNA labeling Kit (Roche 1-004-760) and purified by Quick Spin Columns (Radiolabeled DNA purification kit, Roche 1-273-922 or 1-273-949). Denatured probe was added to the buffer at 10 7 counts per minute (c.p.m.) and hybridized overnight. Then, low-stringency washes involving two 5 min washes at room temperature (25°C) with $\times 2$ SSC containing 0.1% SDS, followed by two 15 min high-stringency washes at 65°C in $\times 5$ SSC and 0.1% SDS were performed. The blot was subsequently exposed to X-ray film (BioMax MR; Kodak 894-1114, Rochester, NY, USA) with intensifying screens at -70°C overnight.

Tissue collection

Newborn mice and mice from PN days 1, 2, 3, 5, 6, 7, 9, 12, 14, 17, and 20 were killed and tissues immediately collected. Whole bodies and/or heads were processed for β -galactosidase staining or *in situ* hybridization, as previously reported (26,27,40-42).

For β -galactosidase staining the dissected tissues were fixed in 4% paraformaldehyde for 30 min to 1.5 h according to the age of mice and size of the tissue blocks, washed with PBS three times, for 15 min each wash, then decalcified in 10% EDTA for up to 7 d, according to the age of mice (43,44). The decalcified tissue was washed in PBS (three washes, 15 min each) and

embedded in optimum cutting temperature (OCT) compound (Sakura Finetek, Torrance, CA, USA). Tissue blocks were cryo-sectioned at 10 μ m thickness.

For *in situ* hybridization, the tissues were dissected on ice and then immediately fixed in ice-cold 4% paraformaldehyde solution for 24 h. Following demineralization in 0.2 M HCl for 1–14 d, the tissues were embedded in paraffin and then cut into 6- μ m sections and mounted on RNase-free glass slides.

Analysis of *Enam-LacZ* activity *in situ*

β -galactosidase staining was performed to determine the spatial and temporal activity of the *Enam-LacZ* reporter constructs, as previously described (42). Briefly, tissue sections were postfixed for 5 min in 4% paraformaldehyde, washed in PBS (three washes of 5 min each), and incubated in freshly prepared X-gal staining buffer [1 mg ml⁻¹ of X-gal, 100 mM Hepes, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 1 mM MgCl₂, 2% Triton X-100, and 5 mM dithiothreitol (DTT), pH 8.0] at 45°C overnight. Tissue sections were rinsed and stored in PBS for counterstaining with hematoxylin and then observed under a dissection microscope (Nikon SMZ1000) or a light microscope (Nikon Eclipse E600). All images were captured using a digital camera (Nikon DXM1200) and ACT1 software (Mager Scientific, Dexter, MI, USA) imaging software.

In situ hybridization of the endogenous enamel mRNA

In situ hybridization was performed, as previously described (26). Briefly, the hybridization mixture, containing 1 $\times$ 10⁶ c.p.m. of ³⁵S-labeled mouse enamel riboprobe, 50% deionized formamide, 10% dextran sulfate, 1 $\times$ Denhardt's, 20 mM Tris-HCl, 0.3 M NaCl, 5 mM EDTA, 0.5 mg ml⁻¹ of tRNA, 20 mM DTT, and 10 mM NaH₂PO₄, was distributed evenly over the sections and incubated at 50°C for 16 h. Then, the slides were washed in $\times$ 4 SSC, 10 mM DTT for 30 min at 42°C; in $\times$ 2 SSC, 50% formamide, and 10 mM DTT for 20 min at 60°C; in riboprobe wash solution for 10 min at 37°C; in RNase solution for 30 min at 37°C; in riboprobe wash solution; in $\times$ 2 SSC; and in $\times$ 0.1 SSC. Finally, the slides were dehydrated, air dried, dipped in NTB-2 emulsion (Eastman Kodak, Rochester, NY, USA), and exposed at 4°C for 2 wk. Then, the photo emulsion was developed, fixed, and the hybridization signal photographed with both dark-field and bright-field optics on an Axioskop photo-microscope (Zeiss, Thornwood, NY, USA).

Comparative interspecies sequence analysis of the 5' mouse enamel regulatory regions

For comparative genomic analysis of the enamel regulatory regions we used a combination of several web-based programs. Vista plot analysis was performed using the rvista 2.0 program at <http://genome.lbl.gov/vista/index.shtml>. Alignment of multiple sequences was performed using the CLUSTALW, ALIGN, PROMOTERWISE, and T-COFFE programs at <http://www.ebi.ac.uk/Tools/sequence.html>. BLAST searches were also conducted by means of software available from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>) and using the <http://genome.ucsc.edu/site> resources.

In silico analysis of the 5' mouse enamel regulatory region

For *in silico* analysis we used human and mouse sequences (GenBank accession numbers AY167999 and AF303737) beginning 5 kb upstream of the putative transcription start site and extending to the translation initiation site within exon 3. To identify potential promoter regions and regulatory modules, the Genomatix (EIDorado) web-based program (Genomatix Software, Ann Arbor, MI, USA) was used at <http://www.genomatix.de/shop/index.html> (45). To identify potential transregulatory elements within the 5' regulatory region of the mouse enamel gene, putative transcription factor-binding sites were searched using the web-based programs

tfsearch (version 1.3) at <http://www.cbrc.jp/research/db/TFSEARCH.html> (46), tess (combined and filtered-search strings) at <http://www.cbil.upenn.edu/cgi-bin/tess/tess> (Schug and Overton, 1997), and GENE2PROMOTER, GEMSLAUNCHER, and MATINSPECTOR (Genomatix) at <http://www.genomatix.de/shop/index.html>. The predicted regulatory sequences were cross-checked between the different programs used and further analyzed using comparative genomics tools. A tooth gene expression database was also used (<http://bite-it.helsinki.fi/>) to verify tooth expression of the potential transcription factors that may bind to the mouse enamelin 5' regulatory region.

Results

DNA constructs

Based upon computer-assisted promoter analyses, we designed two *Enam*–LacZ reporter constructs to investigate the mouse *Enam* promoter region. In these constructs the bacterial LacZ sequence was modified to include a nuclear localization signal (NLS). With respect to the mouse enamelin genomic sequence (AF303737), the 5.2 kb–*Enam*–LacZ construct extended from nucleotide 269 to nucleotide 5,485 (5,217 bp) and the 3.9–*Enam*–LacZ construct extended from nucleotide 1,555 to nucleotide 5,485 (3,931 bp). The 3' ends of the 3.9 kb and the 5.2 kb *Enam* 5' regulatory sequences both terminated in exon 3, just five nucleotides upstream of the *Enam* translation initiation site. The beginning of exon 1 is at nucleotide 4,750, so the 5.2 kb–*Enam*–LacZ and 3.9 kb–*Enam*–LacZ reporter constructs contained 736 bp of transcribed *Enam* sequence (from exon 1 through part of exon 3) and 4.5 or 3.2 kb of 5' untranscribed *Enam* sequence upstream of the NLS–LacZ reporter code, as shown in Fig. 1A.

Detection of enamelin–LacZ activity *in vitro*

Histochemistry of an ameloblast-like cell line (LS8) hosting the 3.9 kb or the 5.2 kb constructs 48 h following transfection showed positive β -galactosidase activity, demonstrating that both fragments contain the necessary basal promoter elements (Fig. 1B). Staining was restricted to the nuclei, indicating specific expression of the nuclear-localized reporter molecule. Cells transfected with the promoterless LacZ vector (control) showed no β -galactosidase activity.

Generation of the 3.9 kb and 5.2 kb *Enam*–LacZ mice

Both *Enam*–reporter constructs were injected successfully into (C57BL/6/6XSJL)F2 blastocysts and the transgenic mice were identified by Southern blotting or PCR analysis. For the 3.9 kb–*Enam*–LacZ transgene, three of the five founders were productive, as shown by PCR analysis (Fig. 1C). For the 5.2 kb–*Enam*–LacZ transgene, five of nine founders were productive, based upon Southern blot hybridization (Fig. 1D). Offspring of the productive founders were evaluated for β -galactosidase expression.

Detection of 3.9 kb *Enam*–LacZ activity *in situ*

The activity of the 3.9 kb reporter construct was determined by β -galactosidase staining of tissue sections from newborn to PN day 20 mice. Whole mounts of newborn and 2-d-old wild-type mice were negative for β -galactosidase activity (Fig. 2A), as were the 5.2 kb transgenic mice, while the 3.9 kb transgenic mice at the same stages of development showed ectopic β -galactosidase expression (Fig. 2B). Staining in the craniofacial region localized to the mandible, nasal region, and papilla of the hair bulbs (Fig. 2B). Histological sections at day 1 showed β -galactosidase staining within the long bones of the extremities, primarily in cells encapsulating developing bones and joints (Fig. 2C). At day 11, osteoblasts of the long bones were still positive for β -galactosidase activity (Fig. 2D–E). Pre-osteoblasts in the alveolar bone surrounding the developing incisors and molars of 1-d-old and 5-d-old mice were positive for β -galactosidase, but no β -galactosidase-positive staining was observed within the incisors or

molars themselves (Fig. 2F,I,J). These results contrast with the ameloblast-specific wild-type *Enam* mRNA localization, as defined by *in situ* hybridization (Fig. 2G,H,K,L) (26). These findings suggest that the *Enam* 3.9 kb sequence was sufficient to promote reporter expression, but insufficient to direct ameloblast-specific expression. Characterization of the three independent 3.9 kb *Enam*–LacZ transgenic lines showed similar expression patterns that differed only in the intensity of the β -galactosidase staining (data not shown).

Detection of 5.2 kb-*Enam*–LacZ activity *in situ*

In contrast to the non-dental localization of reporter activity observed in the 3.9 kb-*Enam*–LacZ transgenic mice, the reporter activity in the 5.2 kb-*Enam*–LacZ transgenic mice correlated well with wild-type *Enam* expression. β -galactosidase activity was first detected in the secretory ameloblasts of day 1 incisors (Fig. 3A,B). The molars at this point were devoid of reporter expression (Fig. 3C). The 5.2 kb-*Enam*–LacZ reporter activity was first detected in the molar secretory ameloblasts of 5-d-old mice (Fig. 3D), which is somewhat later than the onset of normal enamel mRNA expression that occurs on day 2 (26). The 5.2 kb-*Enam*–LacZ transgene expression peaked around PN day 9 in the mandibular first molars (Fig. 3E). β -galactosidase activity was also observed in maturation stage ameloblasts in 14-d-old molars (Fig. 3F). β -galactosidase staining was clearly present within the nuclei of molar ameloblasts (Fig. 3G) and persisted in a few ameloblasts until the molars erupted in 17-d-old mice (Fig. 3H). No β -galactosidase expression was detected in 20-d-old molars (data not shown). In addition, no staining was detected in any other tissues, such as bone, liver, and lung (Fig. 3I–K). Compared with the *in situ* hybridization data (Fig. 2G–H, K–L), the 5.2 kb construct directed tissue- and stage-specific expression of the reporter molecule that recapitulated *Enam* expression. The reporter activity varied depending upon which transgenic mouse line was characterized, but in the majority of lines the expression of the reporter molecule overlapped that of the endogenous enamel mRNA expression.

Comparative sequence analysis of the 5' mouse enamel regulatory regions

Previously, PCR amplification of the 5' ends of enamel cDNA from an enamel organ epitheliaderived cDNA library was used to identify the *Enam* transcription initiation (+1) site (35). The 5.2 kb of the 5' *Enam* regulatory region was analyzed using web-based comparative genomics and promoter analysis programs. Because the wild-type temporal and spatial pattern of *Enam* expression was found to depend upon upstream regulatory elements present on the 5.2 kb-*Enam*–LacZ construct, but not on the 3.9 kb-*Enam*–LacZ construct, for discussion purposes we refer to the 5' sequence unique to the longer construct as the 'distal regulatory region', and the 3.9 kb segment common to both constructs as the 'proximal regulatory region'. Conserved non-coding sequences (CNCs) were identified in the distal and proximal regulatory regions (Fig. 4A). Conserved non-coding sequences have a significantly increased probability of playing a role in gene regulation or other functions *in vivo* (47,48).

A putative basal promoter (PBP) region was also predicted to exist between nucleotides –500 and +101 relative to the assigned transcription initiation site (+1), using web-based programs (Fig. 4A). The predicted promoter is located within the most proximal CNC, a region with 80% sequence identity between human and mouse (Fig. 4A). Within the predicted promoter region a putative TATA box (nucleotides –43 to –33) and an inverted CCAAT box (nucleotides –70 to –63) were identified immediately upstream of the transcription initiation site (Fig. 4A). Comparative sequence analyses showed that the PBP contains four conserved regions that show high homology among species (Fig. 4B). These promoter sequences are conserved among rat, rabbit, human, chimp, rhesus, dog, cow, armadillo, elephant, tenrec (lesser hedgehog), and opossum (Fig. 4B). Neither fish species nor chicken showed homology to the mouse enamel predicted promoter.

Four regulatory modules (M1–M4) were identified within the 5.2 kb mouse *Enam* regulatory region, all within conserved sequences among species. One regulatory module (M1) was located in the distal regulatory region and three regulatory modules (M2–M4) were located within the proximal regulatory region (Fig. 4A). M1 contains several potential overlapping binding sites for the vitamin D, estrogen, and androgen receptors, and the Smad4, Stat3, and Pax4 transcription factors. Most of these transcription factors are known to be expressed by ameloblasts and to regulate synergistically with other genes encoding matrix proteins. M2 and M3 are located within the PBP and contain potential binding sites for EGR/NFAT and SREB-1c/NF- κ B transcriptional complexes, respectively. M4 is located downstream of the transcription initiation site and contains a potential binding site for the CEBP/STAT complex. In addition to the above described regulatory modules, two TA-rich regions with multiple binding sites for homeobox-containing proteins were found, one being in the distal regulatory region and one in the proximal regulatory region (Fig. 4A).

The 5.2 kb enamel regulatory region was further analyzed for potential DNA-binding sites for transcription factors that are expressed in the β -galactosidase-positive tissues of the transgenic mice prior to, or concomitant with, enamel expression. Our combined analysis identified five conserved binding sites for the transcription factor Cbfa-1/Runx2, a key gene for skeletogenesis and tooth development. In addition, five potential binding sites for the Dlx and Msx homeodomain-containing transcription factors that regulate enamel and tooth formation were also found. The majority of these putative binding sites are within regions conserved among species. Of interest, four of the five identified Cbfa-1/Runx2 sites and all of the Dlx sites localized within the distal regulatory region that may confer enamel tissue-specific expression. The distal regulatory region also contains a regulatory module (M1) and a TA-rich (homeobox) region (Fig. 4A).

Discussion

While the molecular mechanisms governing tooth morphogenesis have been elucidated in detail (49), the genetic pathways controlling the specification of dental mineralized tissues remain unclear. Teeth are composed of three mineralized tissues (enamel, dentin, and cementum) and are surrounded by a fourth (bone). Dentin and cementum contain a repertoire of extracellular matrix proteins similar to that found in bone. In contrast, enamel proteins are not consistently detected in bone (and no bone phenotype has been detected in their absence), suggesting that regulatory mechanisms might exist that prevent all but marginal expression in other mineralized tissues. Among the enamel proteins, enamel is the most specific, being expressed only by ameloblasts. Our study reveals a *cis*-regulatory region that confers tooth-specific expression to the mouse enamel gene.

To investigate the spatial–temporal regulation of enamel *in vivo*, three independent transgenic lines were generated harboring 3.9 kb upstream of the translation initiation codon of the enamel gene linked to the reporter LacZ gene (3.9 kb). In these mice, enamel- β -galactosidase expression was suppressed in teeth, contrasting with the strong endogenous expression of enamel mRNA. Instead, enamel- β -galactosidase expression was detected within other mineralized tissues. More, precisely, β -galactosidase staining was found in pre-osteoblasts and chondrocytes, but not in ameloblasts or in odontoblasts. Overall, the pattern exhibited by β -galactosidase activity was different from endogenous enamel expression, suggesting that the 3.9 kb segment of the mouse enamel lacks important regulatory elements essential for the developmental and tooth-specific expression pattern of this gene.

In contrast to the results obtained with the 3.9 kb fragment, LacZ reporter mice harboring 5.2 kb of the enamel gene upstream of the translation initiation codon showed β -galactosidase expression exclusively in tooth in all five transgenic lines generated. However, within two of

the five lines examined, the intensity of expression was weaker when compared with the endogenous expression of enamel. The chromosomal position effects of the transgene integration site could be responsible for the differential intensity of β -galactosidase among the five lines, as previously suggested (50). Additionally, the copy number of the inserted fragment and the pattern of methylation of the CpG-rich regulatory sequences may cause variable expression levels between transgenic lines as a result of differences in chromatin structure modulation, as also observed for other transgenes (51). Nevertheless, in the present study, expression of the reporter gene overlapped that of endogenous enamel and most importantly mimicked its tissue-specific pattern. Preservation of the essential aspects of enamel tissue-specific expression from early to late amelogenesis indicates that critical elements for enamel gene expression are contained within the distal *cis*-regulatory region located between 5.2 and 3.9 kb upstream of the translation start site. Interestingly, both the 3.9 kb and the 5.2 kb

DNA constructs showed strong β -galactosidase activity *in vitro*, but also *in situ*, suggesting that both fragments contain *cis*-elements sufficient to drive LacZ expression. To identify candidate regulatory elements that control enamel promoter activity, we compared the enamel 5' regulatory sequences among several species using Vista plot analysis. Detection of genomic sequences that are highly conserved among species has been shown to be an effective approach for identifying important gene regulatory sequences (52). Our analysis for conserved sequences within this region revealed eight CNCs with high homology (over 60%) among species. Previous *in situ* experiments have shown that CNCs have a significant role in regulating dentin matrix protein-1 and dentin sialophosphoprotein (*Dspp*) gene expression (53,54). Therefore, the identified CNCs may contain important regulatory elements for enamel expression.

In an effort to identify candidate *trans*-acting factors that might regulate enamel expression, we also analyzed the 5' enamel regulatory region for potential transcription factor-binding sites using a combination of comparative genomics and computer promoter analysis approaches. Based on our results, two families of transcription factors were identified that may have a potential role in regulating enamel expression: homeobox-containing proteins and Runx domain proteins. Homeodomain proteins belonging to the Msx and Dlx families have been associated with skeletal and/or tooth development. Msx2, Dlx2, and Dlx3 are the only homeodomain proteins known to be expressed in secretory ameloblasts (55-58). Targeted disruption of Msx2 leads to extensive defects of enamel formation (59). Snead's group reported that Msx2 binds to CCAAT enhancer-binding protein α attached to the amelogenin promoter and regulates amelogenin expression (32,33). Moreover, over-expression of Dlx3 has been shown to regulate osteocalcin expression (60) and mutations of *DLX3* have been associated with enamel abnormalities (61,62). Taken together, these data suggest that homeobox proteins are involved in ameloblast differentiation and control enamel matrix formation, potentially through the regulation of genes encoding enamel proteins. Our study identified several Msx/Dlx homeodomain protein-binding sites located in the enamel promoter within highly conserved regions among species. In addition, enamel expression (26) overlaps with Msx2, Dlx2, and Dlx3 expression in ameloblasts, further supporting the hypothesis that these homeobox proteins may participate in enamel gene regulation.

Enamel regulatory regions also contain several canonical binding sites for Cbfa-1/Runx2. Mice lacking Cbfa-1/Runx2 form abnormal enamel owing to a failure of ameloblast differentiation (63). Furthermore, Cbfa-1/Runx2 is involved in the regulation of *Ambn* (64) and *Dspp* (65) promoter activity, suggesting that Cbfa-1/Runx2 might play a role in controlling the expression of genes encoding matrix proteins during tooth development. Cbfa-1/Runx2 is abundantly expressed in osteoblasts during bone formation (66) where enamel remains transcriptionally repressed. Furthermore, Cbfa-1/Runx2 expression is up-regulated in mature ameloblasts (63) when enamel expression is down-regulated (26). Taken together, these data

suggest that Cbfa-1/Runx2 might account for repression of enamel transcription within osteoblasts and mature ameloblasts.

Of interest, four Cbfa-1/Runx2 and two Dlx potential binding sites are located within the distal promoter region between 5.2 kb and 3.9 kb. Both Cbfa-1/Runx2 and Dlx homeodomain proteins have been shown to play important roles in conferring the tissue specificity of other matrix proteins. Tissue-specific transcriptional control of the osteocalcin gene within the osteoblasts requires a specific *cis*-acting sequence that serves as a binding site for Cbfa-1/Runx2 (67). In addition, studies on bone sialoprotein (Bsp) and the $\alpha 1(I)$ collagen genes clearly show that Cbfa-1/Runx2 (68,69) and homeodomain proteins (69,70) are essential in sustaining tissue-specific gene expression. Bone sialoprotein expression in mineralized tissue requires a homeodomain protein-binding site, and the mutation of two nucleotides within this site abolishes the Bsp tissue-specific expression pattern (70). In our studies, deletion of the distal promoter region results in loss of tissue-specific expression and ectopic β -galactosidase expression in bone, suggesting that the three Cbfa-1/Runx2 and the two Dlx binding sites may function together or independently with other regulatory elements to control tissue-specific expression of enamel. Further investigation is necessary to conclude on the role of the Cbfa-1/Runx2- and Dlx-binding sites of the enamel promoter. Nevertheless, our data emphasize the importance of distal *cis*-regulatory elements in the tissue-specific expression of enamel.

In summary, our data support the view that enamel gene regulation at the level of transcription involves not only activation of transcription at the correct developmental stage but also the prevention of transcription in inappropriate cell types. *Cis*-elements within enamel gene regulatory regions may execute these events by recruiting transcriptional activators or repressors in appropriate cell types. The results of this study provide evidence that the distal 5' enamel region contains *cis*-regulatory tissue-specific elements that are necessary for both positive (tooth) as well as negative (bone, cartilage) regulation of transcription. Understanding the mechanisms that regulate the transcription of terminal differentiation genes during amelogenesis will not only shed light in the basic biology of enamel formation but may also be clinically significant. Inherited defects in enamel formation caused by mutations in several enamel matrix proteins such as amelogenin, amelysin, ameloblastin, and enamel have been reported in humans and mice (4,6,9,71). In contrast, mutations in regulatory proteins have so far not been investigated as causes of amelogenesis imperfecta, with the exception of *DLX3*, which is involved in the tricho-dento-osseous syndrome (61). However, because enamel defects are often seen in a multitude of amelogenesis imperfecta families without evident anomaly of their enamel proteins, it is possible that molecular defects of regulatory molecules expressed by ameloblasts may also cause amelogenesis imperfecta. Therefore, identifying the regulatory proteins that control enamel matrix formation will provide the basis for understanding the pathogenesis of genetic diseases affecting tooth development.

Acknowledgements

We thank Malcolm Snead for providing the LS8 ameloblast-like cell line. This study was funded by NIDCR grant projects DE13221 and DE11301.

References

1. Smith CE. Cellular and chemical events during enamel maturation. *Crit Rev Oral Biol Med* 1998;9:128–161. [PubMed: 9603233]
2. Nanci, A. Enamel: Composition, formation, and structure.. In: Nanci, A., editor. *Ten cate's oral histology development, structure, and function..* Mosby; St Louis, MO: 2003. p. 145-191.
3. Fincham AG, Moradian-Oldak J, Simmer JP. The structural biology of the developing dental enamel matrix. *J Struct Biol* 1999;126:270–299. [PubMed: 10441532]

4. Gibson CW, Yuan ZA, Hall B, Longenecker G, Chen E, Thyagarajan T, Sreenath T, Wright JT, Decker S, Piddington R, Harrison G, Kulkarni AB. Amelogenin-deficient mice display an amelogenesis imperfecta phenotype. *J Biol Chem* 2001;276:31871–31875. [PubMed: 11406633]
5. Masuya H, Shimizu K, Sezutsu H, Sakuraba Y, Nagano J, Shimizu A, Fujimoto N, Kawai A, Miura I, Kaneda H, Kobayashi K, Ishijima J, Maeda T, Gondo Y, Noda T, Wakana S, Shiroishi T. Enamelin (*Enam*) is essential for amelogenesis: ENU-induced mouse mutants as models for different clinical subtypes of human amelogenesis imperfecta (AI). *Hum Mol Genet* 2005;14:575–583. [PubMed: 15649948]
6. Fukumoto S, Kiba T, Hall B, Iehara N, Nakamura T, Longenecker G, Krebsbach PH, Nanci A, Kulkarni AB, Yamada Y. Ameloblastin is a cell adhesion molecule required for maintaining the differentiation state of ameloblasts. *J Cell Biol* 2004;167:973–983. [PubMed: 15583034]
7. Hu JC, Yamakoshi Y. Enamelin and autosomal-dominant amelogenesis imperfecta. *Crit Rev Oral Biol Med* 2003;14:387–398. [PubMed: 14656895]
8. Kim JW, Simmer JP, Lin BP, Seymen F, Bartlett JD, Hu JC. Mutational analysis of candidate genes in 24 amelogenesis imperfecta families. *Eur J Oral Sci* 2006;114(Suppl 1):3–12. [PubMed: 16674655]
9. Hu JC, Yamakoshi Y, Yamakoshi F, Krebsbach PH, Simmer JP. Proteomics and genetics of dental enamel. *Cells Tissues Organs* 2005;181:219–231. [PubMed: 16612087]
10. Orsini G, Lavoie P, Smith C, Nanci A. Immunochemical characterization of a chicken egg yolk antibody to secretory forms of rat incisor amelogenin. *J Histochem Cytochem* 2001;49:285–292. [PubMed: 11181731]
11. Fincham AG, Moradian-Oldak J, Simmer JP, Sarte P, Lau EC, Diekwisch T, Slavkin HC. Self-assembly of a recombinant amelogenin protein generates supramolecular structures. *J Struct Biol* 1994;112:103–109. [PubMed: 8060728]
12. Beniash E, Simmer JP, Margolis HC. The effect of recombinant mouse amelogenins on the formation and organization of hydroxyapatite crystals in vitro. *J Struct Biol* 2005;149:182–190. [PubMed: 15681234]
13. Margolis HC, Beniash E, Fowler CE. Role of macromolecular assembly of enamel matrix proteins in enamel formation. *J Dent Res* 2006;85:775–793. [PubMed: 16931858]
14. Hu C-C, Fukae M, Uchida T, Qian Q, Zhang CH, Ryu OH, Tanabe T, Yamakoshi Y, Murakami C, Dohi N, Shimizu M, Simmer JP. Sheathlin: cloning, cDNA/polypeptide sequences, and immunolocalization of porcine enamel proteins concentrated in the sheath space. *J Dent Res* 1997;76:648–657. [PubMed: 9062558]
15. Lyngstadaas S. Synthetic hammerhead ribozymes as tools in gene expression. *Crit Rev Oral Biol Med* 2001;12:469–478. [PubMed: 11806517]
16. Cerny R, Slaby I, Hammarstrom L, Wurtz T. A novel gene expressed in rat ameloblasts codes for proteins with cell binding domains. *J Bone Miner Res* 1996;11:883–891. [PubMed: 8797107]
17. Tanabe T, Aoba T, Moreno EC, Fukae M, Shimizu M. Properties of phosphorylated 32 kd nonamelogenin proteins isolated from porcine secretory enamel. *Calcif Tissue Int* 1990;46:205–215. [PubMed: 2106381]
18. Hu C-C, Fukae M, Uchida T, Qian Q, Zhang CH, Ryu OH, Tanabe T, Yamakoshi Y, Murakami C, Dohi N, Shimizu M, Simmer JP. Cloning and characterization of porcine enamel mRNAs. *J Dent Res* 1997;76:1720–1729. [PubMed: 9372788]
19. Fukae M, Tanabe T. Separation of non-amelogenin component from purified amelogenin preparation of immature porcine enamel. *Jpn J Oral Biol* 1985;27:1249–1251.
20. Fukae M, Tanabe T. Nonamelogenin components of porcine enamel in the protein fraction free from the enamel crystals. *Calcif Tissue Int* 1987;40:286–293. [PubMed: 3034388]
21. Uchida T, Tanabe T, Fukae M, Shimizu M. Immunocytochemical and immunochemical detection of a 32 kda nonamelogenin and related proteins in porcine tooth germs. *Arch Histol Cytol* 1991;54:527–538. [PubMed: 1793666]
22. Yamakoshi Y. Carbohydrate moieties of porcine 32 kda enamel. *Calcif Tissue Int* 1995;56:323–330. [PubMed: 7767845]
23. Yamakoshi Y, Pinheiro FH, Tanabe T, Fukae M, Shimizu M. Sites of asparagine-linked oligosaccharides in porcine 32 kda enamel. *Connect Tissue Res* 1998;39:39–46. [PubMed: 11062987] Discussion 63–37

24. Fukae M, Tanabe T, Murakami C, Dohi N, Uchida T, Shimizu M. Primary structure of the porcine 89-kda enamelin. *Adv Dent Res* 1996;10:111–118. [PubMed: 9206327]
25. Fukae M, Tanabe T, Uchida T, Yamakoshi Y, Shimizu M. Enamelins in the newly formed bovine enamel. *Calcif Tissue Int* 1993;53:257–261. [PubMed: 8275354]
26. Hu JC, Sun X, Zhang C, Simmer JP. A comparison of enamelin and amelogenin expression in developing mouse molars. *Eur J Oral Sci* 2001;109:125–132. [PubMed: 11347656]
27. Hu CC, Hart TC, Dupont BR, Chen JJ, Sun X, Qian Q, Zhang CH, Jiang H, Mattern VL, Wright JT, Simmer JP. Cloning human enamelin cDNA, chromosomal localization, and analysis of expression during tooth development. *J Dent Res* 2000;79:912–919. [PubMed: 10831092]
28. Gibson CW, Collier PM, Yuan ZA, Chen E. DNA sequences of amelogenin genes provide clues to regulation of expression. *Eur J Oral Sci* 1998;106(Suppl 1):292–298. [PubMed: 9541239]
29. Gibson CW. Regulation of amelogenin gene expression. *Crit Rev Eukaryot Gene Expr* 1999;9:45–57. [PubMed: 10200911]
30. Papagerakis P, Hotton D, Lezot F, Brookes S, Bonass W, Robinson C, Forest N, Berdal A. Evidence for regulation of amelogenin gene expression by 1,25-dihydroxyvitamin d(3) in vivo. *J Cell Biochem* 1999;76:194–205. [PubMed: 10618637]
31. Papagerakis P, MacDougall M, Hotton D, Bailleul-Forestier I, Oboeuf M, Berdal A. Expression of amelogenin in odontoblasts. *Bone* 2003;32:228–240. [PubMed: 12667550]
32. Zhou YL, Lei Y, Snead ML. Functional antagonism between Msx2 and CCAAT/enhancer-binding protein alpha in regulating the mouse amelogenin gene expression is mediated by protein–protein interaction. *J Biol Chem* 2000;275:29066–29075. [PubMed: 10859305]
33. Zhou YL, Snead ML. Identification of CCAAT/enhancer-binding protein alpha as a transactivator of the mouse amelogenin gene. *J Biol Chem* 2000;275:12273–12280. [PubMed: 10766866]
34. Caton J, Bringas P Jr, Zeichner-David M. IGFs increase enamel formation by inducing expression of enamel mineralizing specific genes. *Arch Oral Biol* 2005;50:123–129. [PubMed: 15721138]
35. Hu JC, Zhang CH, Yang Y, Karrman-Mardh C, Forsman-Semb K, Simmer JP. Cloning and characterization of the mouse and human enamelin genes. *J Dent Res* 2001;80:898–902. [PubMed: 11379892]
36. Dymecki S. A modular set of FLp, FRT and LacZ fusion vectors for manipulating genes by site-specific recombination. *Gene* 1996;171:197–201. [PubMed: 8666272]
37. Chen LS, Couwenhoven RI, Hsu D, Luo W, Snead ML. Maintenance of amelogenin gene expression by transformed epithelial cells of mouse enamel organ. *Arch Oral Biol* 1992;37:771–778. [PubMed: 1444889]
38. Yang XC, Fan MW. Identification and isolation of human dental pulp stem cells. *Zhonghua Kou Qiang Yi Xue Za Zhi* 2005;40:244–247. [PubMed: 15938892]
39. Lawson KA, Dunn NR, Roelen BA, Zeinstra LM, Davis AM, Wright CV, Korving JP, Hogan BL. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev* 1999;13:424–436. [PubMed: 10049358]
40. Lu P, Gao XJ, Jia HT, Wang JD, Gao Y, Wei MJ. Enamelin transcriptional expression in developing postnatal rat tooth germ. *Zhonghua Kou Qiang Yi Xue Za Zhi* 2004;39:414–417. [PubMed: 15498352]
41. Papagerakis P, Peuchmaur M, Hotton D, Ferkdadjji L, Delmas P, Sasaki S, Tagaki T, Berdal A. Aberrant gene expression in epithelial cells of mixed odontogenic tumors. *J Dent Res* 1999;78:20–30. [PubMed: 10065942]
42. Lu Y, Zhang S, Xie Y, Pi Y, Feng JQ. Differential regulation of dentin matrix protein 1 expression during odontogenesis. *Cells Tissues Organs* 2005;181:241–247. [PubMed: 16612089]
43. Lobe CG, Koop KE, Kreppner W, Lomeli H, Gertsenstein M, Nagy A. Z/ap, a double reporter for cre-mediated recombination. *Dev Biol* 1999;208:281–292. [PubMed: 10191045]
44. Rosen AD. End-point determination in EDTA decalcification using ammonium oxalate. *Stain Technol* 1981;56:48–49. [PubMed: 6785901]
45. Quandt K, Frech K, Karas H, Wingender E, Werner T. Matind and matinspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* 1995;23:4878–4884. [PubMed: 8532532]

46. Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acids Res* 1998;26:362–367. [PubMed: 9399875]
47. Gottgens B, Barton LM, Gilbert JG, Bench AJ, Sanchez MJ, Bahn S, Mistry S, Grafham D, McMurray A, Vaudin M, Amaya E, Bentley DR, Green AR, Sinclair AM. Analysis of vertebrate SCL loci identifies conserved enhancers. *Nat Biotechnol* 2000;18:181–186. [PubMed: 10657125]
48. Loots GG, Locksley RM, Blankespoor CM, Wang ZE, Miller W, Rubin EM, Frazer KA. Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science* 2000;288:136–140. [PubMed: 10753117]
49. Thesleff I, Sharpe P. Signalling networks regulating dental development. *Mech Dev* 1997;67:111–123. [PubMed: 9392510]
50. Dobie K, Mehtali M, McClenaghan M, Lathe R. Varie-gated gene expression in mice. *Trends Genet* 1997;13:127–130. [PubMed: 9097721]
51. Bestor TH, Tycko B. Creation of genomic methylation patterns. *Nat Genet* 1996;12:363–367. [PubMed: 8630488]
52. Hardison RC, Oeltjen J, Miller W. Long human–mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Res* 1997;7:959–966. [PubMed: 9331366]
53. Yang W, Lu Y, Kalajzic I, Guo D, Harris MA, Gluhak-Heinrich J, Kotha S, Bonewald LF, Feng JQ, Rowe DW, Turner CH, Robling AG, Harris SE. Dentin matrix protein 1 gene cis-regulation: use in osteocytes to characterize local responses to mechanical loading in vitro and in vivo. *J Biol Chem* 2005;280:20680–20690. [PubMed: 15728181]
54. Godovikova V, Li XR, Saunders TL, Ritchie HH. A rat 8 kb dentin sialoprotein–phosphophoryn (DSP–PP) promoter directs spatial and temporal LacZ activity in mouse tissues. *Dev Biol* 2006;289:507–516. [PubMed: 16310176]
55. Dassule HR, Lewis P, Bei M, Maas R, McMahon AP. Sonic hedgehog regulates growth and morphogenesis of the tooth. *Development* 2000;127:4775–4785. [PubMed: 11044393]
56. Lezot F, Davideau JL, Thomas B, Sharpe P, Forest N, Berdal A. Epithelial Dlx-2 homeogene expression and cementogenesis. *J Histochem Cytochem* 2000;48:277–284. [PubMed: 10639494]
57. Zhao X, Zhang Z, Song Y, Zhang X, Zhang Y, Hu Y, Fromm SH, Chen Y. Transgenically ectopic expression of bmp4 to the msx1 mutant dental mesenchyme restores downstream gene expression but represses shh and bmp2 in the enamel knot of wild type tooth germ. *Mech Dev* 2000;99:29–38. [PubMed: 11091071]
58. Ghoul-Mazgar S, Hotton D, Lezot F, Blin-Wakkach C, Asselin A, Sautier JM, Berdal A. Expression pattern of Dlx3 during cell differentiation in mineralized tissues. *Bone* 2005;37:799–809. [PubMed: 16172034]
59. Bei M, Stowell S, Maas R. Msx2 controls ameloblast terminal differentiation. *Dev Dyn* 2004;231:758–765. [PubMed: 15499554]
60. Hassan MQ, Javed A, Morasso MI, Karlin J, Montecino M, van Wijnen AJ, Stein GS, Stein JL, Lian JB. Dlx3 transcriptional regulation of osteoblast differentiation: temporal recruitment of Msx2, Dlx3, and Dlx5 homeodomain proteins to chromatin of the osteocalcin gene. *Mol Cell Biol* 2004;24:9248–9261. [PubMed: 15456894]
61. Price JA, Bowden DW, Wright JT, Pettenati MJ, Hart TC. Identification of a mutation in Dlx3 associated with trichodonto-osseous (TDO) syndrome. *Hum Mol Genet* 1998;7:563–569. [PubMed: 9467018]
62. Dong J, Amor D, Aldred MJ, Gu T, Escamilla M, MacDougall M. Dlx3 mutation associated with autosomal dominant amelogenesis imperfecta with taurodontism. *Am J Med Genet A* 2005;133:138–141. [PubMed: 15666299]
63. D'Souza RN, Aberg T, Gaikwad J, Cavender A, Owen M, Karsenty G, Thesleff I. Cbfa1 is required for epithelial–mesenchymal interactions regulating tooth development in mice. *Development* 1999;126:2911–2920. [PubMed: 10357935]
64. Dhamija S, Krebsbach PH. Role of cbfa1 in ameloblastin gene transcription. *J Biol Chem* 2001;276:35159–35164. [PubMed: 11451945]

65. Chen S, Rani S, Wu Y, Unterbrink A, Gu TT, Gluhak-Heinrich J, Chuang HH, Macdougall M. Differential regulation of dentin sialophosphoprotein expression by runx2 during odontoblast cytodifferentiation. *J Biol Chem* 2005;280:29717–29727. [PubMed: 15980071]
66. Stein GS, Lian JB, van Wijnen AJ, Stein JL, Montecino M, Javed A, Zaidi SK, Young DW, Choi JY, Pockwinse SM. Runx2 control of organization, assembly and activity of the regulatory machinery for skeletal gene expression. *Oncogene* 2004;23:4315–4329. [PubMed: 15156188]
67. Gutierrez S, Javed A, Tennant DK, van Rees M, Montecino M, Stein GS, Stein JL, Lian JB. CCAAT/enhancer-binding proteins (C/EBP) beta and delta activate osteocalcin gene transcription and synergize with runx2 at the C/EBP element to regulate bone-specific expression. *J Biol Chem* 2002;277:1316–1323. [PubMed: 11668178]
68. Kern B, Shen J, Starbuck M, Karsenty G. Cbfa1 contributes to the osteoblast-specific expression of type I collagen genes. *J Biol Chem* 2001;276:7101–7107. [PubMed: 11106645]
69. Roca H, Phimpilai M, Gopalakrishnan R, Xiao G, Franceschi RT. Cooperative interactions between runx2 and homeodomain protein-binding sites are critical for the osteoblast-specific expression of the bone sialoprotein gene. *J Biol Chem* 2005;280:30845–30855. [PubMed: 16000302]
70. Benson MD, Bargeon JL, Xiao G, Thomas PE, Kim A, Cui Y, Franceschi RT. Identification of a homeodomain binding element in the bone sialoprotein gene promoter that is required for its osteoblast-selective expression. *J Biol Chem* 2000;275:13907–13917. [PubMed: 10788516]
71. Bartlett JD, Beniash E, Lee DH, Smith CE. Decreased mineral content in mmp-20 null mouse enamel is prominent during the maturation stage. *J Dent Res* 2004;83:909–913. [PubMed: 15557396]

Fig. 1.

Enamelin promoter/LacZ reporter constructs. (A) Map of transgene constructs. Numbering of the enamel sequences corresponds to the sequence shown in GenBank (accession no. AF303737). The smaller construct (3.9 kb-*Enam*-LacZ) contained 3,931 bp of Enam 5' untranslated region extending from bp 1,555 to bp 5,485. The larger construct (5.2 kb-*Enam*-LacZ) contained 5,217 bp of Enam from bp 269 to bp 5,485. Both Enam constructs contained the putative transcription initiation site at bp 4,750, non-coding exons 1 and 2, and the region of exon 3 up to, but not including, the translation initiation site, which was supplied by the bacterial β -galactosidase cDNA (β -gal) fused to a nuclear localization sequence (NLS) and followed by a polyadenylation sequence. (B) Transient transfection in an ameloblast-like cell line (LS8). Both the 3.9 kb and the 5.2 kb constructs were sufficient to drive the expression of β -galactosidase in the nucleus of the LS8 line. Cells transfected with a promoterless LacZ construct lacked β -galactosidase staining (left; neg. control). (C) Polymerase chain reaction (PCR) confirmation of successful integration of the 3.9 kb-*Enam*-LacZ transgene in three founders (a, b, c). (D) Southern blotting of transgenic mouse founders carrying the 5.2 kb construct.

Fig. 2.

Comparative detection of the 3.9 kb *Enam*-LacZ activity and the endogenous enamelin mRNA. (A, B) At postnatal (PN) day 2, staining of long bones of the extremities was distinct, mostly in the cells that encapsulate the developing long bones and joints. Staining in the craniofacial region was localized to the nasal primordial and papilla of the hair bulbs. (A) PN day 2 control mouse. (B) PN day 2 3.9 kb transgenic mouse. (C-E) The expression of β -galactosidase was first found in pre-osteoblasts of the long bone in PN day 1 mice (C). β -galactosidase activity was continuously detected, as demonstrated here by whole-mount staining of the forearm and digits from a PN day 11 sample (D). Positive staining of β -galactosidase activity was specifically located in the nuclei of the osteoblasts (E). (F, I, J) β -galactosidase activity was also detected in alveolar bone (AB) surrounding the incisors (F) and molars (I, J) of PN day 1

(F, I) and PN day 5 (J) transgenic mice. In contrast, β -galactosidase expression was not detected in either ameloblasts (AM) or the dental pulp (DP) of day 1 incisors (F) or day 1 and day 5 molars (I, J). (G, K) *In situ* hybridization was performed in PN day 1 and 5 mice. Enamelin mRNA was detected in the AM of both PN day 1 incisors (G) and PN day 5 molars (K). Enamelin mRNA was not detected in DP or in AB cells. (H, L) Bright field views corresponding to the *in situ* hybridization data shown in panels G and K.

Fig. 3.

Tissue-specific detection of the 5.2 kb *Enam*-LacZ activity. (A, B) β -galactosidase activity was detected in ameloblasts (AM) of postnatal (PN) day 1 incisor. Dental pulp (DP) and alveolar bone (AB) cells were devoid of staining (A). Positive staining of β -galactosidase activity was specifically located in the nuclei of the AM (B). (C-H) In molars, β -galactosidase reporter gene activity was not found in day 1 molars (C) and it was first detected in the AM of PN day 5 mice (D). The intensity of β -galactosidase activity peaked at PN day 9 (E) and it continued to be detected in maturation stage AM (F), being located within the nuclei (G) and until PN day 17 (H). DP (C-G) and AB (C, E) cells were devoid of staining. (I-K) No β -galactosidase activity was detected in long bones (I), lung (J) or liver (K).

Fig. 4.

Comparative sequence analysis of the 5,226 kb mouse enamel regulatory region. This region covers the DNA sequences including the 3.9 kb (proximal regulatory region) and the 5.2 kb fragments. The sequences located between 5.2 and 3.9 kb are designated as the distal regulatory region. (A) Conserved non-coding sequences (CNCs) showing over 60% sequence homology among mouse, rat, human, macaca, canine, bovine, and porcine enamel *cis*-regulatory regions were identified using Vista plots. Only the tops of the plots (showing 50–100% identity) are shown. Five detected CNCs are located within the enamel proximal regulatory region (1–5) and three are located within the distal regulatory region (a–c). A potential basal promoter (PBP) was detected within the most proximal CNC. Within the promoter a TATA box (T) and an inverted CCAAT box (C) were found immediately upstream of the transcription initiation site. In addition, one regulatory module (M1) was identified in the distal regulatory region and three regulatory modules (M2–M4) were located within the proximal regulatory region. Furthermore, two TA-rich regions with multiple binding sites for homeobox (H)-containing proteins were found. Finally, several potential binding sites for the transcription factors Cbfa-1/Runx2 (R), Dlx (D), and Msx (M) were detected by combining three web-based promoter analysis programs. Only putative binding sites with perfect homology to known consensus sequences are reported. (B) The PBP sequence was located within 500 bp prior to the transcription start codon (+ 1) and partially overlapped exon 1. Phylogenetic analysis showed that the predicted promoter contains four (1–4) highly conserved regions among species. In addition, three other conserved regions (5–7) were detected within the enamel 5'-untranslated (UTR) region.