Production and purification of recombinant amelogenins for investigating proteopathic mechanisms associated with amelogenesis imperfecta

Claire Madialin Gabe

Submitted in accordance with the requirements for the degree of
Doctorate of Philosophy

The University of Leeds
Faculty of Medicine and Health
School of Dentistry

November 2018

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. *The published work is integrated in the Materials and Methods, Results and Discussion sections of the thesis.*

The method Sections 2.1.1.1 (expression of recombinant amelogenin), 2.1.1.4 and 2.1.2.3 (polyacrylamide gel electrophoreses), 2.1.2.1 (nickel column chromatography), and the Figures 41, 42, 43 and 73 are based on work from the jointly authored publication:

GABE, C. M., BROOKES, S. J. & KIRKHAM, J. 2017. Preparative SDS PAGE as an Alternative to His-Tag Purification of Recombinant Amelogenin. *Frontiers in Physiology,* 8**,** 424.

All three authors Miss Claire Madialin Gabe (CMG), Dr. Seven Joseph Brookes (SJB) and Prof. Jennifer Kirkham (JK) contributed to the design of the experiments.
CMG carried out the practical work (data acquisition).
All three authors CMG, SJB and JK contributed to drafting the manuscript and approved the final submitted version.

Acknowledgements

It has been a great and enjoyable experience to work under the supervision of Dr. Steven Brookes and Prof. Jennifer Kirkham. I will always be thankful to them for their friendliness, invaluable help and constant support, patience and great scientific advice thoughout my PhD. I have enjoyed very much research as part of the Enamel research team.

I would like to thank as well Dr. Sarah Myers for her excellent help and support, Mr. Matty Percial for his great help on day-to-day basis, and other colleagues from Oral Biology department.
I would like to thank Dr. Thuy Do for her teaching of bioinformatics software.
I am very thankful to Dr. Xuebin Yang, PGRT, whose support has been very important in achieveing my PhD.
I am grateful to all my colleagues and friends in Oral Biology, whom it has been a pleasure to meet and work with on a daily basis.

I am grateful to my wonderful friends Lucia, Sam, Jorge, Oscar and Lizzie, Rodrigo and many others for their kindness and support, thanks to whom my Leodisian PhD years were very enjoyable.
I am grateful to Sasha, Adriana, Ricardo for introducing me to the exciting world of research and bringing to me an inspiring experience, thanks to which I knew I would want to enrol in PhD.

I would like to dedicate this thesis to Phyllis and Alan, for their constant love and help, to whom I owe a lot of my English language,
I would like to dedicate this thesis to my parents and whole family, to whom I am grateful for their love and constant support, in all circumstances, which have been important to reach this step.
Last but not least, I would like to dedicate my thesis as well to my beautiful Goddaughter Emilie, hoping that it can inspire her in the future, whichever will be her choices.

Abstract

In a mouse model of amelogenesis imperfecta (AI), an amelogenin p.Y64H mutation was reported to cause the abnormal retention of amelogenin in the ameloblast secretory pathway. This was hypothesised to be due to enhanced pathological aggregation of mutant amelogenin. The aim of this thesis was to develop purification methodologies to deliver large amounts of wild-type (WT) and mutant recombinant amelogenins (r-amelogenins) and develop microplate based binding assays to study protein-protein interactions of these recombinants to elucidate the effect of the mutation on aggregation. His-tagged r-amelogenin was extracted from *Escherichia coli* (*E. Coli*) with 3% acetic acid. This extract was subject to nickel affinity chromatography (targeting the His-tag); the 'gold standard' technique for purifying recombinant proteins from bacterial contaminants. The His-tag was then removed but cleavage was only ~50% efficient. Cleaved r-amelogenin unexpectedly still bound the nickel column (presumably due to the presence of di and tri histidine motifs in the amelogenin sequence) which precluded its isolation from uncleaved contaminant. Size exclusion chromatography was also trialled and also found to be ineffective. Finally, preparative SDS PAGE was found to produce cleaved r-amelogenin at single band purity on analytical SDS PAGE. After optimising the purification regime, simple and cost-effective microplate binding assays were developed initially using amelogenin rich porcine enamel matrix derivative (EMD) as a surrogate. Initially the aim was to immobilise EMD on the plate and then measure the binding of FITC-labelled EMD by simple end point fluorescence measurements. An alternative method trialled UV spectrophotometry to monitor the loss of EMD from free solution in real time as it bound EMD immobilised to the well surfaces. This latter method provided an adaptable, simple and cost-effective means of monitoring amelogenin binding and aggregation. It provided pilot data suggesting that p.Y64H mutant r-amelogenin was clearly more aggregative than WT r-amelogenin.

Table of Contents

List of Tables

List of Figures

Abbreviations

AFM	Atomic force microscopy
AI	Amelogenesis imperfecta
ATF6	Activating transcription factor 6
ATMP	Amelogenin trityrosyl motif peptide
BiP	Binding immunoglobulin protein
BSA	Bovine serum albumine
CD	Circular dichroism
DEJ	Dentine enamel junction
DLS	Dynamic light scattering
E. coli	*Escherichia coli*
ELISA	Enzyme-linked immunosorbent assay
EMD	Enamel matrix derivative
EMP	Enamel matrix proteins
ER	Endoplasmic reticulum
ERAD	ER-assisted degradation
ERAF	ER-assisted folding
FITC	Fluorescein isothiocyanate
FTIR	Fourier transform infrared
GlcNAc	N-Acetylglucosamine
hemagglut	Hemagglutination and hemagglutination inhibition test
His-tag	Polyhistidine tag
HR-SEC	High-resolution size exclusion chromatography
HSP	Heat shock protein
IDP	Intrinsically disordered protein
IEE	Inner enamel epithelium
IPTG	Isopropyl 1-thio-β-D galactopyranoside
IRE1α	Inositol-requiring enzyme 1α
ITC	Isothermal titration calorimetry
ITGB6	Integrin β6
KLK4	Kallikrein-related peptidase 4
KO	Knock-out
LRAP	Leucine-rich amelogenin peptide
MIH	Molar incisor hypomineralisation
MMP-20	Matrix metalloproteinase 20
NFDM	Non-fat dry milk proteins

Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
ODAM	Odontogenic ameloblast associated
PBS	Phosphate-buffered saline
PERK	Protein kinase RNA-like ER kinase
PPII	Polyproline II
QCM	Quartz crystal microbalance
r-amelogenin	Recombinant amelogenin
SAXS	small-angle X-ray scattering
SCPP	Secretory calcium-binding phosphoproteins
SDS	Sodium dodecyl sulphate
SDS PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SEM	Scanning electron microscopy
SLC24A4	Solute carrier family 24 member 4
SPR	Surface plasmon resonance
TBS	Tris buffered saline
TEM	Transmission electron microscopy
TEMED	N,N,N',N'-tetramethylethylenediamine
TRAP	Tyrosine-rich amelogenin peptide
TTBS	TBS containing 0.05 % Tween 20
UPR	Unfolded protein response
WDR72	WD repeat-containing protein 72
WT	Wild-type
XBP1	X-box binding protein 1
XRD	X-ray diffraction
Y2H	Yeast two-hybrid
PCA	Protein fragment complementation assay
FRET	Fluorescence resonance energy transfer

Chapter 1 Introduction

1.1 Dental enamel

Enamel is the hardest mineralised tissue in vertebrate biology and is evolutionarily conserved between species (Sire et al., 2007). It can form the outer layer of teeth in sarcopterygians (tetrapods and lobe finned bony fishes) and has been associated with the dermal bones and scales of fossilised lobe finned fish (Qu et al., 2015). It protects teeth from physical and chemical stresses encountered during feeding. Such a vital function would explain the high degree of conservation amongst the proteins involved in generating enamel (amelogenesis). Enamel's hardness and strength originates from its highly structured organisation of hydroxyapatite crystallites. The formation of enamel *per se* occurs in four discrete stages. In humans, amelogenesis begins *in utero* for deciduous teeth and at 0-1 years for permanent teeth (Nanci and Ten Cate, 2013). Developing enamel consists initially of a protein matrix secreted by the enamel forming ameloblasts which partially mineralises on secretion. Later, the protein matrix is degraded and the partially mineralised tissue undergoes secondary mineralisation. This process is described in detail later in the Introduction.

1.1.1 Structure and function of mature enamel

Teeth are an essential component of the digestive system; their primary function is food mastication, which breaks food into small particles accessible to digestive enzymes. Teeth are also important in speech, aesthetic appearance (in humans), self-defence or predation in some species. Mastication, which consists of tearing and crushing food, requires extremely resistant teeth, of an appropriate shape and hardness. To carry out millions of masticatory cycles over a life time (in humans), a robust structure is necessary as each cycle generates forces between 28 and > 1200 N (Ferrario et al., 2004, He et al., 2013).

Enamel is the external covering of teeth (Figure 1A). It comprises over 90% inorganic mineral by weight (De Menezes Oliveira et al., 2010), its mineral phase being essentially hydroxyapatite crystals (Deakins and Burt, 1944). Such a highly mineralised structure protects the dentine from physical and chemical stress which would otherwise cause demineralisation and pain due to the damaging effects of acid drinks and foods, and bacterial colonisation (Chun et al., 2014).

A study of human maxillary second molars reported enamel hardness values between 3 and 6 GPa and a Young elasticity modulus ranging from 70-115 GPa (Cuy et al., 2002). At the microscopic scale, enamel is a highly ordered structure where thousands of elongated hydroxyapatite crystallites are organised into bundles called prisms (or rods) interspersed with similar crystallites which comprise the interprismatic enamel (He and Swain, 2008) (Figure 1B). Prism organisation varies amongst species and tooth type and gives rise to histological banding features in enamel known as Hunter-Schreger bands. In rodent incisors, prisms are inclined relative to each other at about 60° generating an extreme decussating pattern. This arrangement confers mechanical strength, hardness and an ability to resist crack propagation and wear (Bajaj and Arola, 2009, Yahyazadehfar et al., 2013). This capacity to deflect crack stress propagation also protects the dentine enamel junction (DEJ) and the bulk of the tooth which is composed of dentine (Palmer et al., 2008). This creates a compound structure able to withstand mastication forces (He and Swain, 2008). In terms of evolution, the decussating structure of the enamel (as well as the thickness) have been linked to species selections and adaptation to hard-object diets (Dumont, 1995, Teaford and Ungar, 2000, Lambert et al., 2004).

Enamel *per se* is, however, brittle, with a fracture toughness measured at 0.6 -1.5 MPa (Park et al., 2008). It would be susceptible to cracking if it was not supported by the dentine, which is tougher and softer, and absorbs applied masticatory stresses (Chun et al., 2014). Studies carried out in primates suggested that this is further facilitated by the scalloped structure of the dentine-enamel junction, which resists delamination (Shimizu and Macho, 2007).

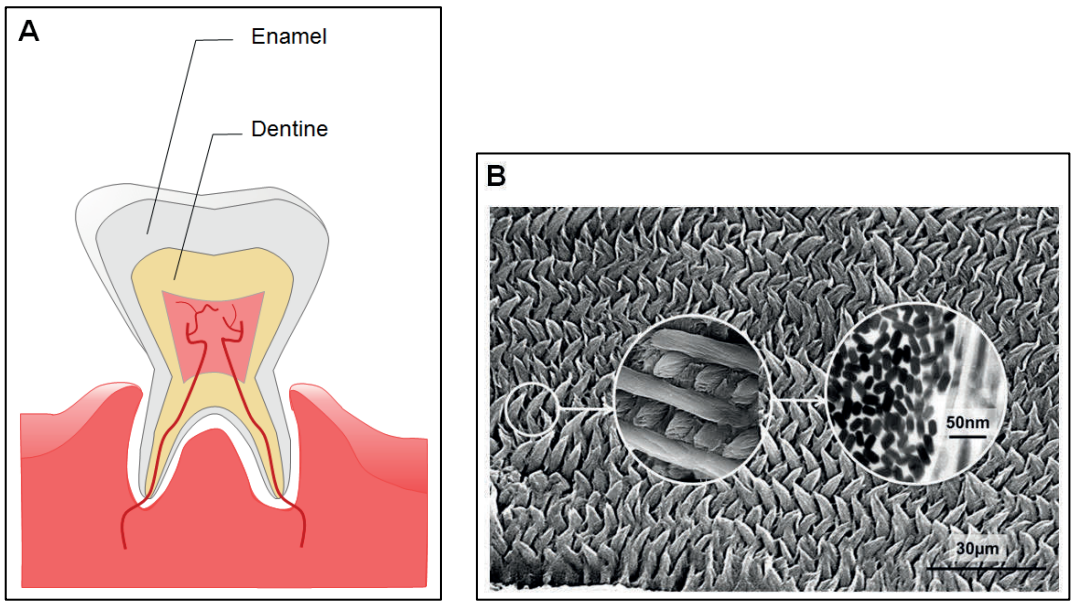


Figure 1 Schematic showing enamel on a simplified tooth. (A) Enamel is the outer mineral layer of teeth. Its colour is transfull/opaque. It protects the tissues of the inner layers. (B) Tissue architecture of rodent enamel relative to mineral crystals: enamel structure is highly ordered, organised into decussating prisms and interprismatic enamel, which makes it one of the toughest tissues in the organism and confers its mechanical properties.

1.1.2 Development of enamel: cellular and extra-cellular events

This Section provides a brief over view of amelogenesis as part of tooth formation and as applicable to human deciduous teeth, permanent teeth and also to tooth development in other species. A distinctive characteristic of tooth development in rats and mice is that their incisors continuously grow. In these cases, the enamel forms continuously but amelogenesis occurs via a similar route (further description is provided in Section 1.3.4.1.1 pp. 66 - 67).

1.1.2.1 Embryonic origins of enamel tissue

Human tooth formation is initiated after 37 days of development *in utero*, as the oral epithelium thickens at the internal surface of the upper and lower jaws. It forms primary epithelium bands that grow into the underlying mesenchyme comprising neural crest cells, forming the dental lamina. Tooth formation originates from this structure and advances through three stages: bud, cap, and bell stage (Figure 2). It is at the cap stage (8 weeks development *in utero*, deciduous teeth) that the enamel organ can be distinguished (Figure 2C). Its peripheral inner epithelial layer differentiates during the bell stage into ameloblasts, which are responsible for enamel formation.

During the bell stage, the inner enamel epithelium (IEE) directly faces the dental papilla and these two structures will form the enamel and dentine respectively. Amelogenesis can be described as a four-step process: (1) first, during a pre-secretory stage, the IEE cells differentiate into ameloblasts. (2) Second, during the secretory stage, the ameloblasts acquire secretory functions and secrete proteins forming the enamel matrix. The secretory stage is followed by (3) a transition stage, during which the secretion of enamel proteins stops and the ameloblasts undergo morphological changes reflecting a change of function. (4) Finally, the enamel matrix achieves full mineralisation throughout the maturation stage. All these steps are illustrated in Figure 3 and are discussed in detail below.

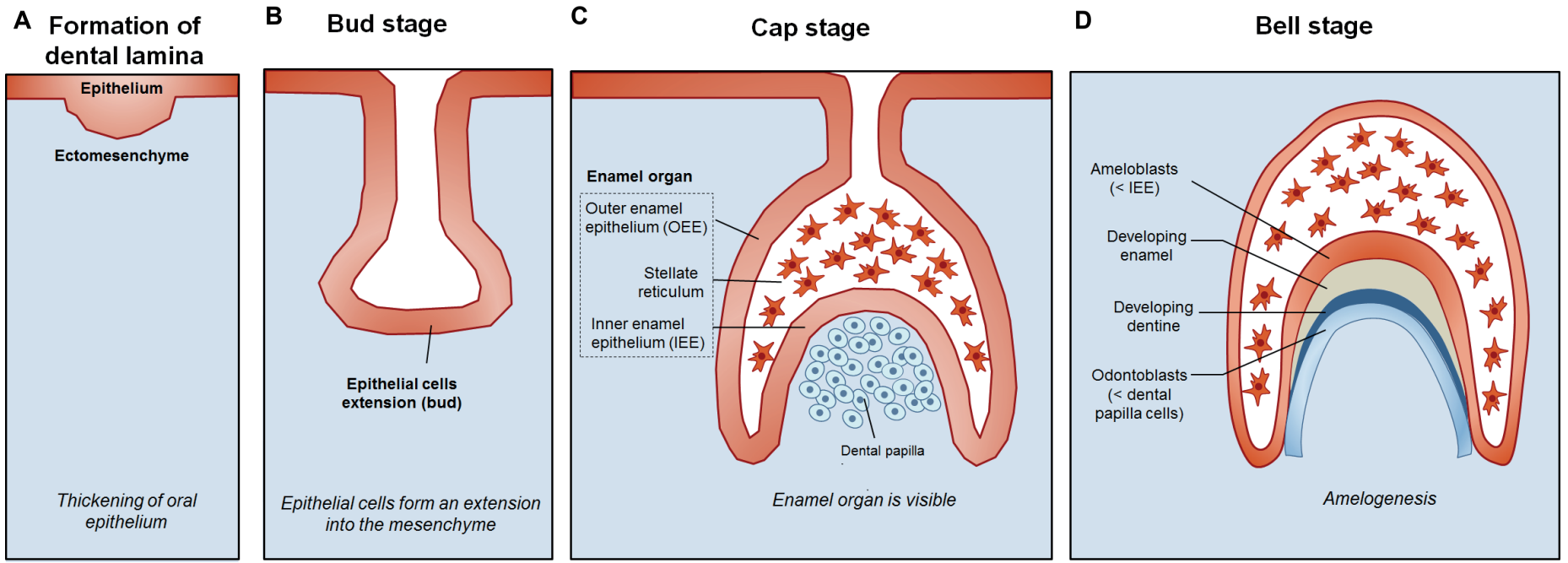


Figure 2 Major stages of deciduous tooth development (occurring in utero). (A) After 37 days of embryonic development, the oral epithelium thickens and forms the dental lamina. (B) Then, the epithelial cells multiply and migrate to form an extension into the underlying mesenchyme, which has the shape of a bud. (C) At 8-10 weeks development, the bud grows and shows the appearance of a 'cap' overlaying a 'ball' of condensed ectomesenchyme. This is when the enamel organ is visible: following series of histodiffferentiations, the centre of the enamel organ forms the stellate reticulum, while the peripheral layer forms the inner and outer enamel epithelium. (D) After 10 weeks, the tooth germ develops further, taking the shape of a bell (bell stage).

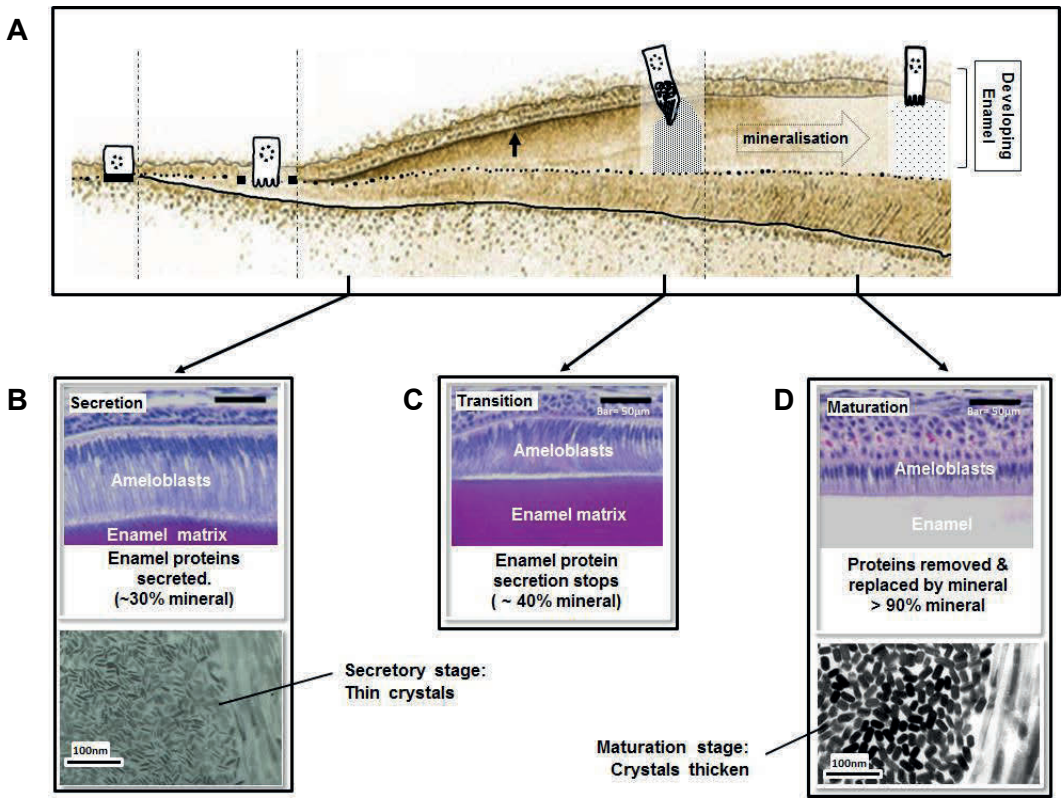


Figure 3 (A) Enamel develops incrementally with secretion of matrix proteins, associated with changes of cell morphology and function. Upon differentiation of epithelial cells, presecrectory ameloblasts are formed. They eventually start to incrementally secrete enamel matrix proteins (EMPs), into the forming enamel matrix. (B) During the secretory stage, EMPs are secreted via the amelobasts Tomes' processes. As matrix proteins are secreted, long, thin mineral crystals are deposited. The matrix at this stage is 33% mineral /33% protein and 33% water. (C)Tomes' processes resorb during the transition/maturation stage where the enamel has reached its full thickness. (D) At the maturation stage, the ameloblasts have changed in morphology (now half height, no Tomes' process), cells are more resorptive than secretory. Crucially, cells now pump large amounts of mineral ions into the matrix so pre-existing thin crystals grow in width and thickness to occlude >90% of the tissue. The panel A is adapted with permission from Hu et al. (2007), Cells Tissues Organs 2007;186:78–85. S.G.Karger AG, Basel[1] (itself adapted from Uchida et al. (1991))[2]. The histology analyses of mouse incisors in panels B, C, D are reproduced from Barron et al. (2010), Human Molecular Genetics 2010;19 (7):1230-47 by permission of Oxford University Press[3].

---

[1] HU, J. C., CHUN, Y. H., AL HAZZAZZI, T. & SIMMER, J. P. 2007. Enamel formation and amelogenesis imperfecta. *Cells Tissues Organs*, 186, 78-85.

[2] UCHIDA, T., TANABE, T., FUKAE, M. & SHIMIZU, M. 1991a. Immunocytochemical and immunochemical detection of a 32kDa nonamelogenin and related proteins in porcine tooth germs. *Archives of Histology and Cytology,* 54**,** 527-538.

[3] BARRON, M. J., BROOKES, S. J., KIRKHAM, J., SHORE, R. C., HUNT, C., MIRONOV, A., KINGSWELL, N. J., MAYCOCK, J., SHUTTLEWORTH, C. A. & DIXON, M. J. 2010. A mutation in the mouse Amelx tri-tyrosyl domain results in impaired secretion of amelogenin and phenocopies human X-linked amelogenesis imperfecta. *Human Molecular Genetics,* 19**,** 1230-47.

1.1.2.2 Presecretory stage: Inner enamel epithelium cells differentiate into ameloblasts

At the onset of amelogenesis, the cuboidal cells of the IEE are separated from the underlying mesemchymal dental papilla cells by a basement membrane. The cells engage in reciprocal intercellular signalling across the basement membrane which triggers the differentiation of IEE cells into presecretory ameloblasts and the dental papilla cells into odontoblasts which are responsible for dentine formation. Just before the first layer of dentine forms, differentiating IEE cells secrete proteins possibly involved in the epithelial-mesenchymal signalling (Thesleff, 2003), which drives the terminal differentiation of odontoblasts.

Initially low columnar cells, with a large central nuclei and underdeveloped secretory apparatus, and resting on the basement membrane (Figure 4A), the differentiating IEE cells undergo morphological and functional changes (Smith and Nanci, 2003). As odontoblasts begin to secrete a predentine matrix, IEE cells become taller and columnar. Their nuclei migrate proximally (towards the side bordering the stratum intermedium) while their Golgi and the rest of their secretory apparatus migrate towards the distal end: the IEE become ameloblasts, organised as a monolayer, separated from the odontoblasts by a basement membrane (Figure 4B).The ameloblasts become secretory, sending projections through the basement membrane, which is degraded (Figure 4C) (Reith, 1967).

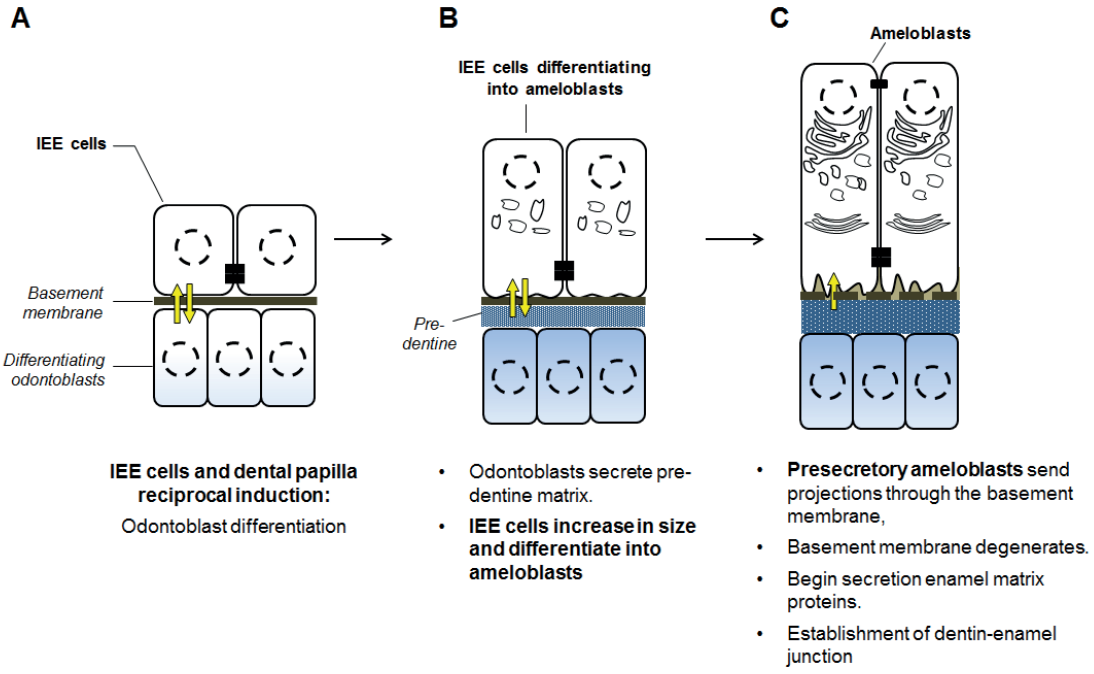


Figure 4 IEE cells differentiate into ameloblasts following reciprocal induction (highlighted by two yellow arrows) with odontoblasts (and their mesenchymal precursors).

1.1.2.3 Secretory stage: ameloblasts acquire secretory function in contact with pre-existing dentine, then secrete enamel matrix

The newly differentiated ameloblasts initially secrete low levels of proteins. However, following initiation of dentinogenesis the ameloblasts become highly active secretory cells and develop the so-called Tomes' processes which appears to direct the secretion of enamel proteins towards formation of discrete prismatic and interprismatic enamel (Nanci and Warshawsky, 1984, Warshawsky et al., 1987)

The secretory ameloblasts are tall columnar cells attached to each other by junctional complexes, forming a monolayer. They show typical characteristics of specialised secretory cells including an expansive Golgi and rough endoplasmic reticulum (ER) and membrane- bound secretory granules packed with EMPs (Reith, 1961, Reith, 1970, Sasaki et al., 1984). Responsible for the synthesis of EMPs, ameloblasts are adapted to manage the large amount of proteins transiting through the rough ER and the Golgi apparatus (Warshawsky, 1968). The secretory vesicles migrate to the Tomes' process, where they are released and added to the extracellular enamel matrix already secreted (Kallenbach, 1973). By incrementally secreting EMPs, the ameloblasts migrate away from the dentine surface (Figure 5A). This incremental apposition of matrix proteins leads to the deposition of the full thickness of enamel.

The enamel matrix is highly proteinaceous during the secretory stage: it comprises mainly amelogenin, ameloblastin, and enamelin; amelogenin is the major protein, 90% (Termine et al., 1980), and catalytic amounts of matrix metalloproteinase 20 (MMP-20) are present, which is responsible for extracellular processing of the EMPs (detailed later in Section 1.2.3.1, pp. 30 -32). Enamelin is expressed and released into the matrix from the early secretory stage to the early maturation stage. A study in mouse molars showed that its expression finishes before that of amelogenin (Hu et al., 2001). After their secretion, the nascent amelogenin, ameloblastin and enamelin molecules are soon processed by MMP-20 so that the bulk of the tissue is comprised of a range of enamel protein-derived polypeptides (Smith et al., 1989, Fukae and Tanabe, 1998, Uchida et al., 1998) which may also have architectural function (Bartlett, 2013).

Almost simultaneously with protein secretion, thin plate-like enamel apatite crystals elongate in length (c-axis growth) measuring at least 100 µm (Nylen et al., 1963) and possibly spanning the full thickness of the enamel (Daculsi et al., 1984, Margolis et al., 2006). Their growth is restricted to the c-axis as the EMPs likely act as a constraint to crystal growth in thickness and width (Robinson et al., 1998).

In effect, the crystal architecture is lightly "sketched in" during the secretion stage (Figure 5B) but how these crystals are nucleated is not fully established. Initially, calcium and phosphate ions are present in the earliest enamel secreted (Deakins and

Burt, 1944, Robinson et al., 1974, Takano et al., 1986). Mineral is nucleated in the extracellular environment to form an ordered crystal array (Fincham et al., 1999). The crystals appear thicker at the dentine-enamel junction and thinner at the secretory front . The mechanisms for this have not been fully elucidated, the hypotheses drawn to date are detailed in Section 1.2.3.1.2 (pp. 33 – 35).

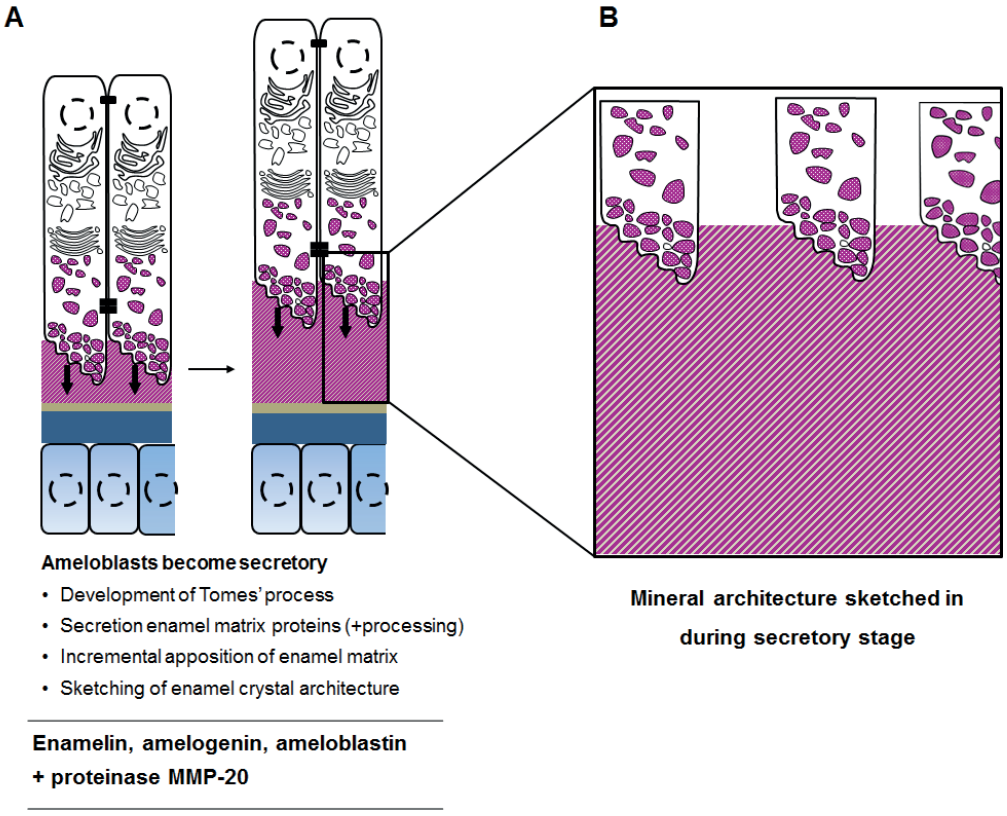


Figure 5 Secretory stage of amelogenesis (A) Secretory ameloblasts are long columnar cells, specialised in EMPs secretion. Enamel matrix forms incrementally. (B) enamel crystallites are 'sketched in' during the secretory stage.

1.1.2.4 Transition and maturation stages: reduction of secretory activity and secondary mineralisation

The transition stage starts just before enamel has reached its full thickness. The ameloblasts retract their Tomes' process and stop secreting secretory stage proteins such as amelogenin and ameloblastin, enamelin and MMP-20 (Bartlett, 2013, Smith et al., 2017b). They begin to decrease in height and start secreting a final layer of aprismatic enamel matrix before finally secreting a new basement membrane onto the surface of the enamel (Figure 6A). The basement membrane is present throughout the maturation stage, where maturation-stage-specific secretory calcium-binding phosphoproteins (SCPP) are found.

These SCPPs include amelotin and odontogenic ameloblast associated (ODAM) protein. Amelotin was identified in 2005 by differential display polymerase chain reaction analysis of mouse ameloblasts. Its encoding gene *AMTN* is conserved between species located close to the genes encoding other members of the SCPP family such as *ENAM* and *AMBN* on chromosomes 5 in mouse or 4 in humans (Iwasaki et al., 2005). Both mRNA expression pattern and protein localisation analyses showed that its expression is specific to maturation stage ameloblasts (Somogyi-Ganss et al., 2012). The predicted translation product has an elevated content of proline, leucine, glutamine and threonine (52%) which may drive hydrophobic interactions promoting amelotin aggregation (Iwasaki et al., 2005, Somogyi-Ganss et al., 2012). Amelotin has been found in the outer enamel layer during the transition towards maturation stage, and at the basement membrane-like structure of maturation stage ameloblasts (Moffatt et al., 2006b, Somogyi-Ganss et al., 2012), where it interacts with ODAM and secretory calcium-binding phosphoprotein proline-glutamine rich 1 (Holcroft and Ganss, 2011, Fouillen et al., 2017). Its aggregative nature suggests that amelotin may mediate the attachment of ameloblasts to the mineralising enamel during the maturation stage (Moffatt et al., 2014, Smith et al., 2017b). It was also shown to promote hydroxyapatite mineralisation (Abbarin et al., 2015).

ODAM, also known as APin, is another protein secreted by ameloblasts specifically during the maturation stage. Although it is overexpressed in some epithelial tumours (Kestler et al., 2008), its amelogenesis-specific expression is physiological and it was identified and characterised by secretome analysis of the rat enamel organ (Moffatt et al., 2006a, Moffatt et al., 2008). ODAM is highly conserved in mammals and has a rich content of proline and glutamine (28%) and hydrophobic residues (30%). During the maturation stage it co-localises with amelotin at the basement membrane (Park et al., 2007, Moffatt et al., 2008) and a yeast two-hybrid (Y2H) assay showed that they may interact (Holcroft and Ganss, 2011), possibly contributing to ameloblast adhesion to

the mineralising enamel matrix. Other proteins involved in cell adhesion, predominant during maturation stage include integrin β6 (ITGB6), at the distal membrane of ameloblasts (Wang et al., 2014a)

At the onset of the maturation stage, the enamel matrix has reached its full thickness. The ameloblasts decrease further in height until they are only half the height of secretory stage ameloblasts. Their apical membranes oscillate between ruffle and smooth states (Figure 6B) (Josephsen and Fejerskov, 1977, Reith and Boyde, 1981). They begin to secrete kallikrein-related peptidase 4 (KLK4), a serine protease which further degrades the spectrum of proteins and peptides generated during the secretory stage processing (Hu et al., 2002). Ruffle-ended ameloblasts secrete bicarbonate ions and are responsible for pumping calcium into the enamel matrix whereas smooth-ended ameloblasts allow for calcium intercellular transit towards the enamel matrix. The bicarbonate ions are assumed to buffer the decrease in pH in the matrix caused by the release of protons as the hydroxyapatite crystals grow in width and thickness (Smith, 1998).

The maturation stage is further characterised by an increase in mineral content and the concomitant loss of the protein degradation products generated by KLK4 proteolysis activity as they are endocytosed by ruffle-ended ameloblasts (Bartlett and Simmer, 2014, Pham et al., 2017). WD repeat-containing protein 72 (WDR72), whose predicted structure suggests a role in vesicle formation, may be involved in EMPs degradation products removal (Katsura et al., 2014). Fluid (enamel fluid) replaces the degraded enamel proteins. This fluid is kept supersaturated with respect to hydroxyapatite as the ameloblasts pump mineral ions into the enamel which drives the growth in thickness and width of the enamel crystals until the enamel becomes fully mineralised. Maturation stage ameloblasts also express solute carrier family 24 member 4 (SLC24A4), which is a sodium/potassium/calcium exchanger, playing an important role during maturation stage by transporting Ca^{2+} ions into the enamel matrix (Wang et al., 2014b).

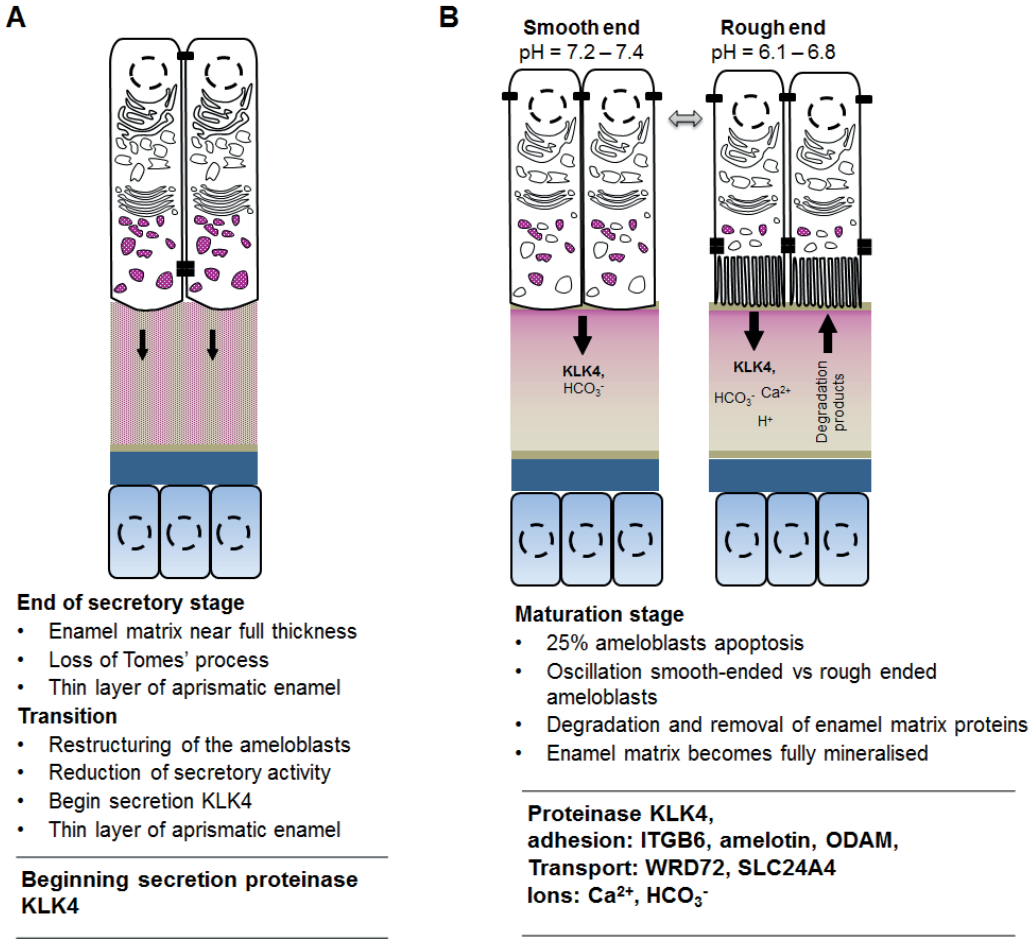


Figure 6 Transition and maturation stages. (A) ameloblasts reduce their secretory activity and start secreting KLK4 which is responsible for protein degradation. (B) during the maturation stage, ameloblasts maximally secrete KLK4 and oscillate between smooth-ended and rough-ended morphologies as they pump mineral ions into the matrix and resorb matrix protein degradation products. The enamel becomes fully mineralised.

The above description (Section 1.1.2 pp. 4 - 12) provides a brief overview of all stages of amelogenesis. However, the main focus of this thesis concerns the EMPs predominant at the secretory stage and the next Section describes these proteins in more detail.

1.2 Secretory stage enamel matrix proteins

The previous Section (1.1.2) provided a brief overview of amelogenesis. This Section details the state of knowledge about the synthesis, structure, assembly and function of the EMPs at the secretory stage, especially amelogenin, which is the main focus of this thesis.

During the secretory stage, enamel is ~30% protein by weight, comprising mainly proteins secreted by ameloblasts (Fukae and Shimizu, 1974). The major structural matrix proteins characterised to date are amelogenin, ameloblastin and enamelin. which can be found in the secretory enamel matrix as various isoforms (further information is described in the next Section). Proteinase MMP-20 is also found in the matrix during the secretory stage. Structural matrix proteins are secreted by ameloblasts on the preformed dentine (Hu et al., 2007) where they self-assemble to form a matrix that supports crystal formation.

1.2.1 Amelogenin, the major enamel matrix protein

Amelogenin is the major enamel protein, accounting for over 80-90% of total proteins in the secretory stage matrix (Eastoe, 1965, Termine et al., 1980).To date, neither the structure nor the precise function of amelogenin is fully understood. It is essential for amelogenesis as knock-out (KO) mice display a chalky white, disorganised hypoplastic enamel. Scanning electron microscopy (SEM), light microscopy and other analyses including sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS PAGE) and western blotting, carried out on these KO mice, suggested that amelogenin is necessary for the generation of enamel of the correct thickness and prismatic structure, but not for the formation of mineral crystals (Gibson et al., 2001). Various mutations in the amelogenin gene have been reported in cases of human AI; this is detailed in the Section 1.3.3 (Figure 12 p. 58). In the enamel matrix, amelogenin exists as a heterogeneous mixture. The factors explaining its diversity are detailed in a review by Brookes et al. (1995); they include gene sexual dimorphism, alternative splicing, and proteolytic processing in the extracellular matrix (Brookes et al., 1995).

1.2.1.1 Sexual dimorphism

Amelogenin is encoded by *AMELX*, located on chromosome X, and by *AMELY* on chromosome Y in human, pig and cow species (Note: *AMELX* in capital letters is the nomenclature for the human gene, specifically). In mouse and rat, the amelogenin

locus is found only on the X chromosome (Lau et al., 1989, Gibson et al., 1992, Veis, 2003, Ikawa et al., 2005). In humans, the amino acid sequence of proteins AMELX and AMELY share 88% similarity by Clustal Omega sequence alignment (by aligning the entries Q99218-1 and Q99217-1) (The UniProt Consortium, 2017). At the gene level, the 5' and 3' UTR regions appear to be much less conserved. At the gene level, the exon and intron structures do not differ significantly, however the 5' upstream regions share only 80% similarity, possibly causing a difference of transcriptional activity between *AMELX* and *AMELY* genes . This may explain why *AMELX* transcripts are 10-fold more abundant than *AMELY* transcripts in male tooth buds (Fincham et al., 1991a, Salido et al., 1992). In this thesis, amelogenin will refer to the proteins encoded by *AMELX* only.

1.2.1.2 Transcription variants of amelogenin (alternative splicing)

Alternative splicing is another factor explaining the diversity of amelogenins. Early studies include Northern hybridisation experiments, which discerned two species of mRNA derived from a single amelogenin gene in bovine ameloblasts (Shimokawa et al., 1987). Alternative amelogenin spliced products were also found in mouse (Lau et al., 1992), pig (Yamakoshi et al., 1994) and human species (Salido et al., 1992). In most species studied, the gene structure comprises seven exons. The primary transcripts obtained are subjected to alternative splicing, and several mRNA isoforms (shown in Figure 7) have been identified, with many being confirmed at the protein level. In most eutherian mammals, the major mRNA transcript for amelogenin comprises exons 1, 2, 3, 5, 6, 7 (Salido et al., 1992) which encodes for a protein of 173-197 amino acids (excluding a N-terminal signal peptide) depending on species. Note: Throughout this thesis, that major splicing isoform of amelogenin (see Figure 7b), minus its N-terminal signal peptide, are refered to as either the "full-length", "nascent" or "parent" amelogenin when it is unprocessed.

Exons 1, 2, 3, 5, 6, 7 correspond to the ancestral gene of amelogenin, they are present in mammals including monotremes (platypus) and non-mammalian species such as frogs and lizards (Sire et al., 2012). They are highly conserved, even though the length of the longest exon, exon 6, varies between species.

Exon 1 contains the 5' untranslated region of the mature mRNA, a region which expands into the 5' end of the exon 2. Downstream, exon 2 comprises the initiation codon followed by the sequence encoding the signal peptide and the first two amino acids of the secreted protein. Exons 3 and 5 encode short amino acid sequences (16-17 amino acids each). Exon 4 is spliced out in most variants (Figure 7, 'b'-'k') which include the dominant isoform ('b'). In some of these splicing variants, exons 3 and 5

have also been spliced out (Figure 7, isoforms ‘c, d, f, i’). Exon 6 encodes a proline-rich, hydrophobic sequence comprising 137 (porcine) to 161 (bovine) amino acids. This difference in size is linked to the number of repeats of a proline-glutamine motif (Brookes et al., 1995). As exon 6 is a large exon, many internal splicing sites have been identified at the transcript level, favouring alternative splicing (see isoforms ‘e’-‘i’ in Figure 7) (Veis, 2003). Removal of the 5’ section of exon 6 generates an alternatively spliced mRNA encoding the 59-amino acid leucine-rich amelogenin peptide (LRAP), which was found in the developing enamel matrix of various mammalian species including mouse, rat and cow (Fincham et al., 1981, Fincham et al., 1982, Fincham et al., 1983, Gibson et al., 1991, Lau et al., 1992, Bonass et al., 1994). LRAP is rather acidic, with a pI = 4.1. It comprises the N- and C-terminal charged regions of the full-length amelogenin. Although its precise function in enamel development is still unknown, *in vitro* studies have been carried out to determine its structure, assembly (Tarasevich et al., 2015) and function, supporting its role in controlling enamel crystal growth (Shaw et al., 2004, Le Norcy et al., 2011).

The rarely used exon 4, encodes a short sequence of 14 - 17 amino acids depending on species. The amelogenin protein isoforms containing exon 4 (Figure 7a) are only found in the matrix in minor amounts (Yamakoshi et al., 1994). As highlighted by Brookes et al (1995), exon 4 is not well conserved: unlike the translated products of other exons, porcine and human (cDNA) exon-4-peptides share only two amino acids homology (Brookes et al., 1995). Present only in mammalian genes, exon 4 was predicted *in silico* to appear in ancestral therians (220 Ma) and be functional from artiodactyl order (104 Ma) onwards (Sire et al., 2012).

Additional exons exist, found in amelogenin genes of eutherian mammals. Sire et al (2012) hypothesise that these are a possible evolutionary trait. Exon 4b is a copy of exon 4, found downstream of Exon 7. It is present in the amelogenin gene of rodents that diverged from squirrels. Not detected in transcripts, it is assumed to be spliced out (Bartlett et al., 2006a). In rodents, a further 2 exons have been identified downstream of exon 7. Exons 8 and 9 were identified in mRNA splicing variants in mouse and rat (Li et al., 1998, Papagerakis et al., 2005, Bartlett et al., 2006a) and at the protein level represent the largest amelogenin gene products secreted during rodent amelogenesis. Exon 8 results from a duplication of exon 5, translocated downstream of exon 7. Although the nucleotide sequences are 91% identical, the resulting amino acid sequence is only 50% conserved (Bartlett et al., 2006a). Transcripts comprising exon 8 also comprised exon 9. Exon 9 was predicted *in silico* to appear in the gene of ancesteral placental mammals (176 Ma) as a downstream sequence, but was activated later (50 Ma) in rodents (Sire et al., 2012) as exons 4b and 8 (Bartlett et al., 2006a) appeared.

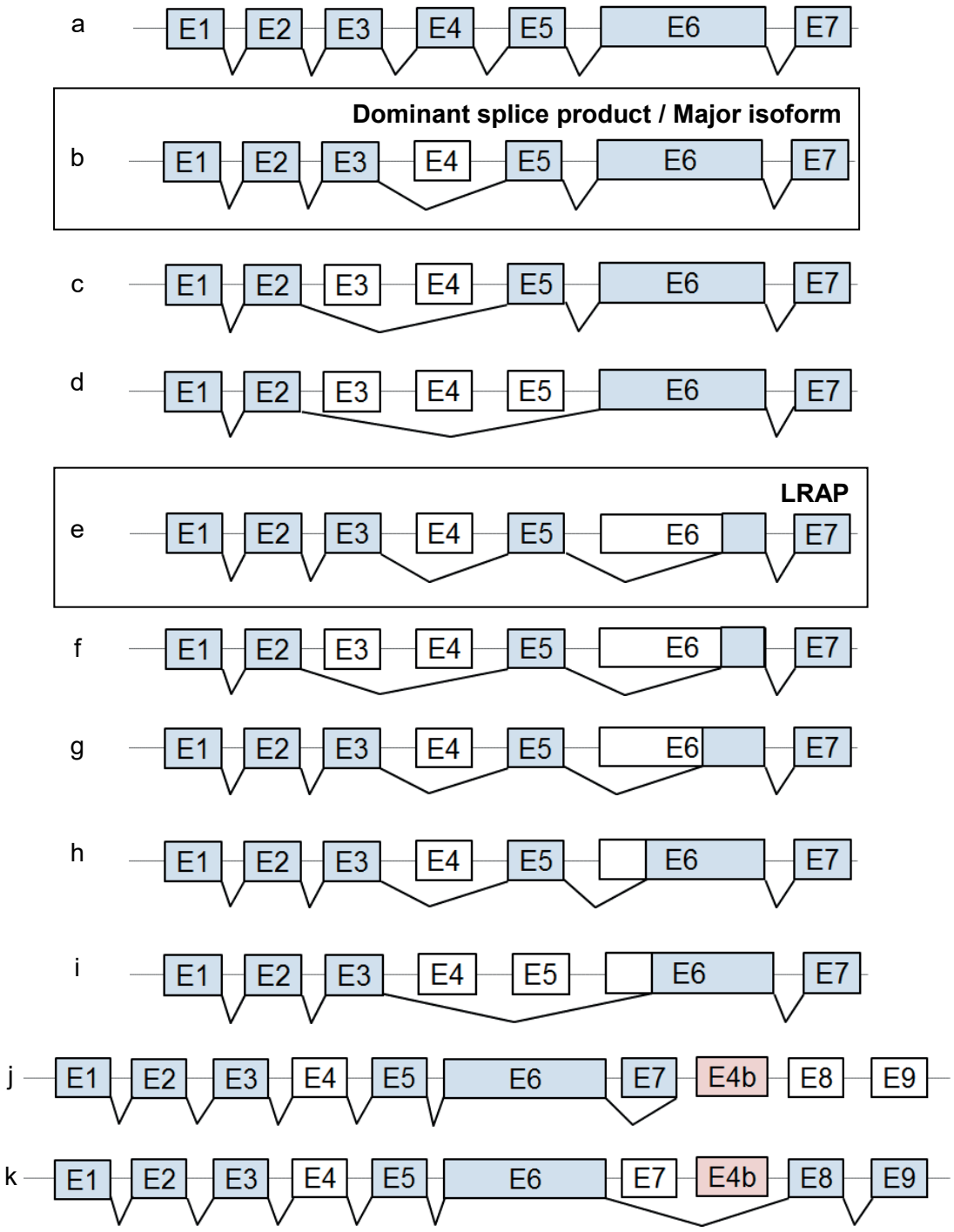


Figure 7 Alternative splicing variants of amelogenin in mammals. The isoforms 'b' (Dominant splice product) and 'e' (LRAP), in boxes, are the most abundant splicing isoforms identified in enamel matrix (Gibson et al., 2011). Rodent genes include an additional three exons downstream of E7: E4b, E8 and E9. In the rodent transcripts identified, E4b is spliced out; E8 and E9 were found, always together (k) but never with E7 (j).

1.2.1.3 Properties of amelogenin protein

The secreted amelogenin product (excluding the N-terminal signal peptide) of the major mRNA transcript is conserved between mammalian species (Brookes et al., 1995). The full-length amelogenin protein is largely hydrophobic, its isoelectric point is in the range of neutrality, e.g. murine amelogenin has a pI of 6.6 (Simmer et al., 1994, Tan et al., 1998) . Its hydrophobicity is a reflection of the high abundance of proline and leucine in amelogenin (Eastoe, 1965, Fang et al., 2011).

1.2.1.3.1 Primary structure

The primary amino acid sequence of mouse amelogenin (the major isoform secreted) has been known since the 1980s, when it was determined by cDNA sequencing (Snead et al., 1985) and by Edman sequencing of translated proteins (Takagi et al., 1984). Major native amelogenin molecules have predicted molecular weights ranging between 16-21 kDa on secretion depending on the animal species but on SDS PAGE their electrophoretic mobility is decreased, which leads to overestimation of the molecular weights. Hence, the full-length amelogenin (major isoform) is often referred to as the "25 kDa" amelogenin in the literature. Such shift in electrophoretic mobility has been linked to amelogenin's high proline content (Termine et al., 1980), after the presence of proline was previously shown to affect the mobility of other proteins (de Jong et al., 1978). A more useful nomenclature has found favour in the literature in recent years. The amelogenin in question is identified with a letter to denote species and a number corresponding to the number of amino acids present in the protein. For example P173 and M180 denote the secreted unprocessed amelogenin translated from the dominant amelogenin mRNA transcripts (see Figure 7b) in pig and mouse respectively.

The most abundant newly secreted amelogenin (excluding exon 4) is predicted to be hydrophobic overall with a hydrophilic C-terminal fragment, based on a Hopp & Woods hydrophilicity plot of recombinant rM179[1] (Fincham et al., 1994). It is divided into three domains, of which the N-terminal and C-terminal are highly conserved between species (Brookes et al., 1994, Brookes et al., 1995, Margolis et al., 2006). The first 45 amino acids comprise a rather hydrophobic domain,though still containing charged residues. Remarkably, this domain also has a relatively high tyrosine content (6 tyrosine residues), hence its name 'tyrosine-rich amelogenin peptide" (TRAP). It comprises a run of 13 amino acids that contain 3 tyrosine residues that has been

[1] rM179 corresponds to the recombinant M180 expressed by *E. coli*, which has lost Met[1] due to *E. coli* endogeneous methionine aminopeptidases (Simmer et al., 1994).

shown to have functional importance: this so-called amelogenin trityrosyl motif peptide (ATMP), was reported to recognise and bind N-acetylglucosamine (Ravindranath et al., 1999, Ravindranath et al., 2000). The importance of the ATMP sequence is discussed in greater details in Section 1.2.2.2.1 pp.21 - 22 . The central domain of amelogenin is particularly hydrophobic, being made up of XXP repeats (wherein X is often glutamine). The number of XXP repeats differs between species, reflecting the variability in length of exon 6 in the mature transcript (Fang et al., 2011, Margolis et al., 2006). The C-terminal domain comprises the last 11-15 amino acids and is highly hydrophilic, with a pI of 4.65. Studies using the LRAP variant showed that this "telopeptide" has the capacity to bind to hydroxyapatite crystals (Shaw et al., 2004, Shaw et al., 2008), which researchers suggested could be associated with the control of crystal growth, and also "a potential candidate for promoting crystal nucleation" (Tarasevich et al., 2007). The latter assumption appears however contradictory with a previous mouse KO study, which reported that amelogenin is not necessary to the formation of mineral crystals (Gibson et al., 2001). A further report suggested that the telopeptide may also drive the ordered supramolecular assembly of amelogenin nanospheres within the enamel matrix (Fang et al., 2011).

To date, one site of amelogenin post-translational phosphorylation has been identified at serine 16 (Ser^{16}) (Fincham and Moradian-Oldak, 1993) . The role of phosphorylation in amelogenesis has therefore been a subject of interest. Phosphorylation of Ser^{16} *per se* has been studied *in vitro* using the 20 kDa porcine amelogenin processing product (P148), the major degradation product found in enamel matrix (Yamakoshi et al., 1994) and full-length amelogenin (P173). In comparison with their respective non-phosphorylated counterparts produced in *E. coli* using recombinant DNA technology, the native phosphorylated P148 and P173 amelogenins are capable of regulating mineral formation and stabilizing amorphous calcium apatite for longer periods of times (>1 day) (Kwak et al., 2009, Kwak et al., 2011). Fourier transform infrared spectroscopy (FTIR) measurements comparing phosphorylated LRAP with non-phosphorylated LRAP showed that phosphorylation affects the secondary structure of this splicing isoform, and strongly suggested that it is a determinant factor involved in the interactions of LRAP with hydroxyapatite and/or amorphous calcium apatite (Yamazaki et al., 2017). Using a quartz crystal microbalance (QCM), researchers found that phosphorylation of amelogenin (native porcine P173) affected amelogenin conformation and binding to hydroxyapatite surfaces (Connelly et al., 2016). The only *in vivo* studies considered to date have consisted of testing the effect of experimentally induced single point mutations of Ser^{16} to alanine leading to a significantly affected enamel, as presented by Dr. H.C. Maroglis's group in Enamel IX International symposium (Kirkham et al., 2017), abstract no. 74. However, one major criticism of

this approach is whether the affected phenotype is caused by the loss of the phosphate group or whether it is due to the presence of alanine, a non-polar amino acid that may cause a gain of toxicity independent from the loss of the phosphate group.

Apart from the phosphorylation on Ser^{16}, no other post-translational modifications are known. Studies investigating potential glycosylation showed that amelogenin is not a glycoprotein. Fincham et al (1991), using immunological procedures, found that amelogenin is not glycosylated (Fincham et al., 1991b).

1.2.1.3.2 Secondary structure

Structural studies using circular dichroism (CD), FTIR and nuclear magnetic resonance (NMR) identified secondary structure motifs such as beta sheet, beta strand or alpha helices in bovine amelogenin (Renugopalakrishnan et al., 1986, Renugopalakrishnan et al., 1989, Delak et al., 2009). These motifs were found locally without forming further any ordered structured domain. Most studies agree that amelogenin is an unstructured or intrinsically disordered protein (IDP) (Delak et al., 2009, Lakshminarayanan et al., 2009, Ruan and Moradian-Oldak, 2015).

Amelogenin has a high proline content, which is a disorder-promoting residue (Theillet et al., 2013) and also possibly responsible for polyproline II (PPII) motifs, which make extended and flexible structures. Lacking stabilising hydrogen bonds, the PPII motif is most likely to be an open and flexible structure (Williamson, 1994, Tompa, 2002). Reviewed by Adzhubei et al, (2013), PPII regions are most common in IDPs. Favouring low affinity interactions, and simultaneously allowing a high specificity of recognition, these motifs are favourable to protein assembly and have been implicated in protein recognition, signal transduction, or protein complex assembly. PPII motifs have also been found in proteins involved in conformational diseases such as Prion PrPC proteins, α-synuclein, tau protein, Alzheimer extracellular amyloid β peptide fragment (Adzhubei et al., 2013).

In amelogenin in particular, PPII structures were identified by CD in recombinant rP172 and in rP148 (identical to the native P173 and P148 respectively but missing the N-terminal methionine and the phosphate group on Ser^{16}; rP148 was engineered with an extra methionine in position 149). Increasing the protein concentration from 5 µM to >62.5 µM was reported to favour formation of β-sheet structures (Lakshminarayanan et al., 2007, Delak et al., 2009). Within the temperature range 5-45°C, β-sheet tended to form “at the expense of PPII structure” (Lakshminarayanan et al., 2009). Therefore, under physiological conditions it is possible that both β-sheet and PPII structures might contribute towards amelogenin assembly and aggregation

properties (though it is unclear what impact the lack of N-terminal methionine and the phosphate group on Ser^{16} would have on these data).

1.2.1.3.3 Tertiary structure

Efforts have been carried out to solve the crystallographic structure of amelogenin. Crystallographic data reported in 2005 showing a birefringent microribbon structure (Du et al., 2005), was later found to be due to the presence of a cellulose contaminant, as highlighted in the erratum by Du et al (2005)The difficulty in crystallising amelogenin may be due to its high proline content since proline-rich regions and/or PPII structures make it difficult to crystallise certain proteins (Williamson, 1994). To date, no amelogenin tertiary structure derived from crystallography has been reported. Due to the lack of any X-ray crystallographic structure or any other template, *in silico* prediction using homology modelling has not been achieved.

1.2.1.3.4 Quaternary structure: amelogenin self-assembly

The formation of quaternary structures of amelogenin is a highly complex phenomenon due to amelogenin self-assembly and aggregation properties, which are of fundamental importance in amelogenesis. This has therefore attracted a great deal of interest and is discussed in greater details in Section 1.2.2 below.

1.2.2 Amelogenin self-assembly

1.2.2.1 Amelogenin propensity to aggregate

The PPII motif, in combination with the hydrophobic domains of amelogenin, appears to favour hierarchical self-assembly of amelogenin as suggested by variable temperature CD and isothermal titration calorimetry (ITC) studies (Lakshminarayanan et al., 2007, Lakshminarayanan et al., 2009). PPII is found in protein regions available for recognition and binding (Adzhubei et al., 2013). Early reports suggested that PPII motifs are responsible for the temperature-dependent but reversible propensity for amelogenin to aggregate (Nikiforuk and Simmons, 1965, Mechanic et al., 1967). This characteristic behaviour of amelogenin has been further characterised by high-resolution size exclusion chromatography (HR-SEC), dynamic light scattering (DLS), atomic force microscopy (AFM) and transmission electron microscopy (TEM) (Fincham et al., 1994, Moradian-Oldak et al., 1994). The size of the amelogenin aggregates increases with temperature in the range 5 - 37°C (Moradian-Oldak et al., 1998).The formation of these aggregates has been linked to the low solubility of amelogenin, shown by quantitative analyses using recombinant mouse amelogenin rM179. The r-amelogenin is only sparingly soluble at physiological pH and temperature (Simmer et

al., 1994, Tan et al., 1998), as physiological pH is around amelogenin's predicted isoelectric point of 6.8. Amelogenin solubility was also found to decrease with higher ionic strength of the solution (Tan et al., 1998).

1.2.2.2 Amelogenin hierarchical assembly

Despite being aggregative, amelogenin self-assembles in an ordered fashion, to form hierarchical structures (Fang et al., 2011), which are presumed to be necessary to guide the formation of enamel crystals. This Section addresses: i) the domains involved in amelogenin self-assembly (Section 1.2.2.2.1) , ii) evidence of the first level of supramolecular structures which form in the ameloblasts (Section 1.2.2.2.2) and iii) the assembly into larger structures found in the extracellular matrix (Section 1.2.2.2.3).

1.2.2.2.1 Domains involved in amelogenin recognition and self-assembly

Amelogenin self-assembly was characterised by Y2H assay conducted by Paine and Snead (1997). The authors confirmed that amelogenin (M180) is capable of binding to itself with a greater affinity than that seen for protein p53 binding to SV40 large T-antigen. p53 and SV40 large T-antigen were used as positive controls as they are well-known binding partners, yielding 150 β-galactosidase (Miller) units in the Y2H assay (Li and Fields, 1993). Two domains were found to be necessary for amelogenin self-assembly: the N terminal 42-residues (Met^{1}-Met^{42}) A domain and the 16-residue (Ser^{157} – Lys^{173}) C-terminal B domain (Paine and Snead, 1997). Deletion experiments *in vitro* and *in vivo* confirmed the importance of the A and B domains in the formation of supramolecular assemblies (Paine et al., 2000). The N-terminal 42 residues, which comprise the majority of the TRAP domain as described in Section 1.2.1.3.1 p. 17 (and part of the hydrophobic core of amelogenin) specifically drives the formation of supramolecular amelogenin assemblies (so-called nanospheres). The C-terminal 16 residues (hydrophilic) were reported to be essential in maintaining the integrity and stability of these nanospheres (Paine and Snead, 1997, Moradian-Oldak et al., 2000) by preventing clustering of hydrophobic cores of amelogenins. Later, another Y2H assay testing different truncated mouse r-amelogenins indicated that Pro^{169} plays a pivotal role in driving self-assembly, but point mutations Pro^{169} →Thr or Pro^{169} →Lys surprisingly did not affect significantly affect the self-assembly (Paine et al., 2003b). *In vitro*, a study comparing the aggregates formed by recombinants rM166 and rM179 suggested that the C-terminal telopeptide was essential to form stable and monodisperse aggregates (Moradian-Oldak et al., 1994) The r-amelogenins rM166 and native P161 lacking the conserved C-terminal telopeptide, had a larger size distribution and a reduced affinity for hydroxyapatite (Moradian-Oldak et al., 2002). They formed larger aggregates than their full-length precursors, as measured by DLS

and small-angle X-ray scattering (SAXS) (Moradian-Oldak et al., 2002, Aichmayer et al., 2005). Native P148, which also lacks the C-terminal telopeptide as well as the Pro^{162} (corrresponding to the mouse amelogenin Pro^{169}) was also capable of forming nanospheres, though of a smaller size (Moradian-Oldak et al., 2002). DLS data and fluorescence studies using recombinant porcine amelogenins found that rP148[2] formed similar oligomers, even though smaller, as rP172 (Bromley et al., 2011). This clearly questions the importance of the C-terminal telopeptide and the Pro^{169} (in mice, or Pro^{162} in pig) in amelogenin supramolecular assembly.

While both A and B domains are present in the full-length amelogenin (major splicing isoform) sequence, alternatively spliced variants such as LRAP lack these domains. Paine and Snead (1997) showed that LRAP does not self-assemble (Paine and Snead, 1997). LRAP lacks the sequence Tyr^{34}-Met^{42}, i.e. the ATMP domain; their finding suggests that the ATMP is a necessary, though not sufficient, driver of amelogenin self-assembly. The importance of the ATMP motif in amelogenin self-assembly has been supported by Wald et al (2017), who focused on the evolutionarily conserved “Y/F-X-X-Y/L/F-X-Y/F” motifs (including the amino acids sequence YPSYGY), present in amelogenin and ameloblastin. Interestingly such motifs are short linear sequences rich in hydrophobic residues and are well-known drivers of IDP assembly (Mészáros et al., 2007). Comparing the WT (YPSYGY) and mutant (GPSGGG) recombinant human amelogenins and using HR-SEC, TEM and surface plasmon resonance (SPR), Wald et al showed that the integrity of the YPSYGY motif is necessary for amelogenin self-assembly, and possibly amelogenin-ameloblastin recognition (Wald et al., 2017). This latter observation supports previous works by Ravindranath et al, discussed in further details later (Section 1.2.5.2, pp. 45 - 49)

1.2.2.2.2 Intracellular assembly of amelogenin *in vivo*

Cross-linking studies in rat enamel organs showed that amelogenin begins to self-assemble within the ameloblast secretory pathway. Diagonal SDS PAGE analysis followed by western blotting with anti-amelogenin antibodies showed the presence of dimers, tetramers, pentamers and hexamers of full-length M180 amelogenin. Through cross-linking, hexamers appeared to be the biggest unit forming intracellularly (Brookes et al., 2006) prior to further assembly steps in the extracellular matrix. These levels of assembly are detailed below, in Section 1.2.2.2.3.

---

[2] For clarification, recombinant rP148 used by Bromley et al. (2011) differs from the native P148 since it lacks Met^{1} and Ser^{16} phosphorylation, and carries and extra methionine at position 149.

1.2.2.2.3 Extracellular assembly of amelogenin observed *in vivo*

Having demonstrated that amelogenin assembly begins in the intracellular secretory pathway (Brookes et al., 2006), Brookes et al. (2006) hypothesised that the hexamers undergo further assembly within the newly secreted matrix. The first stages of hexamer assembly may lead to the formation of transient aggregates, which may constitute the stippled material observed near the Tomes' processes in developing enamel (Fincham et al., 1994). The transient aggregates further assemble into spherical structures, 'nanospheres', each comprising approximately 100 amelogenin units (Moradian-Oldak et al., 2002). Freeze-etching and AFM of developing enamel in rat incisors showed the presence of 30-50 nm diameter nanospheres (Robinson et al., 1981). TEM of mouse, bovine and hamster developing enamel also showed spherical electron-lucent structures of 20-50 nm diameter with uranyl acetate staining; these nanospheres appear themselves to be organised as arrays of beads surrounding the forming enamel crystallites (Fincham et al., 1995). Fincham et al (1995) assumed that the nanospheres observed in the TEM were made up of amelogenin aggregates. They appeared electron-luscent because the hydrophobic nature of amelogenin may prevent them from binding to the heavy metal contrasting agent used to stain the TEM sections by increasing the electron absorption of specific structures comprising the section linking their observations to previous experiments where purified porcine amelogenins showed negative heavy metal staining when contrasted with ammonium molybdate.

In addition to the *in vivo* studies described above, a plethora of *vitro* studies has been carried out to characterise amelogenin assembly. The literature reviewed below (in Section 1.2.2.2.4) used r-amelogenin rM179 (Simmer et al., 1994), an analogue of major mouse amelogenin M180, lacking phosphorylation in Ser^{16} and the N-terminal Met. Amelogenin assembly studies discussed below also used porcine 25 kDa full-length amelogenin, P173, which can be obtained from pig developing teeth obtained as a by-product of the meat industry (Aoba et al., 1987a, Limeback, 1987).

1.2.2.2.4 Characterisation of amelogenin self-assembly *in vitro*

The solution properties of r-amelogenin, i.e, its propensity to self-assemble and aggregate were thoroughly investigated *in vitro* using rM179. The techniques employed include DLS, HR-SEC, TEM and AFM. While the assembly of amelogenins in macromolecular structures is hierarchical, *in vitro* studies reported various assembly modes: nanospheres, micelles, or microribbons, as discussed below and illustrated in Figure 8.

- *Nanospheres*

The self-assembly of amelogenins into nanospheres (briefly mentioned previously in the Section 1.2.2.2.1) has been supported since the 1990s using rM179 (Fincham et al., 1994, Moradian-Oldak et al., 1998). “Spherical aggregate structures” of 12-18 nm diameter were observed using TEM on parlodion-covered carbon-coated copper grid, while 15-20 nm diameter spheres were shown by AFM on mica surface (Fincham et al., 1994, Wen et al., 2001), which is consistent with the nanosphere structures observed *in vivo,* described above in Section 1.2.2.2.3. DLS studies at pH 7-8 in Tris-HCl showed that their molecular weights ranged between 2000 - 3800 kDa, which suggests that they are made up of 100-200 amelogenin monomers {Moradian-Oldak, 1994 #142}. A cryoelectron microscopy study of rM179 At pH 8 in Phosphate-buffered saline (PBS) supported the formation of 15-20 nm diameter nanospheres at room temperature within 10 minutes of incubation (Fang et al., 2011). A combination of DLS and SAXS studies suggested that these nanospheres comprise a dense hydrophobic core (detected by SAXS), surrounded by a looser shell of protein that was hypothesised to be comprised of the C-terminal hydrophilic telopeptide (Aichmayer et al., 2005).

The formation of the nanospheres was shown to be pH, temperature and concentration-dependent. Initial DLS studies at room temperature (25-29°C) indicated that the size and distribution of aggregates varies with pH (Moradian-Oldak et al., 1994, Moradian-Oldak et al., 1998). At pH values below 6, aggregates were relatively small (hydrodynamic radii 4-7 nm in Tris-HCl, or 3.9 - 4.5 nm in sodium acetate at pH 4.4) whereas at pH 7.4, the hydrodynamic radius of the aggregates was increased to 19.2 nm. This might be expected as pH values below 6 would induce a greater charge on rM179, making it more soluble and less prone to aggregation. The increased size of the aggregates at physiological pH is linked to their decreased solubility. At pH values ≥ 7.8, which is significantly above the pI of rM179, the solubility of amelogenin increased, and the size of the aggregates decreased to 15-16 nm hydrodynamic radii (Tan et al., 1998, Moradian-Oldak et al., 1998).

DLS studies also showed that nanosphere assembly is sensitive to environmental temperature: Below 25°C, the nanosphere hydrodynamic radius was 15-18 nm, whereas at 37-40°C, it increased to 60-70 nm (Moradian-Oldak et al., 1998). Moradian-Oldak et al (1998a) observed that variation of pH between values 4.4 and 8 affected the size of amelogenin aggregates significantly more at 37°C than at 25°C. While the size of the aggregates reached a maximum radius of 25 nm at pH=7.4 and a temperature of 25°C; this increased to 35 nm at pH=7.4, and then to 58 nm at pH=8, at 37°C.

The hydrodynamic radius of amelogenin aggregates also increases with ionic strength. For example at pH 7.4, increasing the ionic strength from 0.05 M to 0.16 M increased nanosphere hydrodynamic radius from 19.2 nm to 28.8 nm (Moradian-Oldak et al., 1994). The physiological ionic strength of enamel fluid is 0.164 mM (Aoba and Moreno, 1987), suggesting that physiological temperature and ionic strength favours amelogenin aggregation and the assembly of larger nanospheres.

In light of the effect of temperature on amelogenin interactions we may question the relevance of amelogenin *in* vitro aggregation/binding experiments when these are carried out at non-physiological temperatures. A typical example is the mechanism of amelogenin hierarchical assembly into nanospheres, proposed by Fang et al (2011) from data obtained at 25°C. In their report, the authors hypothesised that 1) amelogenins form dimers via their C-terminal extremities, 2) six dimers associate side-by-side to form dodecamers and 3) dodecamers associate to form nanospheres (Fang et al., 2011). Fang et al's hypotheses contrasts with previous knowledge on amelogenin assembly in that the ATMP motif is necessary in amelogenin self-assembly in first place (discussed in the Section 1.2.2.2.1, p. 22), and the C-terminal hydrophilic peptide DKTKREEVD is not sufficient to trigger amelogenin-amelogenin interaction. For instance, LRAP does not self-assemble as indicated by Y2H studies (Paine and Snead, 1997).

- *Micelles*

Other modes of amelogenin assembly have been proposed. In 2007, Fukae et al (2007) speculated that amelogenins associate as micelles, using porcine 25 kDa amelogenin and its proteolytic processing products (Fukae et al., 2007). Their hypothesis was based on the Kyte and Doolittle amphiphilicity index and Robson secondary structure prediction. They claimed that the N-terminal amino acids (Met^{1} to Trp^{45}) form the hydrophobic core of the micelles while the 13 kDa (Leu^{46} to Ser^{148}) domain acts as a rod-like structure that connects the micellar core to the hydrophilic C-terminal segment comprising Lys^{166}- Asp^{173}, which is exposed at the surface of the micelle facing the aqueous environment. Fukae et al (2007) add that the highly

charged C-terminal segments form α-helices and can interact with each other by intermolecular ionic interactions that mediate amelogenin-amelogenin interactions, and furthermore, inter-micelle assembly and aggregation. There is, however, a lack of evidence supporting this micelle model and there is a caveat to consider. Given that amelogenin is reportedly an IDP and that no tertiary structure has been found to date (see Section 1.2.1.3.3, p. 20), it is difficult to reconcile the *in silico* predicted structural features Fukae et al. (2007) describe with the experimental data obtained by others.

- *Ribbons*

Other groups reported that amelogenin supramolecular structures are ribbon-like with amelogenin monomers either assembling directly into so-called nanoribbons or amelogenin nanospheres assembling into so-called microribbons. A first X-ray structure was published in 2005. The authors of that report claimed that the birefringent microribbons visible by synchrotron X-ray diffraction (XRD) and SEM were supramolecular assemblies of amelogenin nanospheres (Du et al., 2005). The report was later retracted (2005), following Beniash et al's observation that the fibrillar structure observed was actually that of contaminanting cellulose fibres. Moradian-Oldak et al (2006) later maintained that amelogenin supramolecular assembly results in the formation of microribbons using FTIR and Raman microspectroscopy, and proposed that amelogenin nanospheres assemble into collinear arrays which correspond to microribbons. They reported that these features had not yet been verified *in vivo* (Moradian‐Oldak et al., 2006). One criticism of this study is that polyethylene glycol, which is a protein precipitant (Ingham, 1984) was used to drive the amelogenin to form ribbons. It is unclear how relevant these experimental conditions are to the situation *in vivo*. A recent paper by Carneiro et al (2016) claimed that human amelogenin rH174 self-assembles into nanoribbons *in vitro*. Using AFM and TEM; they proposed that an N-terminal segment (Gly^{8} - Thr^{21}) drives amelogenin self-assembly into nanoribbons through β-sheet interactions (Carneiro et al., 2016). However these are *in silico* predictions, and the formation of nanoribbons *in vitro* was observed at acidic pH (4.5-5.5 or 6.5), which again raises questions around the relevance of their findings to the situation *in vivo*.

To summarise, most reports agree that amelogenin assembly is an ordered hierarchical process, though different modes of assembly have been reported (summarised below in Figure 8). The most plausible model of assembly to date is nanospheres of 20-50 nm diameter, which could be the spherical structures observed in developing mouse enamel (Fincham et al., 1995). Further *in vivo* investigations are needed to confirm the precise identity of these structures e.g. are they pure amelogenin or are other enamel proteins involved?

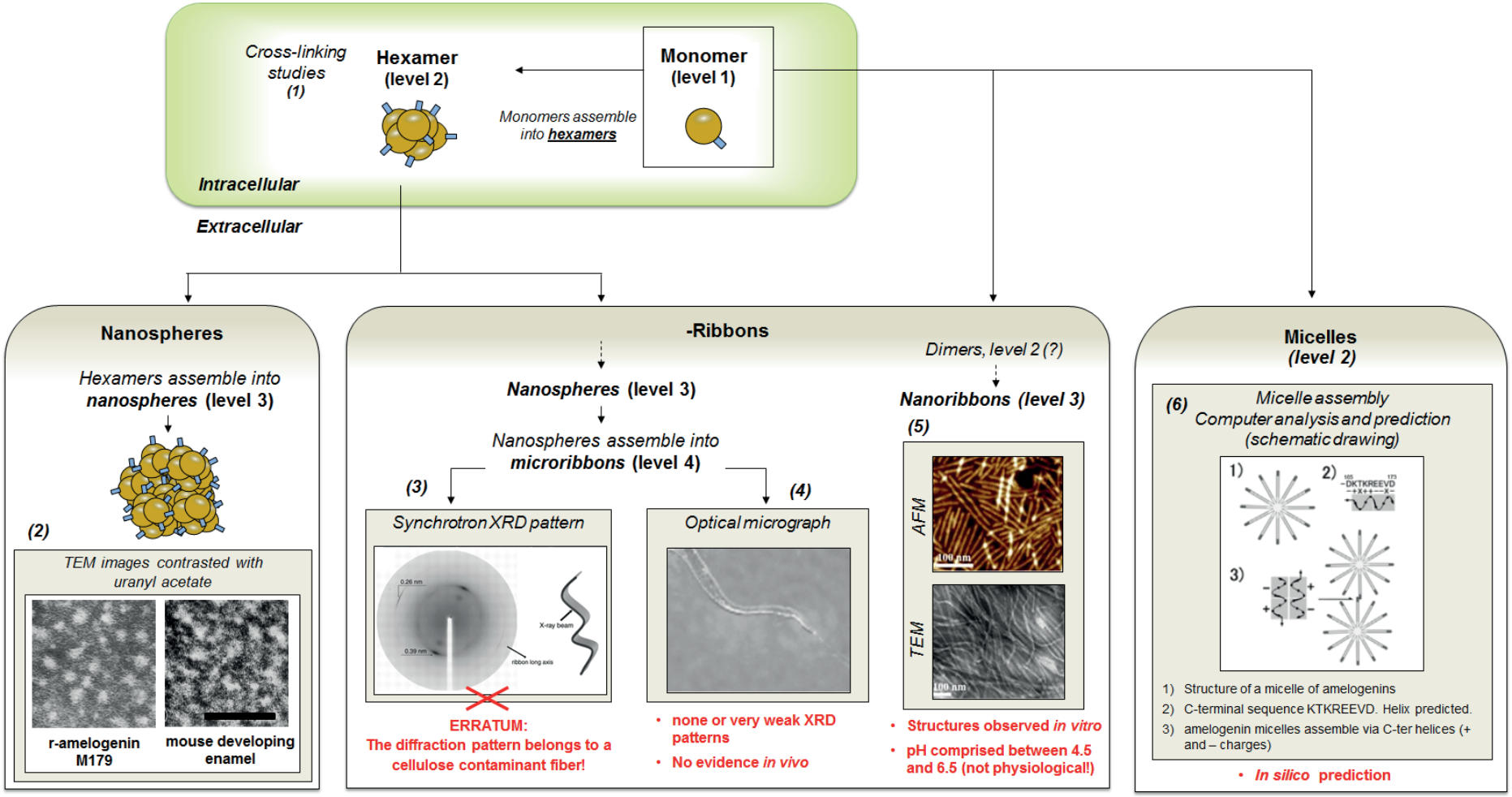

Cross-linking studies (1)
Hexamer (level 2)
Monomer (level 1)
Monomers assemble into hexamers
Intracellular
Extracellular
Nanospheres
Hexamers assemble into nanospheres (level 3)
(2)
TEM images contrasted with uranyl acetate
r-amelogenin M179
mouse developing enamel
-Ribbons
Nanospheres (level 3)
Nanospheres assemble into microribbons (level 4)
(3)
Synchrotron XRD pattern
0.26 nm
0.39 nm
X-ray beam
ribbon long axis
ERRATUM:
The diffraction pattern belongs to a cellulose contaminant fiber!
(4)
Optical micrograph
• none or very weak XRD patterns
• No evidence in vivo
Dimers, level 2 (?)
Nanoribbons (level 3)
(5)
AFM
TEM
100 nm
100 nm
• Structures observed in vitro
• pH comprised between 4.5 and 6.5 (not physiological!)
Micelles (level 2)
(6)
Micelle assembly
Computer analysis and prediction
(schematic drawing)
1)
2)
3)
1) Structure of a micelle of amelogenins
2) C-terminal sequence KTKREEVD. Helix predicted.
3) amelogenin micelles assemble via C-ter helices (+ and – charges)
• In silico prediction

Figure 8 Suggested hierarchical assembly of amelogenin in supramolecular structures: nanospheres, microribbons or micelles. Amelogenin assembles into hexameric units during intracellular trafficking along the secretory pathway. Supramolecular assemblies were described as nanospheres (bottom-left panel), microribbons (bottom-central panel). Another group reported, from *in silico* predictions, that amelogenin assembles into micelles (bottom-right panel). The most accepted model to date is the nanosphere. *(1)* The cross-linking studies were carried out by Brookes et al. (2006). *(2)* The TEM images are taken from Fincham et al. (1995). *(3)* The synchrotron XRD pattern is reproduced with permission from Du et al. (2005). *(4)* The optical micrograph is reproduced from Moradian-Oldak et al. (2006). *(5)* The TEM and AFM images are from Carneiro et al. (2016). *(6)* The schematic drawing is reproduced from Fukae et al. (2007). The full references for the reports and publishers are detailed in the footnote below[1].

[1] References for Figure 8:

(1) BROOKES, S. J., LYNGSTADAAS, S. P., ROBINSON, C., SHORE, R. C. & KIRKHAM, J. 2006. Intracellular nanosphere subunit assembly as revealed by amelogenin molecular cross-linking studies. *European Journal of Oral Sciences,* 114 Suppl 1**,** 280-4; discussion 285-6, 382.

(2) Reprinted from Journal of Structural Biology, Vol 115, FINCHAM, A. G., MORADIAN-OLDAK, J., DIEKWISCH, T. G., LYARUU, D. M., WRIGHT, J. T., BRINGAS, P., JR. & SLAVKIN, H. C. Evidence for amelogenin "nanospheres" as functional components of secretory-stage enamel matrix. Pages no. 50-59, Copyright (1995), with permission from Elsevier.

(3) From DU, C., FALINI, G., FERMANI, S., ABBOTT, C. & MORADIAN-OLDAK, J. 2005. Supramolecular assembly of amelogenin nanospheres into birefringent microribbons. Science, 307, 1450-1454. Reprinted with permission from AAAS.

(4) MORADIAN-OLDAK, J., DU, C. & FALINI, G. Ibid.On the formation of amelogenin microribbons. 114**,** 289-296. Copyright © 1999-2018 John Wiley & Sons, Inc. All rights reserved.

(5) CARNEIRO, K. M., ZHAI, H., ZHU, L., HORST, J. A., SITLIN, M., NGUYEN, M., WAGNER, M., SIMPLICIANO, C., MILDER, M. & CHEN, C.-L. 2016. Amyloid-like ribbons of amelogenins in enamel mineralization. *Scientific Reports*, 6, 23105. Work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copy of the licence provided in link http://creativecommons.org/licenses/by/4.0/.

(6) Material taken from FUKAE, M., YAMAMOTO, R., KARAKIDA, T., SHIMODA, S. & TANABE, T. 2007. Micelle Structure of Amelogenin in Porcine Secretory Enamel. *Journal of Dental Research,* 86**,** 758-763. with permission from SAGE Publications as Gratis reuse.

1.2.3 Matrix biochemistry and processing of amelogenin

Understanding the process of amelogenin self-assembly (discussed in the previous Section) and its interactions with other proteins and mineral within the matrix is a critical step in determining the role of amelogenin in amelogenesis. Amelogenin is a major and essential component of developing enamel matrix, suggested to control the growth of enamel crystals (Wallwork et al., 2001). As mentioned in the Section 1.1.2.3 (pp. 8 - 9), the enamel crystallites sketched in during the secretory stage of amelogenesis appear thicker at the dentine-enamel junction (older enamel) and thinner at the secretion front (newer enamel). This correlates with a progressive proteolytic processing of amelogenin within the enamel matrix, with full-length amelogenin present in the newer enamel, where crystals are thinner, and its processing products present in the older enamel, where the crystals are thicker and wider (Daculsi and Kerebel, 1978, Robinson et al., 1998). Once the enamel matrix has reached its full thickess maturation begins; the amelogenins and processing products are degraded and removed, then the enamel becomes fully mineralised.
This Section addresses i) The proteolytic processing of amelogenin throughout secretory stage and its possible role in the reduction in amelogenin- mineral binding (Section 1.2.3.1). Then it addresses ii) the full degradation of amelogenins during transition and maturation stages (Section 1.2.3.2).

1.2.3.1 Amelogenin processing throughout secretory stage

The full-length “25 kDa” amelogenin is restricted to the outer layer of the forming enamel and is rapidly processed after being secreted (Fukae et al., 1980, Uchida et al., 1991b). This processing is sequential and results in the formation of different processing products with various physico-chemical properties. The properties of the major processing products of amelogenin are described in the Section 1.2.3.1.1 below and how these are involved in the control of enamel crystallites growth is addressed in the Section 1.2.3.1.2.

1.2.3.1.1 Protelolytic processing during secretion

Proteolytic processing has been extensively studied in pig developing enamel. Once secreted into the secretory-stage enamel matrix, the full-length amelogenin (P173), whose apparent mobility on analytical SDS PAGE is 25 kDa, is rapidly degraded into smaller fragments (Fincham et al., 1991b). In pig secretory enamel, these include a 23 kDa, 20 kDa, 13 kDa, 11 kDa, 7 and 5 kDa fragments (Yamakoshi et al., 1994) (Figure

9), most of which were shown to be the result of cleavage by MMP-20 that is co-secreted during the secretory stage (Ryu et al., 1999, Nagano et al., 2009).
The newly secreted 25 kDa amelogenin (P173) has a high degree of polarity, with a major hydrophobic portion and an extremely hydrophilic C-terminal portion (the last 12 amino acids comprising the C-terminal telopeptide), Figure 9. It is quickly degraded by MMP-20 and as a result the intact full-length secreted product is restricted to the newly secreted enamel matrix at the surface of the developing enamel (Uchida et al., 1991b). The first MMP-20 processing step removes the hydrophilic C-terminal telopeptide, producing a 23 kDa amelogenin (P161). P161 is also present transiently (Yamakoshi et al., 1994) and is further processed resulting in the loss of its C-terminal comprising Met^{149}-Trp^{161}. The resulting processing product is P148 (commonly referred to as the 20 kDa amelogenin from SDS PAGE analysis). P148 is relatively stable in the enamel matrix, and therefore accumulates, so that it appears clearly as the major band on SDS PAGE analyses of secretory stage pig enamel (Shimizu M, 1983, Yamakoshi et al., 1994, Brookes et al., 1995, Maycock et al., 2002). It was argued that P148 might be derived from either the 23 kDa fragment or directly from the parent amelogenin. Both P161 and P148 fragments self-assemble *in vitro* to form nanospheres; the P161 fragment appears to form more heterogeneous nanospheres on AFM while the P148 fragment formed smaller nanospheres with hydrodynamic radii of 10-14 nm (Moradian-Oldak, 2001).
Further processing of 20 kDa amelogenin was shown *in vivo* to generate 5 and 13 kDa fragments (Tanabe, 1984, Aoba et al., 1987b, Fincham and Moradian-Oldak, 1993). The 5 kDa fragment is the TRAP domain, comprising the first 45 amino acids. It is relatively insoluble and is highly conserved between species (Fincham et al., 1983, Fincham et al., 1989). It is also the ‘A-domain’ involved in protein-protein interactions identified by Y2H experiments (Paine and Snead, 1997). The loss of TRAP disrupts amelogenin self-assembly, and the resulting 13 kDa fragment, comprising Leu^{46}-Ser^{148} (P46-148), does not self-assemble and is completely soluble in the enamel fluid (Brookes et al., 1995). However, there appears to be conflicting reports as to whether these lower molecular fragments are generated by the cleavage of P148 by MMP-20 (Nagano et al., 2009, Kwak et al., 2016): while Nagano et al (2009) supported that the TRAP and P46-148 result from the cleavage of P148 by MMP-20, Kwak et al (2016) argued that P46-148 would be generated by the cleavage of P173 by MMP-20 instead, supporting that MMP-20 does not cleave P148. Otherwise, a minor alternative processing route was reported to generate an 11 kDa fragment, spanning Ala^{63} - Ser^{148} (Aoba et al., 1987b).

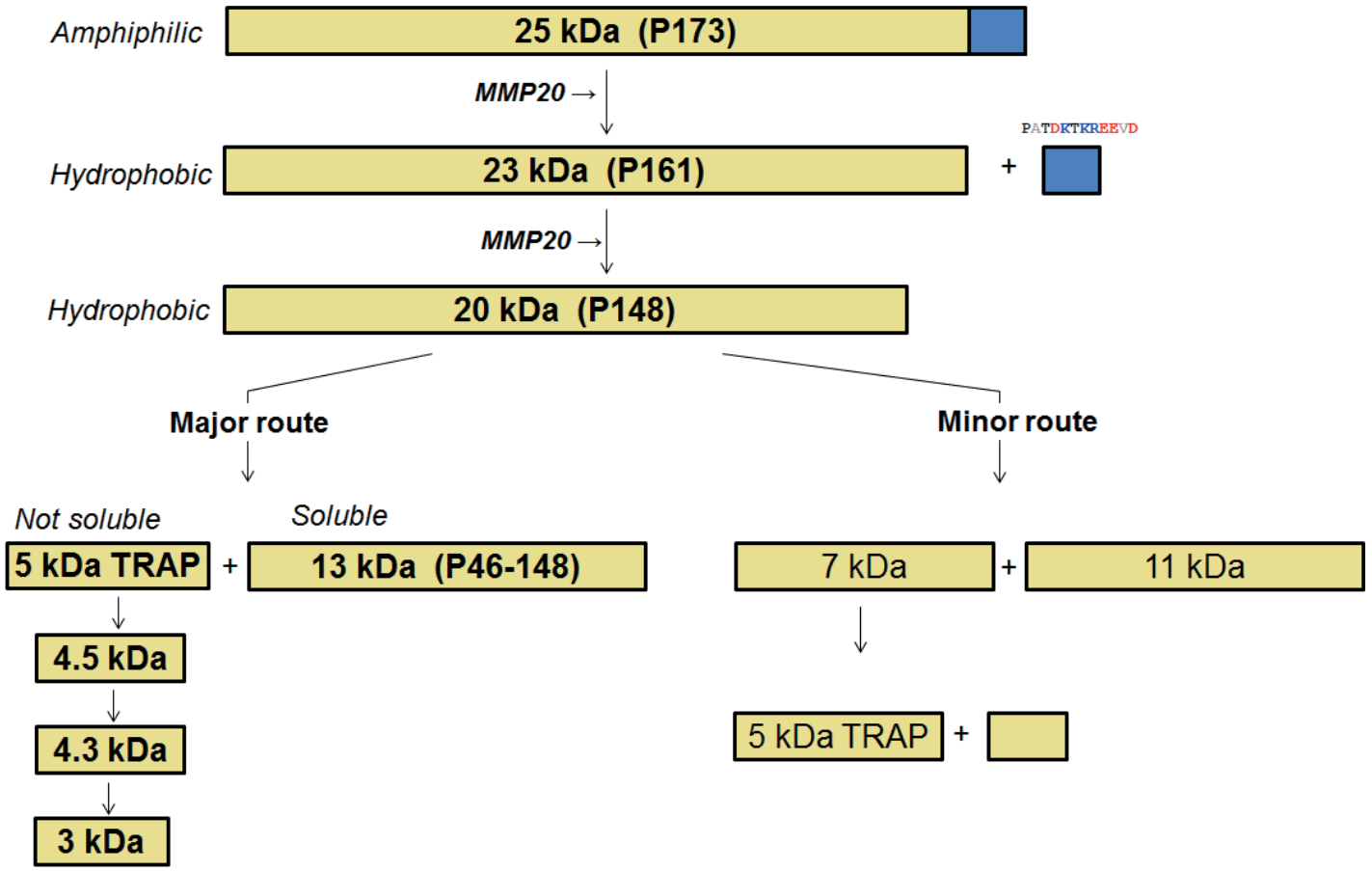


Figure 9 processing of amelogenin in developing enamel matrix (based on the porcine model).

1.2.3.1.2 Effect of amelogenin processing on mineral binding and role of amelogenin in enamel matrix mineralisation

The protein matrix of developing enamel supports and potentially directs the elongation of thin plate-like enamel crystals that grow in length (c-axis growth) (Nylen et al., 1963) as they pursue the retreating ameloblasts. By AFM, amelogenin nanospheres appeared to be organised in linear arrays, aligned with the side faces of the enamel crystals (Fincham et al., 1995, Fincham et al., 1999). In effect, the crystal architecture is lightly “sketched in” during the secretion stage but how these crystals are nucleated is not fully established. Initially, calcium and phosphate ions are present in the earliest secreted enamel (Deakins and Burt, 1944, Takano et al., 1986, Robinson et al., 1974). Crystals nucleate in the extracellular environment to form an ordered crystal array (Fincham et al., 1999), possibly in the initial form of octacalcium phosphate which later changes phase to form hydroxyapatite, which are the basic unit of the prisms observed in mature enamel. As nucleation of hydroxyapatite is not thermodymically favourable, it requires a “catalyst” to proceed. This catalyst may take the form of a proteinaceous template which promotes heterogeneous nucleation by matrix-mineral electrostatic interactions (Kirkham et al., 2002) but it is unclear which protein(s) are involved. They could originate either from the dentine or the DEJ, or from EMPs themselves. Both protein-mineral and protein-protein interactions mediate enamel crystallites and prism formation.

Aoba et al. (1987a) hypothesised that processing and degradation of amelogenins favours crystallite growth in width and thickness– suggesting that full-length amelogenin constrains enamel crystal growth to c-axis only. This hypothesis was based on the smaller crystal thickness observed in the outer layer being concomitant with the location of full-length 25 kDa amelogenin essentially in that region. Its rapid degradation in the deeper layers (Uchida et al., 1991b) was related to reduced protein-to-mineral adsorption which relieved inhibition of crystal growth and allowed the crystals present in the deeper enamel to grow in width and thickness (Aoba et al., 1987a).

However, this theory can be challenged. It must be noted that crystals in the deeper enamel are simply older and have been exposed to the mineralising environment for longer and have had more time to grow. That they appear thicker at the dentine-enamel junction and thinner at the secretion front can be explained on this basis alone in the absence of any crystal growth inhibition associated with mineral bound full-length amelogenin.

A major challenge to the hypothesis proposed by Aoba et al. (1987a) was highlighted in S.J. Brookes’s thesis (1995, pp. 172-175) citing observations and TEM

measurements of enamel crystallites' thickness and width reported by Daculsi and Kerebel in 1978 (shown below in Figure 10). These measurements do indeed indicate that crystallite thickness and width increase with depth; however, the rate of increase in width and thickness is faster in the "outer" 50 µm, (within the first 50 µm depth, from the Tomes' processes, the width and thickness increase by 160% and 220% respectively), which is where the supposedly inhibitory full-length amelogenin is found. In contrast at 200-500 µm depth from the Tomes' processes, where amelogenin degradation products dominate, the crystal width and thickness have only increased by 5.7 and 52.4% in thickness and width respectively (Daculsi and Kerebel, 1978). However, it must be rembered that the youngest crystals will be imperfect and exhibit more growth sites than older crystals so may be growing faster despite the presence of supposedly inhibitory full length amelogenin at this point. It may be that full-length amelogenin is actually damping very active growth near the surface and although the rate of growth is higher, it is less than it would be in the absence of the full length amelogenin.

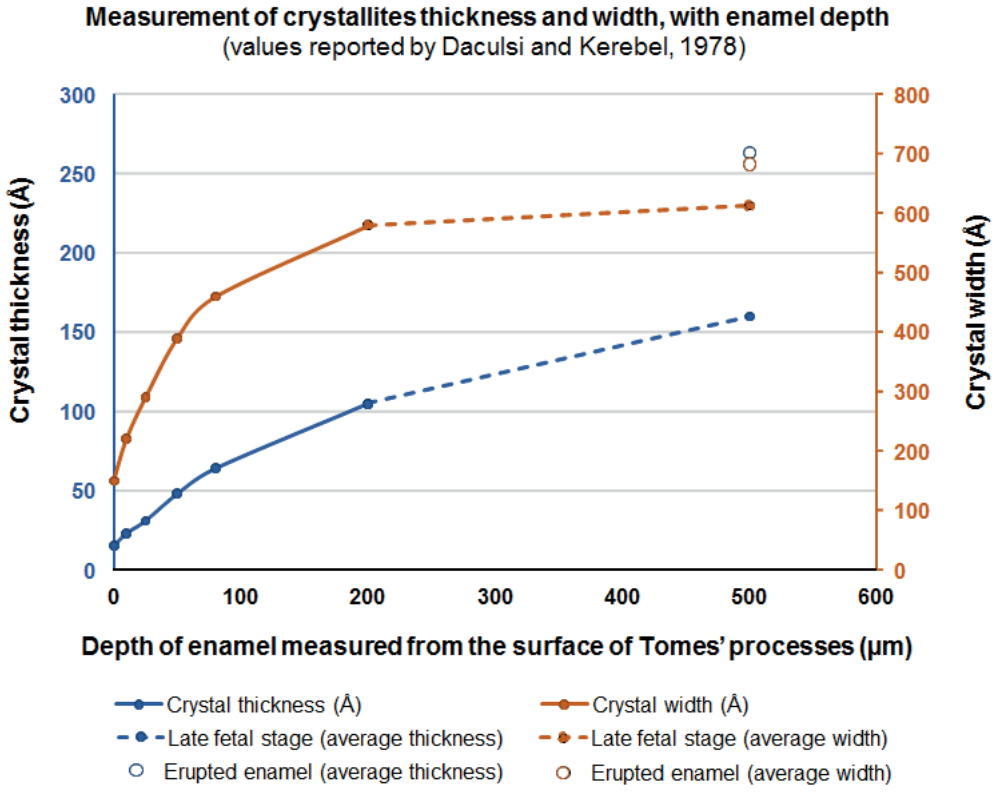


Figure 10 Growth of crystal width and thickness with enamel depth, as measured by Daculsi and Kerebel, 1978. Note the rate of growth is highest near the surface where the supposedly inhibitory full-length amelogenin is localised.

Amelogenin processing products have different biochemical properties to the full-length amelogenin and may therefore have distinct and specific roles in enamel formation. For instance, studies suggested that (25kDa) P173 guides ordered hydroxyapatite crystal formation unlike (20kDa) P148, although both are involved in amorphous calcium phosphate stabilisation (Kwak et al., 2014). The consequences of amelogenin degradation may be more complex than merely generating space for crystallite growth in width and thickness (which Aoba et al (1987) appeared to suggest, as described above).

However, in keeping with previous reports (Aoba et al., 1987a) other studies reported that amelogenin affinity for apatite surfaces is reduced on proteolytic processing. The C-terminal domain, which is charged, has strong affinity for hydroxyapatite crystals (Shaw et al., 2004) and its loss (processing of the full-length 25 kDa amelogenin) results in reduced affinity for apatite surface (Moradian-Oldak et al., 2002).The 23 kDa and 20 kDa amelogenins, present in higher ratio in the enamel matrix (with the 20 kDa amelogenin more stable), have different physico-chemical properties, *per se*, than the parent protein (more hydrophobic, illustrated in Figure 9). Such species form different supra-molecular assemblies *in vitro* to generate nanospheres with different sizes and polydispersity which may alter the structure and biochemical nature of the matrix (Moradian-Oldak, 2001). However, *in vivo*, the processing fragments are already assembled in nanospheres even as they are generated by MMP-20 processing so it is unlikely that monomeric forms of P161 and P148 are available in the matrix and the relevance of these *in vitro* findings is questionable.

In addition to proteolytic processing (detailed above in Sections 1.2.3.1.1-2), amelogenin dephosphorylation might affect structure and function of amelogenin. Dephosphorylation may certainly have a functional role since phosphatases are present in the developing enamel matrix (Brookes et al., 1998). Phosphorylation *per se* can affect the secondary structure in LRAP in terms of its ability to stabilise the mineral phase (Yamazaki et al., 2017). As for P148, De-phosphorylation (or the lack of phosphorylation in rP147) affects the abilty of P148 to control crystal growth *in vitro* (Kwak et al., 2009).

1.2.3.2 Amelogenin degradation during maturation stage

As mentioned in the Section 1.1.2.4 (pp. 11-12), the increase in mineral content during the maturation stage is concomitant with the loss of the proteins in the matrix. The EMPs are degraded by the serine protease KLK4. KLK4 degrades the spectrum of proteins and peptides generated during the secretory stage processing (Hu et al., 2000, Ryu et al., 2002). KLK4 proteolysis generates small peptides that can be easily

removed from the matrix. These peptides are endocytosed by ruffle ended ameloblasts (Bartlett and Simmer, 2014, Pham et al., 2017). Fluid (enamel fluid) replaces the degraded enamel proteins. This fluid is kept supersaturated with respect to hydroxyapatite as the ameloblasts pump mineral ions into the enamel which drives the concomitant growth in thickness and width of the enamel crystals until the enamel becomes fully mineralised.

Using cervical enamel from incisor germs from human fetus (6-8 months old) and from erupted “temporary” teeth, Daculsi and Kerebel (1978) analysed the growth of the enamel crystallites in width and thickness using high resolution TEM. They showed that the crystallites grow mostly in width during the first stage of enamel development, and that until the end of foetal development the width has nearly stabilised to its average value in mature (erupted?) enamel. By then and until complete maturation, the crystallites grow only in thickness (Daculsi and Kerebel, 1978) as illustrated in Figure 10.

As enamel hydroxyapatite crystals grow during the maturation stage, the pH drops since the formation of calcium hydroxyapatite unit releases protons, as shown in equation (1) (Simmer and Fincham, 1995).

$$(1)\ 10\ Ca^{2+} + 6\ HPO_4^{2-} + 2\ H_2O \leftrightharpoons Ca_{10}(PO_4)_6(OH)_2 + 8H^+$$

To buffer the increase of acidity, ruffle-ended ameloblasts secrete bicarbonate ions towards the matrix. The pH of the enamel matrix oscillates between 6.1 and 7.4, as ameloblasts shift between ruffle-ended, which is when they secrete bicarbonate ions to neutralise the protons released, and smooth-ended. This possibly buffers the variations of pH caused by hydroxyapatite formation.

DLS studies showed that amelogenin nanospheres were extremely sensitive to changes of pH (addressed in Section 1.2.2.2.4, pp. 24-25), which may have a functional importance during amelogenesis. While the size of the aggregates were 35 nm at pH=7.4 they increased to 58 nm at pH=8, that is, an increase of 66% over 0.6 pH units (Moradian-Oldak et al., 1998). During the maturation stage, the drop and variations of pH may help break up the amelogenin nanospheres and thereby, facilitate amelogenin degradation and removal.

1.2.4 Other structural proteins of the developing enamel matrix : ameloblastin and enamelin

Ameloblasts secrete three major EMPs during the secretory stage (enamelin, amelogenin and ameloblastin) and two during the maturation stage (amelotin and ODAM, detailed earlier in Section 1.1.2.4, pp. 10 - 11). Evolutionary analyses indicated that the five genes encoding these proteins are related; all are derived from the enamelin encoding gene *ENAM* (Sire et al., 2007), which highlighted their importance in amelogenesis, and the potential necessity of a coordinated action of their expression products. This Section focuses on ameloblastin and enamelin, which are structural proteins specific to the secretory stage. Like amelogenin, they have a relatively high content of proline and glutamine (Smith et al., 2009). Their properties and hypothesised functions are addressed below.

1.2.4.1 Ameloblastin

Ameloblastin is the second most abundant protein in the developing enamel matrix after amelogenin and is another essential component in enamel formation as *ambn*$^{-/-}$ (null) or overexpression mouse models exhibited odontogenic tumours or AI phenotypes respectively (Paine et al., 2003a, Fukumoto et al., 2004). Mutations in ameloblastin gene *AMBN* drove some cases of AI in man, as detailed later in Section 1.3.2.2 (p. 53 - 54).

The ameloblastin gene, in rats, comprises an open reading frame predicted to encode a protein of 422 amino acids, with a molecular weight of 45 kDa. The protein encoded has a high content of proline, glycine, leucine and alanine and a predicted pI of 5.54 (Černý et al., 1996, Krebsbach et al., 1996). Ameloblastin is first expressed during amelogenesis as the IEE differentiates into ameloblasts, through the secretory stage and also during early maturation and its expression is much reduced at the end of maturation stage (Lee et al., 1996). The ameloblastin gene is alternatively spliced (Lee et al., 2003); and two major alternatively spliced transcripts have been reported in human, mouse, rat and pig (MacDougall et al., 2000).

As a secreted protein, ameloblastin was detected by immunohistochemistry transiting the ER, Golgi apparatus and secretory granules within ameloblasts prior to secretion (Uchida et al., 1997). Within the secretory pathway, the newly synthesised ameloblastin undergoes post-translational modifications, O-glycosylation and phosphorylation, which increase its apparent molecular weight on secretion (Uchida et al., 1997). Newly secreted ameloblastin is found mostly at the surface of the ameloblast Tomes' process during the secretory stage (Krebsbach et al., 1996). Once secreted, ameloblastin undergoes processing and degradation, with *in vitro* studies

suggesting that MMP-20 is the protease responsible (Iwata et al., 2007). Sequential extraction studies showed that some of the degradation products interact with enamel mineral and need to be desorbed from enamel crystal surfaces with phosphate buffer in order to be extracted from the tissue (Brookes et al., 2001). Initial cleavages of newly secreted ameloblastin generate a number of N-terminal polypeptides, some of which are found in the prism sheath. The remaining C-terminal polypeptides are quickly degraded and lost after secretion (Uchida et al., 1997). Although ameloblastin has a different spatial distribution within the extracellular matrix than amelogenin (i.e. N-terminal fragments accumulate at prism peripheries) they both share the same secretory pathway in the ameloblasts and are co-secreted (Nanci et al., 1998, Zalzal et al., 2008)

Although the precise function of ameloblastin is not clear, it is assumed to be involved in the control of crystal growth and in ameloblast attachment to the extracellular matrix. The major ameloblastin isoform carries in its sequence the amino acid motifs DGEA and VTKG; sequences found in adhesion molecules (Yamada and Kleinman, 1992). Studies using *Ambn*$^{-/-}$ (null) mouse models highlighted the importance of ameloblastin during ameloblast differentiation and maintaining of the ameloblast monolayer as the absence of ameloblastin resulted in detachment of ameloblasts from the matrix, loss of cell polarity and uncontrolled proliferation. The *Ambn*$^{-/-}$ (null) mouse models developed odontogenic tumours and dysplastic mineralised material was formed at the dentine surface (Fukumoto et al., 2004). A later study found that this so-called null mouse model actually expressed a truncated from of ameloblastin, *Ambn*$^{-5,6/-5,6}$ which lacks exons 5 and 6 (Wazen et al., 2009). This highlighted the importance of exons 5 and 6 as functional domains essential in ameloblast adhesion, control of ameloblast polarity and proliferation as described above (Fukumoto et al., 2004).

In common with other EMPs, ameloblastin has been a topic of interest in tissue regeneration research due to its cell signalling and growth factor-like properties (Zeichner‐David et al., 2006, Lu et al., 2013). It was shown to promote the proliferation of mesenchymal stem cells and osteoclast and their differentiation (Tamburstuen et al., 2010) and showed poential in preventing and healing bone fracture (Lu et al., 2016).

1.2.4.2 Enamelin

Enamelin accounts for 3-5% of EMPs . Its importance in amelogenesis was shown by enamelin-null mouse models where the enamel matrix failed to mineralise, leading to the lack of enamel in erupted teeth, enamelin-null mice did not display any enamel. The loss of enamelin results in the most severe of enamel phenotypes, suggesting

enamelin has a primordial role in enamel matrix mineralisation (Hu et al., 2008, Smith et al., 2009).

Immunofluorescence studies showed that enamelin is mostly located at the secretion front and at the dentine-enamel junction (Brookes et al., 2017b). Enamelin is expressed and secreted before amelogenin at the early secretory stage, though amelogenin and enamelin co-localise later at the “secretory face of the ameloblasts” (Gallon et al., 2013). Full-length enamelin has been identified as a 186 kDa protein that is glycosylated and phosphorylated (Hu et al., 1997). It is also degraded throughout the secretory stage of amelogenesis, and its processing products were shown to have a high affinity for mineral and crystal surfaces (Brookes et al., 2002). This explains why it has been suggested to influence crystal growth and enamel mineralisation. The major processing (32 kDa) product, which makes up 1% of the developing enamel matrix, is spread throughout the enamel layer (Uchida et al., 1991b). How can such a small amount be sufficient to cover the enamel surface?

In vitro studies suggested that enamelin and amelogenin act cooperatively in the control of crystal growth: CD and DLS analyses in PBS at pH 6.5 found direct interaction between recombinant rP148[1] and 32 kDa enamelin (Fan et al., 2011) and the addition of enamelin into an amelogenin gel-like matrix increased the stabilisation of an amorphous calcium phosphate transient phase (Iijima et al., 2010).

It is clear that the enamel matrix structural proteins have the potential to work cooperatively, within the matrix, to deliver an ordered mineralisation process. Their secretion by the ameloblasts also appears to be coordinated; e.g. amelogenin and ameloblastin were shown to be co-transported (Zalzal et al., 2008, Mazumder et al., 2014). The secretion process of EMPs is addressed in greater details below.

[1] The recombinant rP148 differs from the native P148 since it lacks Met^{1} and Ser^{16} phosphorylation, and carries and extra methionine at position 149.

1.2.5 Amelogenin transit through the ameloblast ER for secretion

Amelogenin is trafficked intracellularly towards the ameloblast Tomes' processes for secretion into the extracellular matrix (Simmelink, 1982, Nanci et al., 1985). Its propensity to aggregate requires a tight monitoring within the ameloblast ER, which is the organelle responsible for overseeing protein folding and the assembly of subunits that comprise multimeric proteins. As a secreted protein, amelogenin is synthesised by ribosomes associated with the rough ER and simultaneously translocated into the ER lumen where, like all proteins, presumably adopts an energetically favourable conformational state assisted by the ER folding machinery (chaperone proteins that interact with the client protein to prevent pathological aggregation in order to direct protein folding towards the correct functional end point). It then transits through the Golgi and by vesicular transport reaches the Tomes' process, where it is released into the extracellular matrix. Ameloblasts are cells specialised in secretion and are adapted to manage large amounts of EMPs destined for secretion. The volume of these secretory organelles is maximal during the secretory stage to maintain ameloblast ER proteostasis (Warshawsky, 1968, Tsuchiya et al., 2008). This Section (i) describes the secretory pathway in general (Section 1.2.5.1) and (ii) focuses on intracellular transit of amelogenin during secretion (Section 1.2.5.2).

1.2.5.1 Ameloblasts: cells specialised in protein secretion require an efficient ER trafficking machinery

The ER handles secretory proteins or transmembrane proteins, starting from their synthesis to their transfer to the Golgi apparatus. The ER handles both protein synthesis and folding so that that proteins achieve their functional 3-dimensional conformation. This is important as the ER environment is crowded, with concentrations of proteins reaching 300-400 g/L (Ellis and Hartl, 1999) and up to 30% of WT proteins destined for secretion can spontaneously mis-fold (Schubert et al., 2000). Misfolded proteins may result in in loss of biological activity inactive, or be cytotoxic, or may induce pathological intracellular protein aggregation.

As discussed in more detail in Section 1.3.4.2 (pp. 75 - 77), a first-line mechanism to relieve potentially pathological situations in the ER employs folding machinery comprised of so-called chaperone proteins which interact with client proteins transiting the ER and directing their folding towards the correct conformational endpoint whilst preventing abnormal aggregation. In addition, the ER is able to direct misfolded or aggregated proteins towards degradation which helps restore ER proteostasis. When neither of these factors is sufficient to restore proteostasis, the ER becomes

"stressed". ER stress triggers numerous signalling cascades which comprise the so-called unfolded protein response (UPR) which triggers a number of downstream actions designed to relieve the stress or ultimately commit the cell to apoptosis.

- **ER folding machinery**

The ER folding machinery comprises a variety of protein chaperones, co-chaperones and folding enzymes (foldases) that assist protein folding during synthesis to prevent protein aggregation (Dobson, 2003). This process is called ER-assisted folding (ERAF) and a list of its components is detailed in Table 1. The chaperones recognise the hydrophobic portions exposed at the surface of immature, misfolded or aggregation-prone proteins and then assist their folding (Braakman and Hebert, 2013). One of the most abundant and well-known chaperones, binding protein (BiP)/GRP78, which belongs to the heat shock protein 70 (HSP70) family, actively promotes folding through conformational changes driven by ATPase activity, resulting in ATP hydrolysis (Hartl, 1996). It is noteworthy that the chaperone "Binding immunoglobulin protein" (BiP) has been shown to interact with amelogenin *in vitro* (Fukuda et al., 2013) indicating that amelogenin is likely to be chaperoned during its transit through the ER. Associated co-chaperones include proteins from HSP40 and GrpE families, which modulate the ATPase activity of BiP (Yoshida, 2007). Other chaperones include GRP94 (HSP90 family) which has ATPase activity, and the chaperones calnexin and calreticulin. The latter two are lectins that bind transiently to N-linked glycosylated proteins and act together to help folding of the target proteins in what is called the calnexin-calreticulin cycle (Ou et al., 1993, Helenius and Aebi, 2001). ER folding enzymes, which change the energy landscape and increase the rate of protein folding, include protein disulfide isomerases (such as GRP58/ERp57), which catalyse oxidoreduction reactions (involving disulfide bonds), and peptidylprolyl isomerases (including FKBP family), which catalyse the conversion *cis-trans* of peptide bonds involving proline (Lang et al., 1987, Harding et al., 1989). A remarkable feature is that peptide bonds involving proline adopt either *cis* or *trans* configuration unlike regular peptide bonds (without proline), which spontaneously adopt the *trans* configuration. With proline peptide bonds, *cis-trans* isomerisation reactions are needed to reach the protein's native state. This requires a high activation energy (~20 kcal/mol) (Andreotti, 2003) which curbs the kinetics of protein folding. To overcome that limiting step, peptidylprolyl isomerases in the ER catalyse the *cis-trans* isomerization reaction, preserving the rate of protein folding and transit (Andreotti, 2003). Their role is presumably all the more important when the proteins destined for secretion, such as amelogenin, have an high proline content.

- **ER associated degradation**

Once properly folded, the protein is ready for use or secretion out of the cell. However, in some cases, the protein can evade the ERAF pathway, due to a mutation for example, and misfold. The misfolded protein is detected and then processed for destruction, via ER-assisted degradation (ERAD). The misfolded protein is translocated towards the cytosol, through specific molecular recognition / interaction mechanisms, where it is ubiquitinated and degraded in proteosomes (Yoshida, 2007). ERAD thus helps to reduce the load of misfolded proteins; the ERAD components involved are listed in Table 1. When ERAD or ERAF are not sufficient to restore homeostasis, a third mechanism involving signalling cascades is activated as a last resort: the UPR.

- **The unfolded protein response (UPR)**

The UPR consists of intracellular signalling cascades that control the expression of numerous genes in the cell. Misfolded proteins in the ER lumen activate trans-ER membrane molecular sensors. In higher eukaryotes such as mammals, the UPR comprises three signalling cascades, each of which is initiated upon the activation of an associated sensor protein: protein kinase RNA-like ER kinase (PERK), inositol-requiring enzyme 1α (IRE1α), and activating transcription factor 6 (ATF6). In the resting state, the three sensors are kept inactive through their interaction with the ER chaperone BiP. However, BiP preferentially binds misfolded proteins, so in the presence of misfolded proteins, BiP tends to disassociate from the sensor proteins and the sensor proteins become activated (Bertolotti et al., 2000, Shen et al., 2002) triggering downstream signalling cascades. The UPR aids cells to manage large secretory loads by increasing ER volume (via synthesis of membrane lipids), ERAF and ERAD machinery (Travers et al., 2000).

Figure 11 summarises the transit of secreted proteins such as amelogenin through the ER. The ER quality control machinery is listed in Table 1.

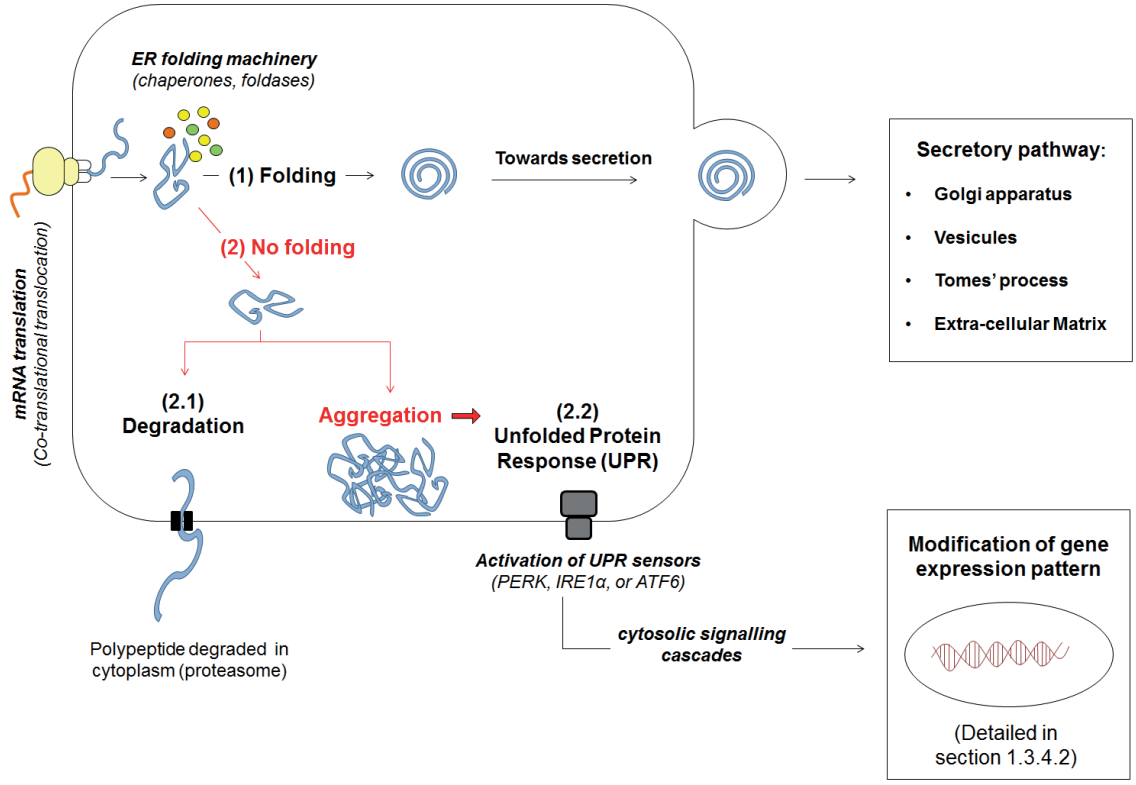


Figure 11 Diagram illustrating the transit of secreted proteins (e.g. amelogenin) through the ER for secretion. The ER handles synthesis and folding of proteins destined to secretion (or transmembrane proteins, not addressed here). Following translocation[1], the proteins fold (1) to adopt the correct conformation, which may require the ER folding machinery (which includes chaperones, co-chaperones and folding enzymes). When the folding machinery fails ('(2) No folding'), the proteins may be targeted for degradation (case '(2.1)') to alleviate the load of misfolded proteins in the ER to limit ER stress. In the case where the ER folding and ER-degradation do not suffice to alleviate ER stress, transmembrane sensors (PERK, IRE1α and ATF6) are activated (case '(2.2)), triggering signalling cascades which impact on gene expression. This is the so-called UPR which attempts in first instance to increase the folding capacity of ER or alleviate the load of misfolded proteins to reduce ER stress. If the ER fails to return to proteostasis then the UPR triggers apoptosis. The balance between the UPR acting in pro-survival mode and pro-apoptotic mode is detailed later in Section 1.3.4.2.

[1] Co-translational translocation is the mechanism used to target proteins into the ER in eukaryotes.

Table 1 ER quality control machinery: list of components of ERAF machinery, the ERAD machinery and UPR sensors. The information presented in this table is compiled from reviews by Hebert and Molinari (2007), Ellgaard and Ruddock (2005) and Ruggiano et al (2014). The references are listed in the footnote below[1].

ERAF	○ **Chaperones** • *Hsp70 family* : BiP/GRP78 • *Hsp90 family*: GRP94/endoplasmin/ERp99 • *Lectins*: Calnexin and calreticulin, EDEM 1 - 3 ○ **Co-chaperones** • *Hsp40 family* : Erdj 1 - 5 • *GrpE-like*: BAP, GRP170 ○ **Folding enzymes** • *Protein disulfide isomerases*: PDI, PDIp, PDIr, PDILT, P5, ERp18, ERp28, ERp44, ERp46, ERp57/GRP58, ERp72, ERdj5, TMX, TMX2 • *Peptidylprolyl isomerases*: CyclophilinB/CypB, FKBP family: FKBP2/FKBP13, FKBP7/FKBP23, FKBP10/FKBP65
ERAD	○ **Substrate recognition:** SEL1, Bip, Derlin-1-3, OS9, XTP3-B ○ **Retrotranslocation:** HRD1, pg78, HERP, Derlin-1-3 ○ **ER ligase:** HRD1, pg78
UPR	○ **PERK pathway** (detailed in Section 1.3.4.2, related to conformational diseases) ○ **ATF6 pathway** (detailed in Section 1.3.4.2, related to conformational diseases) ○ **IRE1α pathway** (detailed in Section 1.3.4.2, related to conformational diseases)

[1] **References**

(1) HEBERT, D. N. & MOLINARI, M. 2007. In and out of the ER: protein folding, quality control, degradation, and related human diseases. *Phys. rev.*, 87, 1377-1408.

(2) ELLGAARD, L. & RUDDOCK, L. W. 2005. The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Rep*, 6, 28-32.

(3) RUGGIANO, A., FORESTI, O. & CARVALHO, P. 2014. ER-associated degradation: Protein quality control and beyond. *Journal of Cell Biology*, 204, 869-879.

1.2.5.2 Protein-protein interactions associated with amelogenin as it transits the ameloblast secretory pathway

To ensure amelogenin's proper transit through the ER and the rest of the secretory pathway, amelogenin needs to adopt a conformation that allows correct assembly and interactions with itself and with other proteins (or put another way, needs to avoid accessing energetically stable conformations that may favour aggregation or other pathological outcomes). This Section focuses on the protein-protein interactions involved during intracellular transit of amelogenin. It addresses: (i) potential binding partners (Section 1.2.5.2.1 below) and (ii) the domain(s) of amelogenin possibly involved in protein-protein interactions (Section 1.2.5.2.2 below).

1.2.5.2.1 Importance of binding partners

Various studies have identified multiple potential binding partners for amelogenin. These include amelogenin itself (self-assembly, detailed previously in Section 1.2.2.2, p. 21), cellular components (such as those involved in the ER-quality control machinery), or specific EMPs.

- **Amelogenin self-assembly**

As detailed previously (Section 1.2.2), amelogenin has a well-known propensity to self-assemble, with an *in vivo* study (discussed in Section 1.2.2.2.2, p. 22) demonstrating the formation of dimers, tetramers, pentamers and hexamers in the ameloblast intracellular secretory pathway (Brookes et al., 2006). However, this is predated by multiple studies reporting the aggregative nature of amelogenin, which has been documented since the 1960s (Nikiforuk and Simmons, 1965, Mechanic et al., 1967).

- **Binding of amelogenin with elements from the ER-folding machinery**

Potential binding partners for amelogenin (other than amelogenin itself) including protein components of ameloblasts and the enamel matrix have been investigated. A proteomics study using affinity chromatography and immunofluorescence microscopy in an SaOS-2 osteoblastic cell line found that HSP70 family proteins (especially BiP/Grp78) and other proteins of the ER folding machinery such as calreticulin, protein disulphide isomerase precursor and tapasin-ERp57 bound amelogenin (Fukuda et al., 2013). Fukuda et al. (2013)'s work was in fact focused on the mechanisms by which amelogenin (in the form of Emdogain) stimulated periodontal regeneration rather anything relating to amelogenin secretion by ameloblasts *per se*. They were actually studying the uptake of Emdogain by osteoblastic cells and identified amelogenin binding partners in the course of this work. Nevertheless, their work provides data that may be relevant to amelogenin secretion

by ameloblasts. Within the same study, other cellular amelogenin-binding proteins identified included cytoskeleton proteins such as actin, vimentin, tubulin. Another group used the Y2H assay, screening a 17-day mouse embryo pretransformed with a Y2H expression library, to identify EMP binding partners. They identified fetuin-A, biglycan and CDC63 as binding partners for amelogenin and also for ameloblastin or enamelin (Wang et al., 2005).

- **Other EMPs destined for secretion**

As mentioned above, the EMPs (amelogenin, ameloblastin, enamelin) were found to have mutual (potential) binding partners such as CDC63, fetuin-A or biglycan (Wang et al., 2005). Whether this is of importance in amelogenesis remains to be clarified. Nonetheless, it is well known that their expression and secretion need to be coordinated and their actions cooperative for amelogenesis (See Section 1.2.4 pp. 37-39). An association between amelogenin and ameloblastin has been put forward, as both are expressed and secreted at the same time, and share the same secretory pathway within the ameloblasts. They were found co-localised in secretory vesicles (Nanci et al., 1998, Zalzal et al., 2008). They were reported to interact via the amelogenin ATMP domain by enzyme-linked immunosorbent assay (ELISA), dosimetry and Scatchard analyses (Ravindranath et al., 2004). Immunofluorescence, CD and fluorescence spectroscopy studies also found that both proteins interact with the interaction being between the N-terminus of amelogenin (TRAP) and the ameloblastin domain encoded by exon 5 (Mazumder et al., 2014, Su et al., 2016). Currently, while amelogenin-ameloblastin interactions have been suggested *in vitro*, no *in vivo* evidence has been published to date.

In contrast, Y2H assays indicated that amelogenin and ameloblastin do not interact (Paine et al., 1998, Bartlett et al., 2006b) .The Y2H system is a popular method to study protein-protein interactions in living yeast cells and has been useful to define the domains involved in amelogenin-amelogenin interaction (see appendix E). It verifies the interaction between two recombinant proteins in a living yeast cells, which activates the expression of a reporter gene should they interact (Fields and Song, 1989). The system is considered an *in vivo* assay (Brückner et al., 2009) but both recombinant proteins involved are fused to Gal4 binding or activating domain which may alter the folding, function or the binding behaviour of the two potential binding partners. In addition to this, protein interaction detected using the Y2H system take place in the yeast nucleus. This is a very different environment to the mammalian rough ER and will be missing ancillary chaperones or part of the folding control machinery, usually resident in the mammalian rough ER. Furthermore, proteins in the Y2H system will not undergo post-translational modification such as phosphorylation. These caveats must be considered when interpreting Y2H system data.

1.2.5.2.2 Importance of the amelogenin tri-tyrosyl motif peptide (ATMP) in mediating amelogenin-protein interactions

The integrity of amelogenin (in terms of its primary sequence) is necessary to allow its folding (or lack thereof), interaction with ancillary chaperones and ultimately its safe transit through the ameloblast secretory pathway. As amelogenin self-assembly has been the most studied protein-protein interaction involving amelogenin, the binding domains responsible for its self-assembly are the best understood.

As detailed previously in Section 1.2.2 amelogenin self-assembly binding domains were identified in N- (A-domain) and C- (B-domain) terminals by Y2H experiments (Paine and Snead, 1997). Within the A domain, the sequence of amino acids Tyr^{34}-Met^{42} , that is: YPSYGYEPM (ATMP domain) was shown to be crucial in driving amelogenin-amelogenin interactions. This domain has also been the focus of reports by Ravindranath et al (1999) who carried out extensive work suggesting that the domain interacted with sugars, glycoproteins, cytokeratins, and also possibly ameloblastin.

These authors reported that the ATMP domain potentially binds glycoproteins by using hemagglutination and hemagglutination inhibition (hemagglut.) tests with sugar, and dosimetry with [14C] N-acetyl-D-glucosamine (GlcNAc). This first report indicated that the amelogenin ATMP domain can bind to GlcNAc (Ravindranath et al., 1999). Then, based on the sequence similarities that they observed between the amelogenin ATMP sequence and wheat germ agglutinin, Ravindranath et al (2000) tested whether amelogenin could bind to GlcNAc-mimicking peptides whose sequence is present in cytokeratins. They found that ATMP was capable of binding to the GlcNAc-mimicking peptide “SFGSGFGGGY” (Ravindranath et al., 2000), a sequence known to be present in cytokeratin 14 (Shikhman et al., 1994).

In a following report, a dosimetric binding assay verified the affinity of cytokeratin 14 for rM179 and for the ATMP sequence and in the same report the authors investigated and corroborated its relevance *in vivo*. Looking at mouse incisors by confocal microscopy, they reported that amelogenin and cytokeratin 14 co-assemble “in the perinuclear region of ameloblasts” at day 0 post-natal, then between day 1 and day 3-5, the complex as observed to migrate towards the ameloblast’s apical region and then disassemble. Dissociation of cytokeratin 14 from amelogenin was observed in the Tomes’ processes by autoradiography with [^{3}H]ATMP and [^{3}H]GlcNAc-mimicking peptide (Ravindranath et al., 2001) which led them to hypothesise that cytokeratin 14 may act as a chaperone for amelogenin during its intracellular transport. In 2003, the same group also identified cytokeratin 5 as a binding partner for amelogenin (Ravindranath et al., 2003). Cytokeratin 5 is a protein that is post translationally

modified with GlcNAc and can pair with cytokeratin 14 to form the intermediate filaments of basal epithelial cells including ameloblasts (Kasper et al., 1989) The role of ATMP binding to cytokeratin 5 was confirmed by ELISA, western blotting and confical laser scanning microscopy on post-natal ameloblasts (Ravindranath et al., 2003). Ravindranath et al. (2003) observed a cytokeratin 5-amelogenin complex migrating towards the periphery of the ameloblasts. Towards the secretory ends of the ameloblasts, confocal microscopy showed amelogenin, cytokeratin 14 and cytokeratin 5 co-localised at day 5 of postnatal growth of enamel. At day 9, the complex appeared dissociated as cytokeratin 14 was no longer found co-localised with either cytokeratin 5 or amelogenin. Cytokeratin 5 may be found either co-localised or dissociated from amelogenin. Remarkably, Ravindranath et al. (2003) raised a question/issue regarding the function of the ATMP domain by suggesting that it may be involved in GlcNAc binding, rather than amelogenin self-assembly, previously reported by Paine and Snead (1997). They opposed the hypothesis of Paine and Snead that the ATMP domain drives amelogenin self-assembly. This raises two questions: (1) Does the ATMP domain drive amelogenin self-assembly or is it the binding site for cytokeratin which appeared to partner amelogenin during its journey through the secretory pathway? And (2) Are these two functions mutually exclusive?

To complicate matters, Ravindranath et al. (2004) reported that recombinant ameloblastin bound to amelogenin and peptides corresponding to the ATMP sequence, despite the fact that their recombinant ameloblastin did not contain GlcNAc or the amino acid sequence that mimics GlcNAc (Ravindranath et al., 2004). This contrasts with the Y2H studies previously described that were unable to demonstrate any interaction between ameloblastin and amelogenin (Paine et al., 1998). However, as already stated, certain caveats need to be considered when interpreting Y2H data. Ravindranath et al (2004) explained the apparent paradox – how can recombinant ameloblastin bind the ATMP sequence when it contains neither GlcNAc nor the GlcNAc mimicking amino acid motif - by hypothesising that the GlcNAc mimicking motif may actually be formed by the relevant amino acids, lying far apart in the primary sequence, coming into the correct juxtaposition to form the GlcNAc mimicking motif once the ameloblastin molecule has folded. Paine et al (1998) may have failed to identify ameloblastin-amelogenin interactions in the Y2H system due to the specific conformation leading to the correct juxtaposition of amino acids to form the GlcNAc mimicking motif being disturbed by fusion of the ameloblastin sequence with Gal4 binding or activating domain on which the Y2H assay depends. Another problem with Y2H system is that there are glycosylated proteins produced by the yeast which may compete with ameloblastin by binding to the amelogenin ATMP.

While a number of studies appeared to suggest that the ATMP domain may mediate amelogenin-protein interactions, convincing *in vivo* evidence is lacking. On the basis that ATMP may drive protein recognition, it is likely to have an essential role in amelogenesis, and it would be of worth to determine how conserved this domain is between species. Wald et al (2017) identified the ATMP domain as a conserved 'Y/F-X-X-Y/L/F-X-Y/F' motif. This study, using HR-SEC, TEM and SPR and site-directed mutagenesis, showed the importance of the integrity of the ATMP domain in mediating amelogenin self-assembly and ameloblastin recognition (Wald et al., 2017). Indeed, mutations in the ATMP domain (for example p.P70T) are associated with AI in humans (Collier et al., 1997) and as discussed later, the p.Y64H mutation in mice appears to cause pathological intracellular aggregation of amelogenin in ameloblasts suggesting altered amelogenin-amelogenin binding characteristic (Barron et al., 2010, Brookes et al., 2014).

The assembly and cooperation between EMPs is clearly a requirement for enamel biomineralisation. This cooperation starts within the ER in the ameloblasts, and proper protein-protein interactions are necessary for transit throughout the secretory pathway of the ameloblasts. Failure of the ER secretory pathway has been shown to the underlying mechanism of enamel pathology in the mouse model carrying the p.Y64H amelogenin mutation.

1.3 Enamel pathologies and therapies to date

Enamel is an acellular tissue, as ameloblasts are lost upon tooth eruption; it cannot cannot undergo cell mediated self regeneration or repair (Moradian-Oldak, 2009, Jayasudha et al., 2014). This is problematic in a variety of conditions including treatment of caries lesions, accidental, environmental or hereditary defects as enamel damage cannot be reversed. These require either recurrent fillings, veneers or tooth replacement with dental implants (Jayasudha et al., 2014). Regeneration studies using peptides have been a recent subject of interest. Self-assembling anionic peptide P_{11}-4 was engineered and shown to promote repair of enamel under simulated intra-oral conditions (Kirkham et al., 2007) and proved to be a safe and effective treatment of early caries lesions in clinical trials (Brunton et al., 2013). Another study used repeated applications of amelogenin-derived peptides, resulting in the formation of aprismatic enamel-like tissue, however these results were only reported, so far, *in vitro* (Mukherjee et al., 2018).

Enamel developmental defects can result from environmental perturbations (physical shocks or chemical stresses) as described above, or be hereditary. Research has also increasingly focused on the role that genetics might play in the aetiology driving enamel conditions linked to 'environmental' factors such as caries, molar incisor hypomineralisation (MIH) or fluorosis (Kirkham et al., 2017).

1.3.1 Enamel pathologies and developmental defects

Dental caries, MIH and fluorosis are relatively common oral diseases usually linked to environmental factors. Growing evidence suggest that genetic factors and heredity may play a significant role in the development of the diseases .

Fluorosis is caused by exposure to excessive fluoride during tooth development. The enamel is hypomineralised and secretory-stage EMPs are abnormally retained in the maturation stage enamel, where they could hinder secondary mineralisation, as reviewed by Brookes et al (2017a) (Brookes et al., 2017a).

Dental caries is caused by acids arising from bacterial metabolism that dissolve tooth minerals. Affecting enamel initially but progressing into the dentine, or directly affecting dentine in the case of root caries, caries can lead to severe inflammation, tooth loss, infection or abscess if left untreated (Bagramian et al., 2009). Affecting 2.3 billion people (permanent teeth) and 560 million children (deciduous teeth) worldwide, it is the most prevalent noncommunicable disease worldwide (WHO, 2017). The cariogenic potential of bacterial biofilms present on teeth is sensitive to environmental influences;

mainly the availability and frequency of fermentable carbohydrate entering the oral cavity. However, it is increasingly clear that there is a genetic predisposition to caries. For example, specific genetic polymorphisms in amelogenin increased caries risk (Kang et al., 2011).

MIH is a condition characterised by hypomineralised enamel on the molar and incisor teeth, which causes tooth sensitivity to cold, warmth, and mechanical stimulation such as toothbrushing (Weerheijm et al., 2001). Its prevalence varies worldwide, from 2.4% in Bulgaria to 40.2% in Brazil (Vieira and Kup, 2016). MIH shows different degrees of severity from minor discoloration in the milder cases to enamel fracture in the most severe cases. Since the late 1970s, MIH has been regarded as an idiopathic disease (Koch et al., 1987). Its aetiology has been a topic of continuing interest. Although the underlying mechanisms are not entirely clear, research to date supports a combination of environmental and genetic factors. Environmental factors include exposure to bisphenols and dioxin, illnesses, and malnutrition and health problems during prenatal, perinatal and neonatal periods; however, reviewers have questioned the robustness of evidence for these factors (Crombie et al., 2009, Alaluusua, 2010). Genetic susceptibility was supported by a genome-wide association study (Kühnisch et al., 2014) and a twin study (Teixeira et al., 2018). Variants of genes involved directly in enamel development such as *AMBN*, *TUFT1* and *TFIP11* were associated with MIH, (Jeremias et al., 2013).

1.3.2 AI: major inherited pathologies of enamel

AI refers to the major group of inherited enamel pathologies that affect both deciduous and permanent dentitions. This implies that AI is non-syndromic, affecting only the teeth, whereas it can be both syndromic and non-syndromic. It is characterised by an altered appearance (hypomineralisation, missing enamel, enamel discolouration and roughness) and sensitivity of the teeth which affects the patients' quality of life. AI-related issues also include difficulties with oral hygiene maintenance and loss of self-esteem (Hu et al., 2007, Coffield et al., 2005). No preventative cure for AI exists to date and its prevalence ranges between 1:14000 in the U.S, to 1:700 in Northern Sweden (Crawford et al., 2007). Different AI classifications have been proposed but the usual criteria for classification include the clinical phenotype, the mode of inheritance, the genetic molecular basis (Aldred et al., 2003).

1.3.2.1 Phenotype and classification

Clinical diagnosis of AI has been largely based on clinical appearance, radiographic assessment and by interrogating the family history (Crawford et al., 2007) though more recently genetic diagnostics have been called for, e.g. targeted 21-gene panel test for AI introduced by the NHS in 2016 (McDowall et al., 2018). Previously however, AI has been classified according to the clinical phenotype exhibited by the patient.

Hypoplastic AI is a defect in the 'quantity' of enamel present, characterised by abnormally thin enamel, which in extreme cases may be absent altogether. It is usually attributed to defective enamel matrix formation during the secretory stage which in turn undermines elongation of the crystals so although the resulting enamel is mineralised, it is extremely thin (Gadhia et al., 2012). The secretory stage may be affected either by mutations in genes encoding structural EMPs such as enamelin, ameloblastin or amelogenin; or it can be affected by genes encoding master controllers of amelogenesis such as *DLX3* or *FAM20A* as reviewed by Smith et al (2017b) (Smith et al., 2017b).

Hypomineralisation is another type of AI, thought to be caused by the enamel matrix reaching normal thickness but failing to mineralise due to a defective maturation stage. The resulting enamel is soft, weak and discoloured. Hypomineralised AI can be divided into two categories: Hypomaturation, which is thought to result from a failure to remove the EMPs; mutations in human and mice KOs of genes encoding enamel matrix proteases MMP-20 and KLK4 were associated with hypomature phenotype (Caterina et al., 2002, Wright et al., 2009, Simmer et al., 2009, Yamakoshi et al., 2011, Hu et al., 2016, Núñez et al., 2016). The second hypomineralisation category is hypocalcified AI, which is where the mineralisation has failed, and is characterised by

a soft enamel (Hu et al., 2007). In both types of hypomineralised AI, enamel may be lost post eruption (Smith et al., 2017b).

1.3.2.2 Genes known to underlie AI

Usually, AI is considered as non-syndromic inherited enamel defects. Actually, AI is a spectrum of inheritied enamel pathologies ranging from the non-syndromic, where enamel only is affected, to syndromic cases, where non-dental tissues may be affected. AI is a heterogeneous group of conditions with at least 18 genes involved identified (Smith et al., 2017b). The patterns of inheritance vary with examples typical of autosomal dominant, autosomal recessive (Wright et al., 2015) and X-linked (Lagerstrom et al., 1991) inheritance patterns.

Smith et al. (2017b) provided a comprehensive review of AI and associated mutations (Smith et al., 2017b). Genes associated with AI include for example those encoding EMPs, for instance *AMELX* (Lagerstrom et al., 1991), *ENAM* (Rajpar et al., 2001, Mardh et al., 2002) or *AMBN* (Poulter et al., 2014). Mutations in genes encoding enamel matrix proteases such as MMP2*0* and KLK4 were shown to cause AI (Kim et al., 2005, Hart et al., 2004). Genes involved in AI also encode proteins involved in cell-cell and cell-matrix adhesion such as amelotin (Smith et al., 2016), and FAM83H (Kim et al., 2008); proteins involved in transport such as WDR72 (El-Sayed et al., 2009) and SLC24A4 (Parry et al., 2013); proteins involved in pH sensing, e.g.GPR68 (Parry et al., 2016) and genes tentatively associated with hydroxyapatite nucleation, e.g. *C4orf26* (Parry et al., 2012). Other genes expressed in ameloblasts, such as *ACPT*, were shown to be associated with AI upon mutation, yet the role of the encoded proteins in amelogenesis is unclear (Smith et al., 2017b). The mutations associated with AI cited in this Section are listed below in Table 2. There have been various point mutations, deletions or insertions - possibly causing frame shifts identified to date. Even silent mutations that affect the DNA but not the amino acid sequence of an encoded protein have been shown to be associated with AI, for instance the inclusion of exon 4 in the mature transcript (for the major amelogenin isoform) lead to hypoplastic and hypomineralised enamel (Cho et al., 2014).

For some genes, specific mutations lead to syndromic conditions while other mutations in the same gene lead to enamel defects only, which blurs the boundaries between the non-syndromic and syndromic enamel defects. Genes involved in AI, such as *ACPT*, *FAM20A* and *DLX3* are expressed in ameloblasts and in other tissues. In the case of autosomal recessive mutation of *FAM20A*, hypoplastic AI can be spotted as an early symptom of enamel renal syndrome (Jaureguiberry et al., 2012, Wang et al., 2013, de la Dure-Molla et al., 2014) and thereby provide a means of early diagnosis since renal problems only manifest later in life.

Table 2 Mutations associated with non-syndromic AI: function of the WT gene expression product, inheritance, amelogenesis stage possibly affected and phenotype reported.

***Gene*, locus**	**Protein** (Function if known)	**Patterns of inheritance**	**Amelogenesis stage**	**Phenotype**	**Publications (AI reports)**
AMELX **Xp22.1-p22.3**	**Amelogenin** EMP	X-linked	Secretory	Hypoplastic Hypomature	(Lagerstrom et al., 1991)
AMBN **4q13.3**	**Ameloblastin** EMP	AR	Secretory	Hypoplastic Thin and aprismatic enamel observed	(Poulter et al., 2014) (Prasad et al., 2016)
ENAM **4q13.3**	**Enamelin** EMP	AD AR	Early secretory	Hypoplastic Mild to severe, localized enamel pitting to generalised hypoplasia	(Rajpar et al., 2001) (Ozdemir et al., 2005)
MMP20 **11q22.2**	**MMP-20** Enamel matrix protease	AR	Secretory	hypomaturation	(Kim et al., 2005)
KLK4 **19q13.4**	**KLK4**, Enamel matrix protease Protein degradation	AR	Maturation	hypomaturation	(Hart et al., 2004)
DLX3 **17q21-q22**	**DLX** transcription factor, Master controller of amelogenesis	AD	Secretory (major) Pre-secretory (weak)	hypoplastic-hypomaturation + taurodontism	(Dong et al., 2005)
WDR72 **15q21.3**	Intracellular vesicle coating (predicted)	AR	Maturation (major) Secretion (weaker)	Hypomaturation Hypodontia, delayed tooth eruption	(El-Sayed et al., 2009)

SLC24A4 **14q32.12**	**Sodium/calcium exchanger** Active transport of Ca^{2+} into enamel matrix	AR	maturation	Hypomaturation, Hypomineralisation	(Parry et al., 2013)
C4orf26 **4q21.1**	Acidic phosphoprotein (predicted) HA crystallisation	AR	Secretory and maturation	Hypomineralisation	(Parry et al., 2012)
FAM83H **8q24.3**	Regulates keratin cytoskeleton, desmosomes Vesicle trafficking?	AD	Ubiquitous Pre-secretory and secretory (major)	Hypocalcified	(Kim et al., 2008)
FAM20A **17q24.2**	**FAM20A** Pseudokinase, Master controller of amelogenesis	AR	Secretory stage	Hypoplastic	(Jaureguiberry et al., 2012) (Wang et al., 2013) (de la Dure-Molla et al., 2014)
ACPT **19q13.33**	**Acid phosphatase, testicular** Orthophosphoric esters hydrolysis	AR	Secretory stage	Hypoplastic	(Seymen et al., 2016) (Smith et al., 2017a)
AMTN **4q13.3**	**Amelotin** Cell-matrix attachement Hydroxyapatite mineralisation	AD	Transition and maturation stage	Hypomineralisation	(Smith et al., 2016)
GPR68 **14q32.11**	**GPR68** pH sensing	AR	All stages of amelogenesis	Hypomineralisation	(Parry et al., 2016)

1.3.3 X-linked AI: Amelogenin mutations and AI

Given that the focus of this thesis is around amelogenin and in particular the p.Y64H amelogenin mutation a detailed account of AI linked to amelogenin mutation is provided below.

1.3.3.1 X-linked AI: Amelogenin mutations and AI

The gene encoding amelogenin *AMELX* is located on the short arm of the X chromosome in the region p22.1-p22.3 (Lau et al., 1989). X-linked AI accounts for 5% of all AI cases. The phenotypes involve various degrees of hypoplasia and/or hypomaturation (Hart et al., 2002, Wright et al., 2003). Within a single family, affected heterozygous females showed a milder phenotype than males or homozygous females. In heterozygous females, the phenotype is characterised by alternatinging stripes of affected and normal enamel. This is caused by “Lyonisation” which is the process by which one of the two X chromosomes present in every female cell is randomly deactivated. This leads to the differentiation of cohorts of ameloblasts that are either expressing a copy of the defective amelogenin gene or an unaffected WT copy (Berkman and Singer, 1971).

To date, over 16 mutations causing X-linked human AI have been reported. The known amelogenin mutations are detailed in Figure 12. Three single point mutations were found in the sequence encoding the signal peptide resulting in a lack of amelogenin secretion and hypoplastic AI. Eight mutations were found in the N-terminal and central domains. One deletion causes truncation of the amelogenin gene to 18 amino acid codons (Lagerstrom et al., 1991, Lagerstrom-Fermer et al., 1993) and the resulting phenotype comprised hypomineralised enamel with various levels of hypoplasia. In another case, a mutation affected amelogenin mRNA splicing so that the major amelogenin mRNA transcript included exon 4 (exon 4 is normally spliced out). This mutation was associated with enamel hypomaturation and to a lesser extent enamel hypoplasia in affected patients and hypomineralised AI in a transgenic mouse model (Cho et al., 2014).

Although genotyping has been an important tool in understanding the causes of AI, there is a need to better understand genotype-phenotype links. Clinical studies currently consist of recruiting AI patients, evaluating their medical, dental and family histories, and establishing their clincial enamel phenotype following examination by principal investigators (Hart et al., 2002). To date, the understanding of underlying mechanisms is poor (Wright et al., 2015) apart from genotype-phenotype correlations. Mutations in the signal peptide and toward the C terminus (green and blue portions in Figure 12B) are associated with hypoplastic AI while mutations in N-terminal and

central domains (yellow portion in Figure 12B) cause hypomaturation AI with variable hypoplasia (Hart et al., 2002, Smith et al., 2017b, Kim et al., 2017) studies focusing on the aetiological mechanism involved (e.g. the impact of a mutation on protein folding and assembly, protein to mineral binding, intra- and extra-cellular transit and matrix function) have been carried out (described overleaf in Section 1.3.3.2). They were intended to provide not only a better understanding of the normal function of the protein in question but also open the possibility of developing therapeutic interventions to overcome the deleterious effect of the mutation during enamel development.

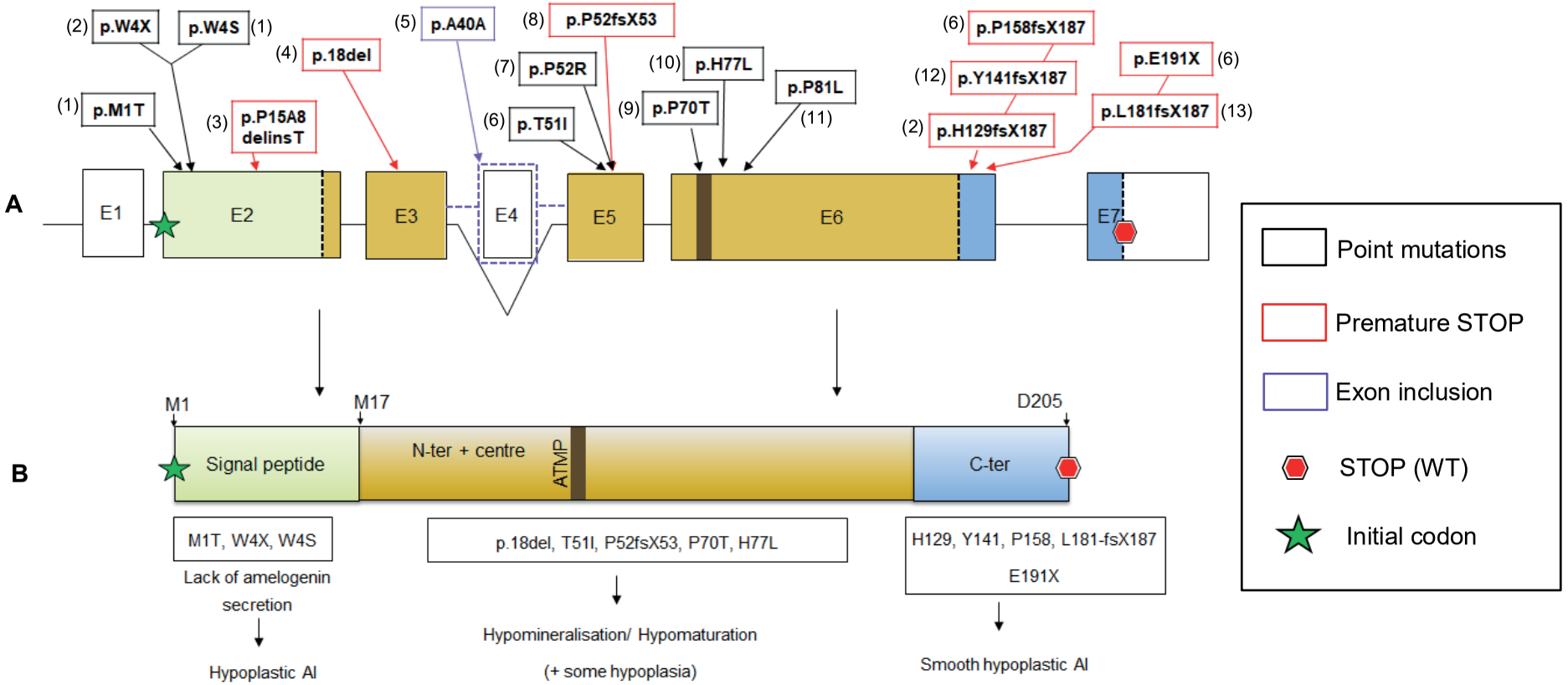


Figure 12 Amelogenin mutations associated with X-linked AI: representation as (A) exon map and (B) protein domains. The nomenclature of the mutations is based on the sequence derived from mRNA formed by the colinear splicing of all 7 human amelogenin exons. The references for the reports identifying the mutations are below in the footnote[1]

[1] **References (next page)**:

(1) KIM, J.-W., SIMMER, J., HU, Y., LIN, B.-L., BOYD, C., WRIGHT, J., YAMADA, C., RAYES, S., FEIGAL, R. & HU, J.-C. 2004. Amelogenin p. M1T and p. W4S mutations underlying hypoplastic X-linked amelogenesis imperfecta. *Journal of Dental Research,* 83**,** 378-383.
(2) SEKIGUSHI, H. 2001. A new mutation in the amelogenin gene causes X-linked amelogenesis imperfecta. Ibid.80**,** 617.
(3) LAGERSTRÖM-FERMÉR, M., NILSSON, M., BÄCKMAN, B., SALIDO, E., SHAPIRO, L., PETTERSSON, U. & LANDEGREN, U. 1995. Amelogenin signal peptide mutation: correlation between mutations in the amelogenin gene (AMGX) and manifestations of X-linked amelogenesis imperfecta. *Genomics,* 26**,** 159-162.
(4) LAGERSTROM, M., DAHL, N., NAKAHORI, Y., NAKAGOME, Y., BACKMAN, B., LANDEGREN, U. & PETTERSSON, U. 1991. A deletion in the amelogenin gene (AMG) causes X-linked amelogenesis imperfecta (AIH1). Ibid.10**,** 971-5.
(5) CHO, E. S., KIM, K.-J., LEE, K.-E., LEE, E.-J., YUN, C. Y., LEE, M.-J., SHIN, T. J., HYUN, H.-K., KIM, Y.-J., LEE, S.-H., JUNG, H.-S., LEE, Z. H. & KIM, J.-W. 2014. Alteration of Conserved Alternative Splicing in AMELX Causes Enamel Defects. *Journal of Dental Research,* 93**,** 980-987.
(6) LENCH, N. J. & WINTER, G. B. 1995. Characterisation of molecular defects in X-linked amelogenesis imperfecta (AIH1). *Human Mutation,* 5**,** 251-9.
(7) KIDA, M., SAKIYAMA, Y., MATSUDA, A., TAKABAYASHI, S., OCHI, H., SEKIGUCHI, H., MINAMITAKE, S. & ARIGA, T. 2007. A novel missense mutation (p. P52R) in amelogenin gene causing X-linked amelogenesis imperfecta. *Journal of Dental Research,* 86**,** 69-72.
(8) LENCH, N., BROOK, A. & WINTER, G. 1994. SSCP detection of a nonsense mutation in exon 5 of the amelogenin gene (AMGX) causing X-linked amelogenesis imperfecta (AIH1). *Human Molecular Genetics,* 3**,** 827-828.
(9) COLLIER, P. M., SAUK, J. J., ROSENBLOOM, J., YUAN, Z. A. & GIBSON, C. W. 1997. An amelogenin gene defect associated with human X-linked amelogenesis imperfecta. *Archives of Oral Biology,* 42**,** 235-242.
(10) HART, P. S., ALDRED, M. J., CRAWFORD, P. J. M., WRIGHT, N. J., HART, T. C. & WRIGHT, J. T. 2002. Amelogenesis imperfecta phenotype–genotype correlations with two amelogenin gene mutations. Ibid.47**,** 261-265.
(11) KIM, Y.-J., KIM, Y. J., KANG, J., SHIN, T. J., HYUN, H.-K., LEE, S.-H., LEE, Z. H. & KIM, J.-W. 2017. A novel AMELX mutation causes hypoplastic amelogenesis imperfecta. Ibid.76**,** 61-65.
(12) GREENE, S., YUAN, Z., WRIGHT, J., AMJAD, A., ABRAMS, W., BUCHANAN, J. & GIBSON, G. 2002. A gene deletion resulting in amelogenin with nine cysteine residues leads to amelogenesis imperfecta. Ibid.47**,** 211-217.
(13) KINDELAN, S., BROOK, A., GANGEMI, L., LENCH, N., WONG, F., FEARNE, J., JACKSON, Z., FOSTER, G. & STRINGER, B. 2000. Detection of a novel mutation in X-linked amelogenesis imperfecta. *Journal of Dental Research,* 79**,** 1978-1982.

1.3.3.2 Characterisation of the effects of amelogenin mutations *in vitro*: Effect of point mutation p.P70T

Investigations are currently focusing on the effect of amelogenin mutations on protein assembly, protein-to-mineral binding, mineral nucleation, protein binding, and intra- and (mostly) extra-cellular processing in X-linked AI. Two point mutations causing AI in humans have been investigated by multiple *in vitro* studies: p.T51I (Lench and Winter, 1995) and p.P70T (Collier et al., 1997). P.T51 and p.P70T are both located in the conserved N-terminal portion of amelogenin, part of the 'A' domain which reportedly drives amelogenin self-assembly (Paine and Snead, 1997). The following Section focuses on the p.P70T mutation, which affects a part of the ATMP domain (the importance of this domain is detailed in the Section 1.2.2.2.1 p. 22 and 1.2.5.2.2 pp. 47 - 49); it is also the mutation for which *in vivo* studies have been carried out (mice carrying the pY64H mutation). *In vitro*, p.P70T mutation has been the focus of multiple studies aiming to understand the underlying mechanisms of AI. The techniques employed include Y2H, SPR, DLS, NMR, Hemaglut., western blotting or ELISA. The r-amelogenin-amelogenin interactions and supramolecular assembly were significantly affected (Moradian-Oldak et al., 2000, Paine et al., 2002, Lakshminarayanan et al., 2010). However, there does seem to be some confusion in the literature regarding the precise location of the p.P70T mutation and this is rectified below (Section 1.3.3.2.1).

1.3.3.2.1 Clarification of amelogenin mutation nomenclature with respect to the p.P70T mutation

The p.P70T mutation was originally described as a p.P41T mutation by Collier et al 1997.The authors described the mutation relative to the amino acid sequences of rodent, pig and cow amelogenin. However, the human sequence is lacking Met[29] so the actual mutation in the human sequence is more accurately described as p.P40T. P40 is highlighted in yellow in the sequence below corresponding the major amelogenin isoform lacking the signal sequence and exon 4 secreted form:

```
1                                        40                                  70
MPLPPHPGHPGYINFSYEVLTPLKWYQSIRPPYPSYGYEPMGGWLHHQIIPVLSQQHPPTHTLQPHHHIP

VVPAQQPVIPQQPMMPVPGQHSMTPIQHHQPNLPPPAQQPYQPQPVQPQPHQPMQPQPPVHPMQPLPPQP

PLPPMFPMQPLPPMLPDLTLEAWPSTDKTKREEVD
```

This nomenclature is confusing because the canonical form listed in the Uniprot data base , Q99217-1, includes the signal peptide but still lacks exon 4 (The UniProt Consortium, 2017). It is based on the sequence derived from mRNA formed by the dominant splice product (Figure 7b, p.16), except that the exon 1 is not translated since it is part of the 5'-untranslated region. Using this canonical sequence the p.P40T mutation becomes the p.P56T mutation as shown below, with the signal peptide highlighted in green:

```
1                                                    56            70
MGTWILFACLLGAAFAMPLPPHPGHPGYINFSYEVLTPLKWYQSIRPPYPSYGYEPMGGWLHHQIIPVLS

QQHPPTHTLQPHHHIPVVPAQQPVIPQQPMMPVPGQHSMTPIQHHQPNLPPPAQQPYQPQPVQPQPHQPM

QPQPPVHPMQPLPPQPPLPPMFPMQPLPPMLPDLTLEAWPSTDKTKREEVD
```

The problem with this canonical sequence is that it provides no means of describing any mutation present in exon 4.To get round this problem some authors describe mutations relative the amelogenin sequence derived from the collinear splicing of all 7 exons that generates a mRNA transcript that includes exon 4. Using this sequence the p.P40T mutation becomes the p.P70T mutation as shown below with exon 4 highlighted in blue:

```
1                                                                       70
MGTWILFACLLGAAFAMPLPPHPGHPGYINFSYENSHSQAINVDRTALVLTPLKWYQSIRPPYPSYGYEP

MGGWLHHQIIPVLSQQHPPTHTLQPHHHIPVVPAQQPVIPQQPMMPVPGQHSMTPIQHHQPNLPPPAQQP

YQPQPVQPQPHQPMQPQPPVHPMQPLPPQPPLPPMFPMQPLPPMLPDLTLEAWPSTDKTKREEVD
```

This nomenclature has been used several times in the literature to describe this mutation as p.P70T, e.g. (Gibson et al., 2007, Zhu et al., 2011) and is the one used in Figure 12A (p. 58), which lists the amelogenin mutations causing AI in human.

The difference of nomenclatures has led to some confusion in the literature. For example Buchko and Shaw (2015) published a report using NMR to study the effects of the p.P70T mutation. They used site-directed mutagenesis against recombinant mouse amelogenin to convert Pro^{71} (analogous to human Pro^{70}) to Thr^{71}. However, they based the amino acid numbering on the sequence corresponding to the major amelogenin isoform lacking the signal sequence and exon 4 which resulted in the wrong proline residue (highlighted in red) being substituted as shown below:

```
1                                        40                                  70
MPLPPHPGHPGYINFSYEVLTPLKWYQSIRPPYPSYGYEPMGGWLHHQIIPVLSQQHPPTHTLQPHHHIP

VVPAQQPVIPQQPMMPVPGQHSMTPIQHHQPNLPPPAQQPYQPQPVQPQPHQPMQPQPPVHPMQPLPPQP

PLPPMFPMQPLPPMLPDLTLEAWPSTDKTKREEVD
```

The authors reported that their “p.P71T” mutation caused amelogenin self-association to occur at lower protein concentrations. However, this has no bearing on the aetiology of the true p.P70T mutation in humans since the actual mutation investigated was p.P100T as shown below:

```
1                                                                        70
MGTWILFACLLGAAFAMPLPPHPGHPGYINFSYENSHSQAINVDRTALVLTPLKWYQSIRPPYPSYGYEP
                                 100
MGGWLHHQIIPVLSQQHPPTHTLQPHHHIPVVPAQQPVIPQQPMMPVPGQHSMTPIQHHQPNLPPPAQQP

YQPQPVQPQPHQPMQPQPPVHPMQPLPPQPPLPPMFPMQPLPPMLPDLTLEAWPSTDKTKREEVD
```

1.3.3.2.2 Amelogenin p.P70T mutation impairs amelogenin self-assembly and mineral binding

P.P70T mutation has been widely studied *in vitro* using recombinant proteins. It was reported to affect the secondary structure of amelogenin (Lakshminarayanan et al., 2010). A Y2H assay also showed that this point mutation affected amelogenin self-assembly by reducing the strength of amelogenin-amelogenin binding by 25% (Paine et al., 2002); a consistent finding with SPR data obtained with r-amelogenins (Paine et al., 2002).

Multiple studies analysed the effect of the p.P70T mutation on the formation of supramolecular amelogenin assemblies: NMR studies carried out in a buffer of 2% deuterated acetic acid ($C^2H_3COO^2H$), 7% deuterated water (^2H_2O) and 91% water (H_2O), at pH 3.0, found that amelogenin carrying the mutation self-assembled at lower concentration and ionic strength than the WT amelogenin (Buchko et al., 2013) suggesting an increased propensity to self-assemble. However, the data should be interpreted carefully since the NMR buffer did not correspond to physiological conditions. DLS and AFM studies using recombinant mouse (M180) or human (H174) amelogenins carrying a polyhistidine tag (His-tag) reported that amelogenin formed larger and more heterogeneous aggregates while carrying the p.P70T mutation (Moradian-Oldak et al., 2000, Zhu et al., 2011). Zhu et al (2011) also found that the p.P70T mutation increased the affinity of amelogenin for apatite and that it delays and constrains apatite crystal growth (Zhu et al., 2011).

1.3.3.2.3 Amelogenin p.P70T mutation impairs interaction with other proteins

As detailed previously in Section 1.2.5.2.2 (pp. 47 - 49), the ATMP domain may be critical in mediating amelogenin interaction with other proteins. Hemagglut. studies showed that the p.P70T mutation precluded ATMP binding to GlcNAc, GlcNAc-mimicking peptides, cytokeratins 5 and 14 or ameloblastin (Ravindranath et al., 1999, Ravindranath et al., 2000, Ravindranath et al., 2001, Ravindranath et al., 2003, Ravindranath et al., 2004). Substrate competition assay, pull-down assays and SPR studies showed that the p.P70T mutation also precluded the interaction between amelogenin and MMP-20 (Tanimoto et al., 2008) , which may significantly hinder the cleavage of amelogenin by MMP-20 as previously reported (Li et al., 2001).

Together, the studies cited above showed the consequence of the p.P70T mutation on amelogenin binding behaviour *in vitro*. Apart from Y2H studies, most of these studies used recombinant (His-tagged) amelogenins, they were carried out under conditions that did not reflect the environment *in vivo*. To validate these findings and elucidate the pathogenesis of AI, *in vivo* evidence is needed.

1.3.4 Pathogenic mechanisms driving AI

A plethora of *in vitro* studies has been carried out (described in Section 1.3.3.2, above) to understand mechanisms underlying AI but their relevance to the situation *in vivo* is never clear. As illustrated in Figure 13 below, combinations of *in vitro, in cellulo* and *in vivo* studies would be better suited for characterising the role of mutations (for example, single point mutations in amelogenin) on the pathogenic mechanisms driving AI. The molecular mechanisms involved in amelogenesis and its pathologies are a current priority in enamel research (Kirkham et al., 2017) as understanding these mechanisms would provide a breakthrough in understanding genotype-phenotype correlations and help develop diagnostic tools, therapeutic interventions as well as informing on better ongoing restorative patient care.

In vitro studies and *in silico* prediction tools do not necessarily require ethical approval and are cost-effective compared with preclinical and clinical studies. They can provide information on protein structure and function and as such are important for gaining the rational understanding required to identify therapeutic targets and develop drugs against these targets. However, *in vitro* and *in silico* studies alone are unable to elucidate with complete confidence the details of EMPs' functions and the effect of mutations on these proteins *in vivo*. Never the less, *in vitro* and *in silico* studies remain useful adjuncts to *in cellulo* and *in vivo* studies using animal models and are often used at the start of a research project to develop hypotheses and inform on downstream research strategies employing *in cellulo* and *in vivo* model systems. Compared to *in vitro* studies, *in cellulo* studies provide a broader picture of the biological and chemical mechanisms involved. They can be carried out in parallel with *in vitro* experiments to guide the design of *in vitro* experiments and to validate the results from the latter. Naturally, at some point the research pipeline must include *in vivo* studies focusing on the whole tissue, organ or the entire organism. Animal models are necessary for the initial characterisation of a pathological phenotype at all stages of amelogesis as human-derived material is limited to clinical observations of teeth *in situ* or analysis of exfoliated teeth. Animal models are also a means to validate *in vitro* or *in cellulo* data at the later stages of research and of course any pharmacological interventions developed must be trialled in a suitable enamel model prior to human trials. To study amelogenesis and AI, mice have been the most useful model as they have the particularity of having continuously growing incisors (Smith et al., 2017b). A good example of the above strategy is the detailed phenotypic analysis of mice phenocopying human AI together with *in cellulo* and *in vitro* experiments allowed the underlying aetiological mechanism to be identified; in this case the p.Y64H mutation in amelogenin was shown to lead to pathological ER stress in the ameloblasts and

ameloblast apoptosis. Moreover, elucidation of the aetiological mechanism involved allowed for a therapeutic intervention to be applied which effectively rescued the phenotype in affected heterozygous female mice (Brookes et al., 2014). This is discussed in more detail in the following pages.

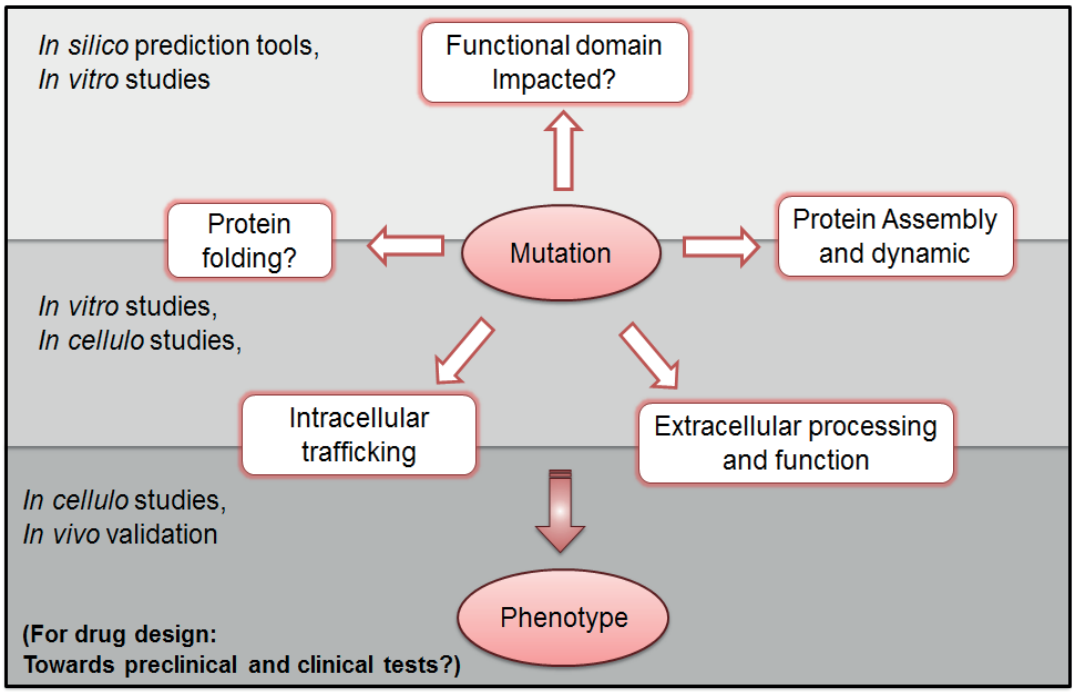


Figure 13 Combination of approaches to consider to study the effect(s) of a single point mutation on amelogenin assembly and function(s) and trafficking.

1.3.4.1 *In vivo* studies: the use of mouse models to study amelogenesis and AI

Phenotyping of teeth from AI patients is potentially confounded by post-eruptive changes (e.g. mechanical damage and demineralisation) that might occur during the time the enamel has spent in the mouth. In extreme cases, there may be no enamel left that can be analysed as a prelude to developing testable hypotheses as to the aetiological factors driving that particular case of AI. Obtaining unerupted genotyped human embryonic teeth would be almost impossible and certainly ethically questionable and precludes the direct study of human amelogenesis. However, where enamel survives on exfoliated human AI teeth (especially if the tooth is unerupted e.g. an impacted 3rd molar) enamel composition and ultrastructure can provide a record of sorts as to what went wrong during amelogenesis, allowing hypotheses to be developed and subsequently tested using *in vitro, in cellulo* or *in vivo* animal models. Rodents, whose incisor enamel forms continuously, have proved to be an invaluable animal model in AI research as all stages of development are present on a single incisor as detailed below (Section 1.3.4.1.1).

1.3.4.1.1 Rodent incisors as models for studying amelogenesis and aetiologies driving AI

Rodent incisors have evolved to grow continuously in response to attrition at the biting edge of the tooth caused by gnawing on hard foods (Hu et al., 2014). In addition, rodents actively wear the incisal edges to keep them chisel-sharp during thegosis; a behavioural process which is distinct from mastication related to feeding (Byrd, 1997). Continuous tooth formation and growth originates from the labial cervical loop located at the apical end of the incisor. It comprises permanent reservoirs of stellate reticulum cells, providing a stem cell niche. These cells generate so-called “transit-amplifying cells”, whose progeny ultimately differentiate into ameloblasts. (Harada et al., 1999). Tooth development occurs in the apical-to-distal (or root to biting edge) direction, with the ‘newest’ dental tissue forming towards the apical end, and the most mature tissues at the distal end (illustrated in Figure 14). Thus, all steps of amelogenesis can be visualised simultaneously on a single incisor and at any time. This is not feasible in other models where teeth are not continuously growing (e.g. pig). In such cases, developing teeth have to be dissected from the jaws of several animals at different developmental stages to capture all stages of amelogenesis.

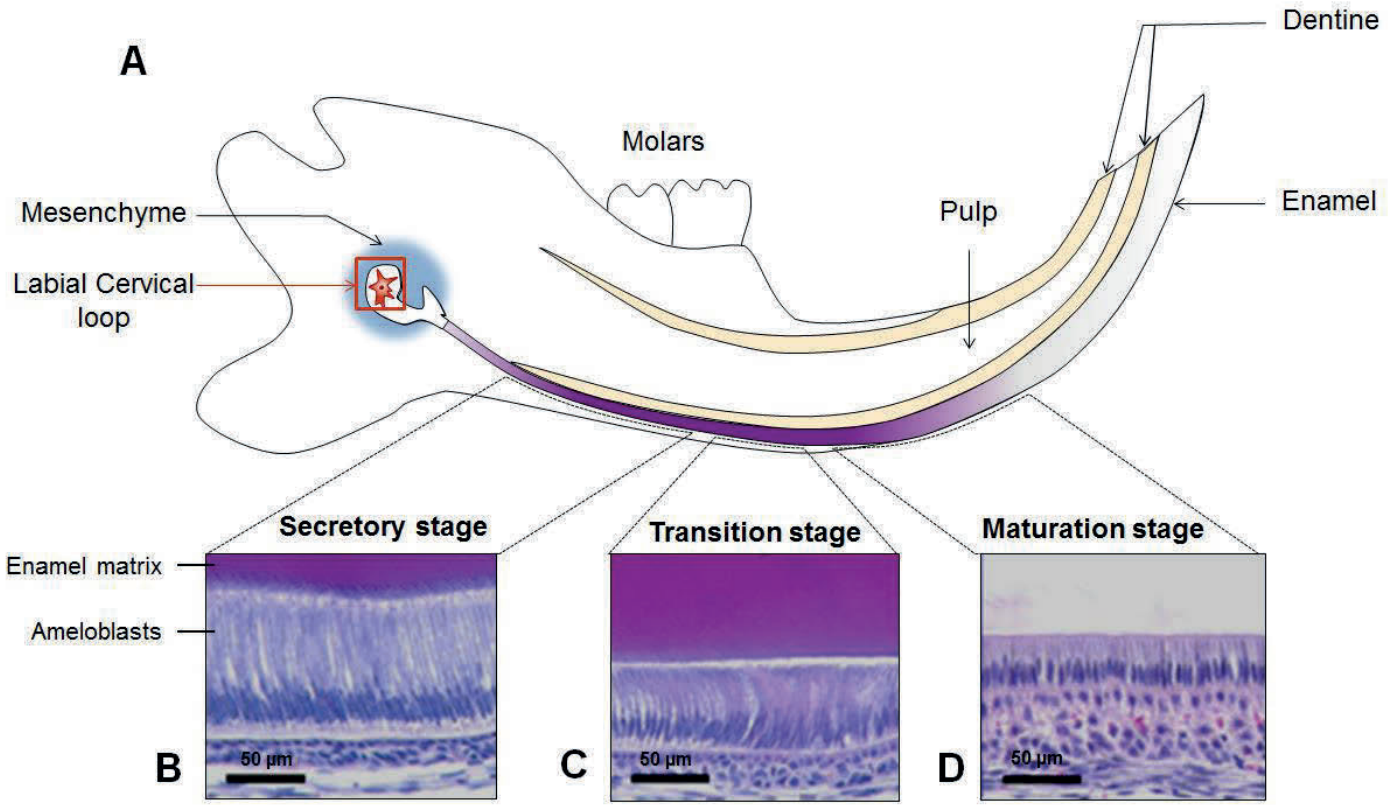


Figure 14 (A) Schematic cross section and (B–D) histology of the murine incisor at secretory, transition and maturation stages. Elements from this figure have been taken from previous publications: panels B-D adapted from (Barron et al., 2010)[1].

[1] The histology analyses of mouse incisors in panels B, C, D are reproduced from Barron et al. (2010), Human Molecular Genetics 2010;19 (7):1230-47 by permission of Oxford University Press.

1.3.4.1.2 Use of mouse models in studying amelogenesis and AI – current state of the art

The advantage of using mouse models in preclinical studies is that mice are relatively easy to modify genetically and various mutations can be introduced to investigate correlations between genotype and phenotype, understand protein function and the aetiological mechanisms driving AI. Mouse and human developing enamel proteomes are highly conserved and the cellular and extracellular event occurring during amelogenesis are very similar. The enamel produced in both species is essentially identical at the biochemical level though the structural organisation of rodent incisors is somewhat unique in that the prisms adopt a more obvious decussating arrangement. Mouse 'null' models or KO models are those where the expression of a specific gene has been abolished. Although the KO models cannot explain fully the exact details around phenomena such as protein-protein interactions and cellular (or extra-cellular) processing, they can show that the protein in question has a specific role to play in amelogenesis overall or more usefully in a specific facet of amelogenesis.
Various murine KO models have been engineered that lack the expression of genes involved in amelogenesis and the resulting enamel (developing or mature) characterised. The KO models studied to date include KOs of amelogenin (Gibson et al., 2001), ameloblastin (Fukumoto et al., 2004), enamelin (Hu et al., 2008, Smith et al., 2009), MMP-20 and KLK4 (Caterina et al., 2002, Wright et al., 2009, Simmer et al., 2009, Yamakoshi et al., 2011, Hu et al., 2016, Núñez et al., 2016), amelotin (Núñez et al., 2016), FAM83H (Wang et al., 2016), WDR72 (Katsura et al., 2014), SLC24A4 (Stephan et al., 2012), GPR68 (Parry et al., 2016), FAM20A (Vogel et al., 2012) and DLX3 (Morasso et al., 1999).
The relevance of KO mouse models in reflecting human AI has been discussed by Wright et al, (2009). Mice KOs models might be expected to accurately model human cases of AI where the human mutation effectively abrogates expression of the protein in question (e.g. a mutation that destroys a promoter region). However, rather than knocking out a protein completely, numerous cases of human AI involve more subtle changes to the protein in question e.g. deletion of part of a protein, amino acid substitutions, affected splicing recognition sites resulting in faulty splicing and the introduction of premature stop codons resulting the expression of a truncated protein. These mutations may cause a different AI phenotype to that obtained when the protein in question is simply knocked out. This is because a KO model is based on loss of function due to complete absence of protein whereas an abnormal protein that is expressed may provide a 'gain of toxicity' meaning that although the protein is present it may actually be frankly toxic in addition to any loss of function the mutation may

have caused. In short, mouse KO models can only tell us so much; to understand the aetiology specific cases of human AI there is a need for mice models carrying mutations that replicate the specific mutations that cause human AI.

1.3.4.1.3 Mouse model carrying the p.Y64H mutation

As briefly eluded to above, a recent mouse model was described that carried the substitution of Tyr^{64} into a histidine (p.Y64H) mutation in amelogenin. The mice phenocopied human X-linked AI (Barron et al., 2010) and although the p.Y64H mutation has not been identified in any cases of human AI to date, it lies very close to the p.P70T mutation known to be associated with human AI (Collier et al., 1997) with both mutations lying in the previously discussed ATMP domain. The mice exhibited severe defects of enamel bio-mineralisation, loss of ameloblast phenotype and disruption of the ordered ameloblast monolayer due to premature apoptosis of affected ameloblasts. Affected ameloblasts failed to secrete amelogenin into the extracellular matrix, instead, amelogenin accumulated intracellularly, disrupting the secretory organelles, rough ER and Golgi apparatus (Barron et al., 2010). The observed abnormal retention of p.Y64H mutant amelogenin (along with ameloblastin and potentially other matrix proteins) led to the identification of ER stress as a possible factor driving AI in this model. The ameloblasts eventually succumbed to apoptosis due to the ER stress activating the so-called UPR, which had tipped towards pro-apoptotic mode (the UPR is mentioned in Section 1.2.5.1, p.42 and UPR fate is discussed in greater details in Section 1.3.4.2, pp. 75 - 79). A concomitant increase of UPR markers was shown in the ameloblasts expressing mutant amelogenin (Brookes et al., 2014). The same group demonstrated that 4-phenylbutyrate, a histone deacetylase inhibitor and so-called chemical chaperone, could rescue the phenotype in heterozygous female mice, as ameloblast cell viability was maintained (apoptosis was dramatically decreased) and the mineralisation of the enamel was restored (Brookes et al., 2014).

The precise molecular mechanisms leading to abnormal retention in the secretory pathway, ER stress and UPR-triggered apoptosis and pharmacological rescue of the phenotype in these mice remain to be unravelled. The UPR in theory attempts to relieve the ER stress in the first instance, but upon prolonged ER stress, such as the abnormal retention of p.Y64H mutant amelogenin, the UPR triggers ameloblast apoptosis which here would disrupt the ameloblast monolayer and compromise amelogenesis. The hypothesis, as it applies to female mice heterozygous for the mutation, in which only ~50% of the ameloblasts are affected (due to random inactivation of the X chromosome – Section 1.3.3.1 p. 56) is illustrated in Figure 15 (pp. 72 - 73). This figure also shows the hypothesised means by which intervention

with 4-phenylbutyrate rescues the phenotype; the drug inhibits apoptosis so the ameloblast monolyayer is maintained and even though half the ameloblasts expressing the p.Y64H mutant fail to secrete protein, the remainder expressing the WT amelogenin gene are able to complete amelogenesis as normal. The next Section (1.3.4.2, pp. 74 - 79) describes the UPR signalling cascades in more detail, and addresses the molecular basis of how the UPR switches from pro-survival to pro-apoptotic mode.

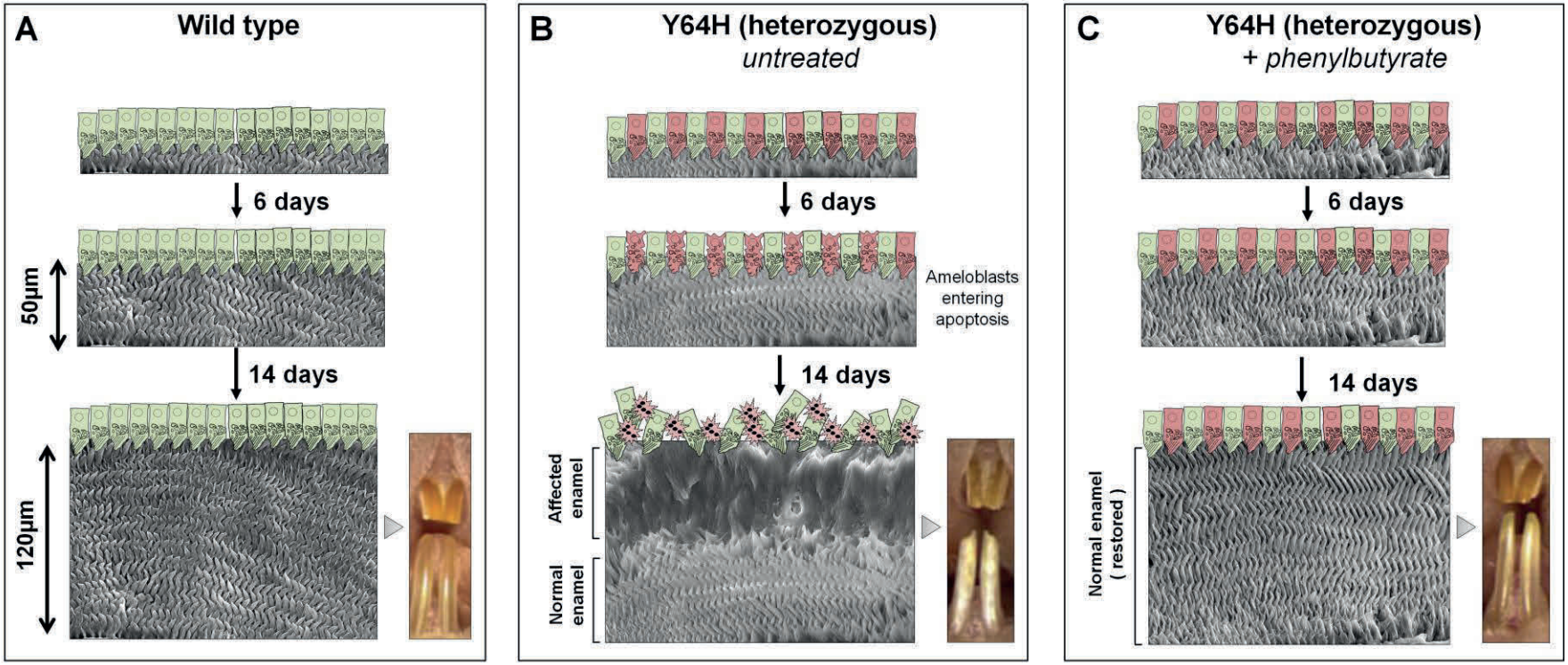

A
Wild type
6 days
50µm
14 days
120µm
B
Y64H (heterozygous)
untreated
6 days
Ameloblasts entering apoptosis
14 days
Affected enamel
Normal enamel
C
Y64H (heterozygous)
+ phenylbutyrate
6 days
14 days
Normal enamel (restored)
WT ameloblast
Affected ameloblast surviving ER stress
Affected ameloblast becoming apoptotic
Affected ameloblast dead

Figure 15 Illustration summarising the aetiological mechanism driving AI in mice carrying the p.Y64H amelogenin mutation. (A) In WT mice, the ameloblasts form a monolayer and secrete EMPs over 14 days and the resulting mature enamel has an ordered decussating structure. (B) for female heterozygous mice the unaffected ameloblasts secrete matrix normally for 6 days while affected ameloblasts are merely bystanders in the process as they endure ER stress with the support of the UPR. Crucially, the ordered monlayer is manintained during this time and the initially secreted enamel is normal. However, after about 6 days the the UPR switches to pro-apoptotic mode and affected ameloblasts die which disrupts the ameloblast monolayer and the enamel secreted thereafter has lost its decussating structure. (C) The addition of phenylbutyrate mitigates UPR induced apoptosis and preserves the integrity of the ameloblast monolayer restoring the phenotype. and maintains (Reproduced from Brookes et al 2014)[1].

[1] Photographies of mice incisors are reproduced from Brookes et al. (2014) with the permission of Oxford University Press under under the terms of the Creative Commons CC BY license (https://creativecommons.org/licenses/by/4.0/).

1.3.4.2 ER stress, the UPR and proteopathic diseases

As detailed previously (Section 1.2.5.1), proteins destined for secretion fold in the ER and attain their correct functional 3-dimensional conformation. This is important since misfolded proteins may be trapped in a conformation that could be inactive and/or promote pathological intracellular aggregation which may be more toxic to the cell than the simple loss of function of the protein. When neither ERAD or ERAF (described in Section 1.2.5.1) comprising the first-line defence mechanism against misfolded proteins is sufficient to alleviate ER stress, the UPR sensors PERK, IRE1α and ATF6 are triggered (detailed in Section 1.3.4.2.1). The UPR is activated and supports the cell as it attempts to handle its secretory burden by triggering the expression of downstream pro-survival genes. However, the UPR is not only triggered when mutated proteins are being trafficked. WT proteins can also misfold and in specialised secretory cells, where the secretory traffic is heavy even under normal conditions, e.g. insulin producing pancreatic cells (Lipson et al., 2006) and WT secretory ameloblasts (Wu and Kaufman, 2006, Tsuchiya et al., 2008), the UPR is activated as a normal physiological response to a particularly heavy period of secretory activity.

During the initial response, the UPR attempts to alleviate ER stress caused by misfolded proteins by increasing the ER folding or degradation capacity of the cell, reducing the rate of global gene expression to reduce the number of client proteins entering the ER and by increasing the actual amount of ER in the cell (Lin et al., 2007). As mutated proteins may show greater propensity to misfold and aggregate, this may cause abnormally lasting ER stress which cannot be alleviated. In such cases the UPR switches from a pro-survival mode to pro-apoptotic mode (Ron and Walter, 2007, Hetz, 2012). This switch underpins the so-called proteopathies or conformational diseases; including some cancers, neurodegenerative diseases (Lindholm et al., 2006), diabetes (Scheuner and Kaufman, 2008) and cystic fibrosis (Younger et al., 2006) and now, AI (Brookes et al., 2014).

The following Section **(i)** describes the three UPR pathways in more detail (Section **1.3.4.2.1**), **(ii)** addresses the balance between pro-survival and pro-apoptotic outcomes of the UPR (**1.3.4.2.2**) and **(iii)** addresses the consideration of UPR in dental pathologies (Section **1.3.4.2.3**)

1.3.4.2.1 Activation of the UPR and signalling cascades and UPR fates

The UPR comprises three signalling cascades (Figure 16) which are initiated upon the activation of a sensor protein: PERK, IRE1α and ATF6. At rest the three sensors are maintained in their inactive state by the bound ER chaperone BiP. BiP preferentially binds to misfolded proteins, so when misfolded proteins are in excess, BiP unbinds from the sensor, which then becomes activated (Bertolotti et al., 2000, Shen et al., 2002) triggering downstream signalling cascades.

- **ATF6 pathway (Figure 16A)**

ATF6 is a transmembrane protein (Haze et al., 1999) that is inactive when bound to BiP under quiescent conditions. It is activated upon release of BiP, after which ATF6 is transported to the Golgi by vesicles (Shen et al., 2002) where it is cleaved by proteases SP1 and SP2 (Ye et al., 2000). Its cytosolic domain (pATF6 N), which contains a basic leucine zipper, reaches the nucleus where it binds directly and specifically to DNA target sequences (Yoshida et al., 1998, Yoshida et al., 2000). which results in the transcription of chaperones BiP, GRP94, ERp57, calreticulin and components involved in the ERAD pathway (Adachi et al., 2008).

- **IRE1 pathway (Figure 16B)**

IRE1 was the first known UPR sensor (initially identified in yeast) and is also the most highly conserved (Cox et al., 1993). The IRE1 pathway is activated on release of bound BiP under conditions of ER stress but can also be independently activated by direct binding of unfolded proteins to IRE1 (Credle et al., 2005, Gardner and Walter, 2011). Activation of IRE1 starts with IRE1 autophosphorylation and oligomerisation (Shamu and Walter, 1996, Bertolotti et al., 2000, Ron and Walter, 2007). This activates the RNase domain of IRE1α (Hetz and Glimcher, 2009), which is in the cytosolic portion of IRE1 and catalyses in the unconventional splicing of X-box binding protein 1 (XBP1) mRNA. This splicing, produces a mRNA that will be further translated into the spliced isoform XBP1 (pXBP1s) (Gonzalez et al., 1999, Uemura et al., 2009). pXBP1s is itself a transcription factor (Yoshida et al., 2001) that upregulates genes involved in ERAD, chaperone synthesis, and expression of components necessary for lipid biosynthesis (necessary for increasing ER membrane biogenesis). In addition, activated IRE1α independently triggers activation of the transcription factors JNK or TRAF/JNK which lead to cell death (Yoshida, 2007) or cell rescue (Ogata et al., 2006) depending on the precise flux and nature of the UPR signalling occurring.

- **PERK pathway (Figure 16C)**

The PERK pathway also involves the release of bound BiP (Bertolotti et al., 2000), with resulting dimerisation and transphosphorylation (Liu et al., 2000). Once activated, PERK phosphorylates and represses eIF2α, a translation initiation factor, which causes an overall reduction in protein translation, in an attempt to alleviate the ER stress (Scheuner et al., 2001). When eIF2 is repressed, ATF4 is preferentially translated and regulates expression of proteins involved in metabolism and oxidative stress resistance, protein folding, and autophagy (Wek et al., 2006). ATF4 also activates transcription of C/EBP homologous protein (CHOP, also named GADD153) which contributes to cascades leading towards cell death (Wek et al., 2006, Brewer, 2014). Other than eIF2 repression cascade, PERK induces NrF2 (Cullinan et al., 2003), a transcription factor that promotes the synthesis of cytoprotective components (Kensler et al., 2007).

A more detailed analysis of the 3 signalling arms is beyond the scope of this review. Suffice to say that the precise scope and nature of the signalling involved, especially with regard to cross talk between the 3 arms, is still under investigation. The essential point is that integration of the signals emerging from the 3 arms dictates whether the UPR remains in pro-survival mode or switches to pro-apoptotic mode; i.e. the “match point” decision.

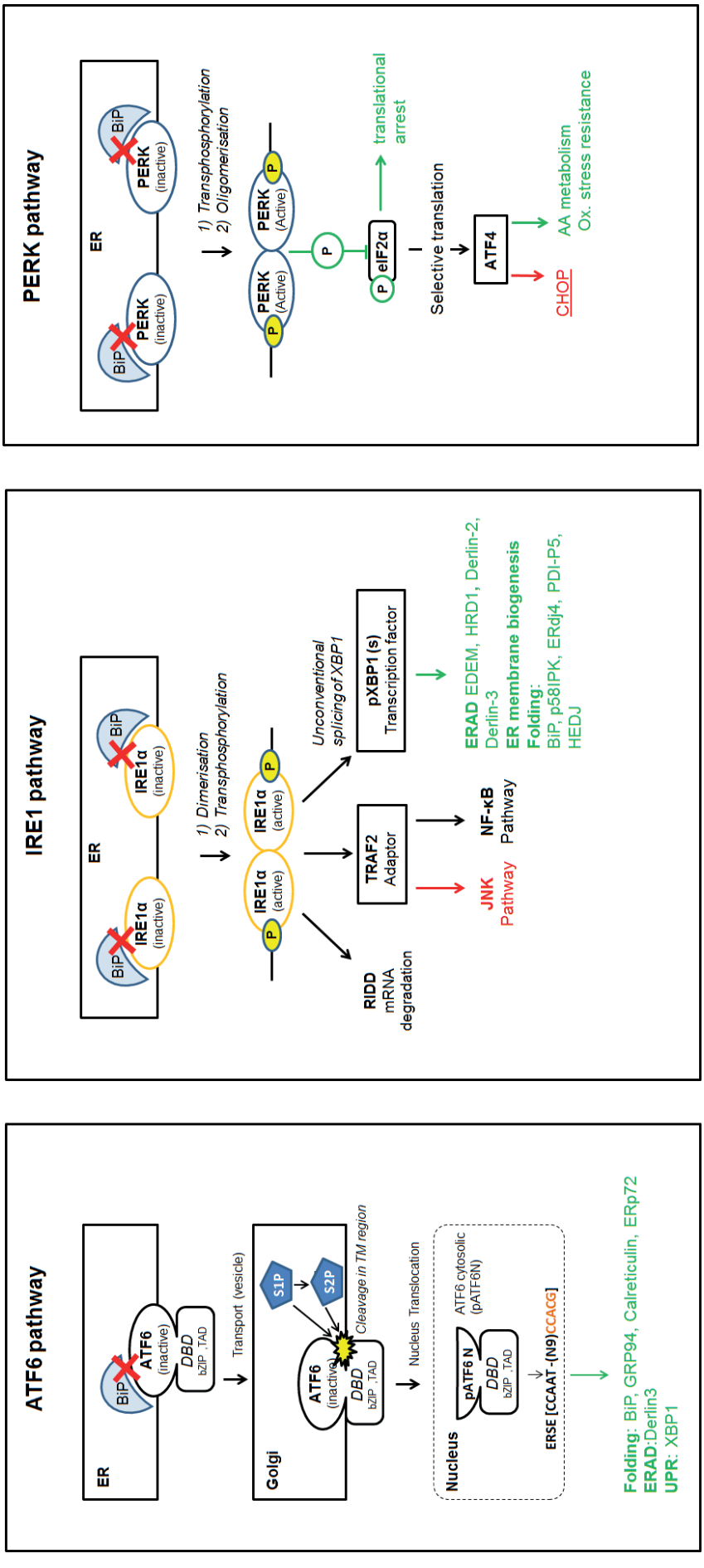


Figure 16 Three arms of the UPR to ER stress, leading to proteostasis or apopotosis.

1.3.4.2.2 UPR fates: “Match-point” decisions

The UPR is an adaptive mechanism that attempts to alleviate ER stress but it can be thought of as a double-edged sword as it switches to pro-apoptotic mode upon prolonged stress. The duration and intensity of stress that trigger the pro-apoptotic mode depend on the specific context of cellular stress. The cross-talk and integrated signals originating from the three sensor-led arms of the UPR (Illustrated in Figure 17) is a key element determining the UPR fate (Tabas and Ron, 2011, Chen and Brandizzi, 2013).

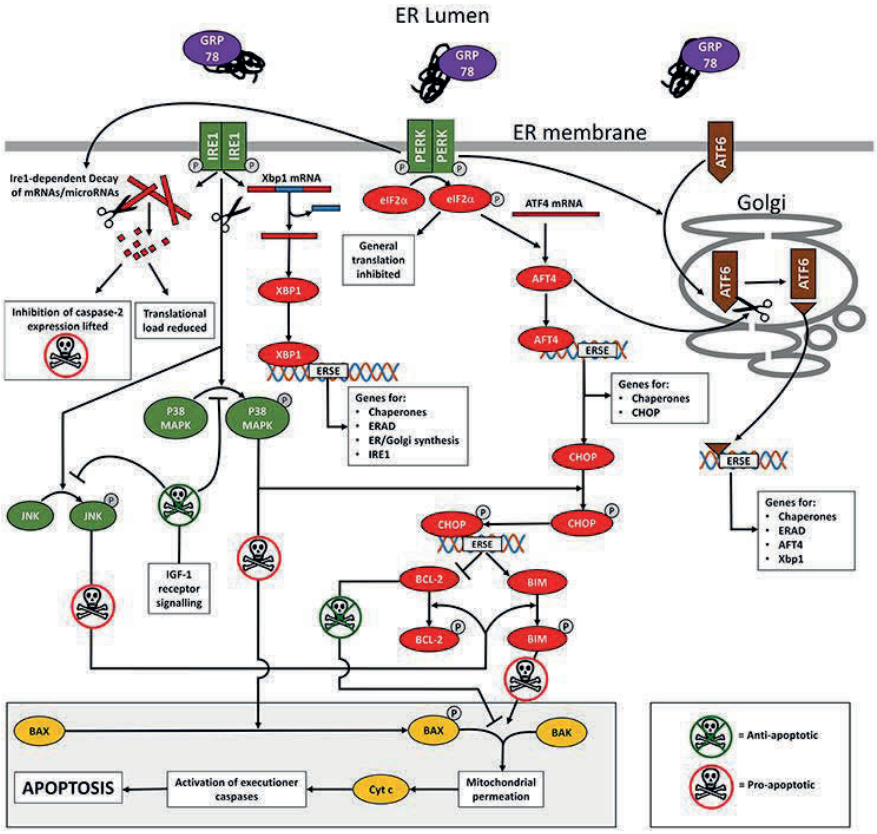


Figure 17 Cross-talk of the integrated signals driven by the 3 arms of the UPR.[1]

PERK and IRE1α pathways are known to be either pro-survival or pro-death, and their prolonged activation reportedly converge towards an apoptotic endpoint (Hetz and Glimcher, 2009, Tabas and Ron, 2011). Current studies focus on signal integration (Hetz and Glimcher, 2009, Brewer, 2014). How the components are regulated, and how crosstalk and kinetics are synchronised, is still a matter of debate though fine-tuning of IRE1 and prolonged PERK signalling are known to play an central role (reviewed by Hetz and Glimcher, 2009).

[1] Figure 17 is reproduced from Brookes et al. (2017) under the terms of the Creative Commons license CC BY (https://creativecommons.org/licenses/by/4.0/).

The molecular mechanisms driving UPR fate (pro-survival or pro-apoptotic modes) are currently topic of interest, as increasing numbers of publications report UPR as a pivotal driver of many human diseases (Wang and Kaufman, 2016).

1.3.4.3 Amelogenin p.Y64H mutation: consequences of single amino acid change on amelogenin binding behaviour and intracellular trafficking

The precise molecular mechanisms leading to abnormal retention of p.Y64H amelogenin, in the secretory of pathway of mouse ameloblasts, remains to be unravelled. For that, two questions are to consider:
(1) At the amino acid level, how can a single substitution of a tyrosine into a histidine possibly alter amelogenin biochemical properties? The properties of both amino acids will be detailed in Section 1.3.4.3.1, to propose avenues for consideration.
(2) When p.Y64H amelogenin was co-transfected with WT ameloblastin in COS-7 cells, the apoptosis was increased (Brookes et al., 2014). Are there any other factors than solely amelogenin aggregation to consider, to explain its abnormal retention in the ER? Section 1.3.4.3.2 proposes avenues for consideration.

1.3.4.3.1 Tyrosine and Histidine chemical properties

At the local (amino acid) level, how can the substitution of Tyr^{64} into a histidine cause such change of amelogenin binding behaviour? At this stage of research, it would be difficult to predict the effect of a single amino acid substitution on amelogenin structure and function, as amelogenin tertiary structure has not been elucidated (see Section 1.2.1.3.3, p. 20). This Section reminds briefly the biochemical properties of both amino acids (shown in Figure 18 overleaf).

- **Tyrosine (Figure 18A)**

Tyrosine is a hydrophobic amino acid. With its aromatic ring, it mediates stacking interactions with other aromatic amino acids. In proteins, it is preferentially buried in protein hydrophobic cores and tyrosine can be substituted by another aromatic amino acid (Betts and Russell, 2003). It comprises a reactive hydroxyl group that confers its polarity and allows hydrogen bonding (Baker and Hubbard, 1984). It allows interaction with non-carbon atoms and provides with a site for phosphorylation, important in the cellular signal transduction (Hunter, 2014).

- **Histidine (Figure 18B)**

Histidine comprises an imidazole ring. It is ionisable and can carry a positive charge. The imidazole ring's tertiary amine comprises a lone pair of electrons that make it a nucleophile and hydrogen bond acceptor and confers its basic properties, while its

secondary amine is an electrophile. This confers histidine/imidazole its chemical versatility and adds complexity to the type of interactions it can engage into: It can be involved in π – π stacking, however there is a charge involved; It can also form (cation – π) interactions, which can be attractive or repulsive depending on histidine protonation state (Liao et al., 2013); (hydrogen – π) bonds and can coordinate with metals (Liao et al., 2013). With its ability to accept or donate protons, histidine is often found in enzyme catalytic triads as a nucleophile activator (Betts and Russell, 2003). It can also coordinate with metal cations, hence its presence in metalloproteins (Strange et al., 1987). With a pKa close to 7, it is present as both protonated (charged) and unprotonated (uncharged) at physiological pH; therefore in proteins, it is not clear whether it will be exposed to the surface or buried in the hydrophobic core. Due to these complex and unique properties, histidine cannot easily substitute to any other amino acid.

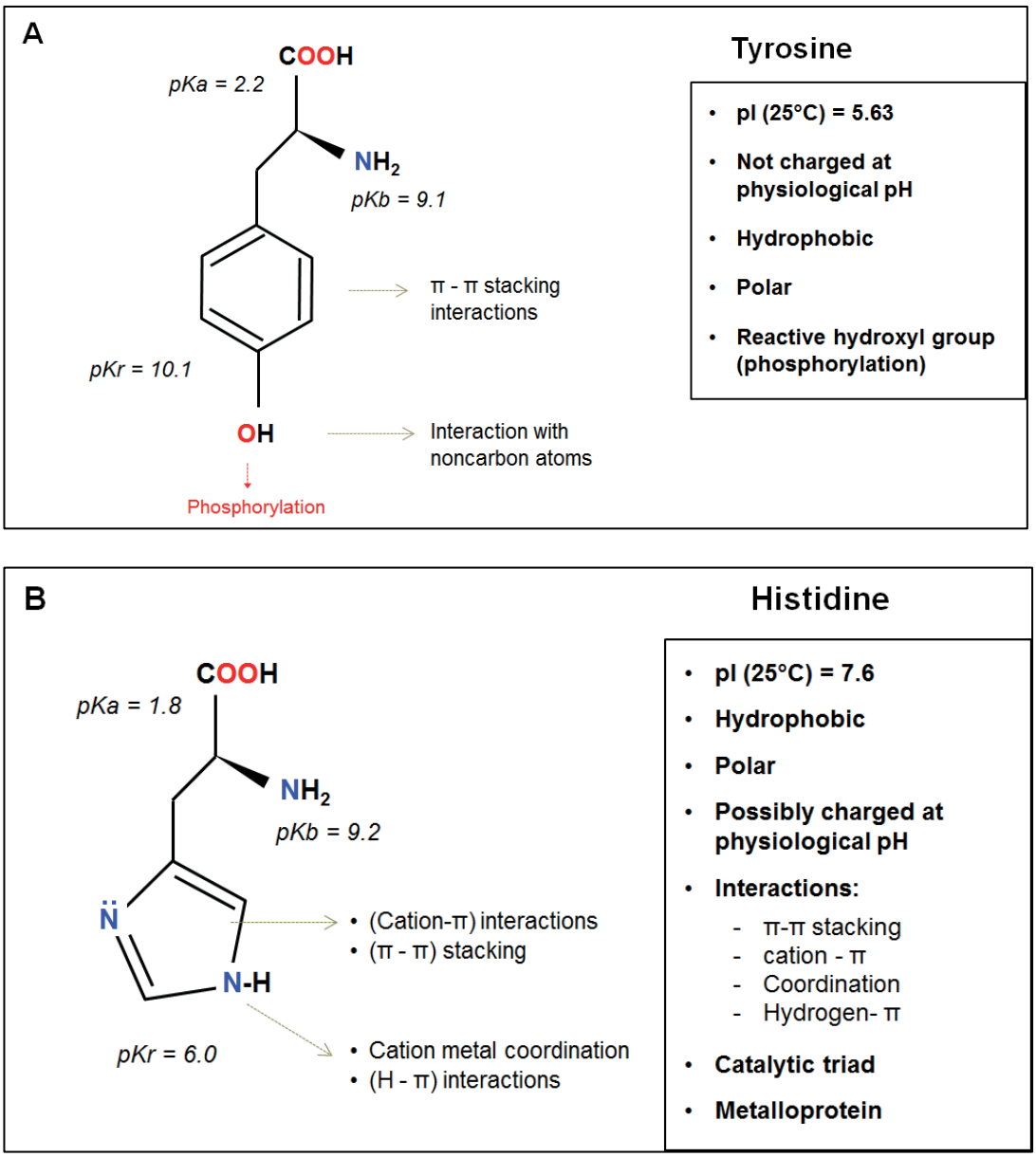


Figure 18 Tyrosine and Histidine: structures and chemical properties.

So, the properties of tyrosine and histidine differ in that:

- Tyrosine is more likely buried in hydrophobic core and can be involved in π - π stacking interactions. Such properties are not that obvious for histidine since histidine can be protonated.

- It is not always clear when histidine will be buried in the hydrophobic core or be exposed at the surface. It may thereby be hypothesised that the single substitution p.Y64H could expose to the solution a domain of amelogenin which possibly should be buried.

Another point to consider is the location of the point mutation. It occurs in the ATMP domain, which, as Wald et al (2017) suggested, might be critical in driving amelogenin binding behaviour (see Section 1.2.2.2.1, p.22, for detailed explanation).

These hypotheses cannot be yet confirmed, as the amelogenin 3-dimensional structure has not been solved (Section 1.2.1.3.3, p.20). Instead, it is more realistic to consider studying amelogenin binding behaviour, and how this is affected by the p.Y64H mutation. This is indeed among the aims of this thesis.
Amelogenin binding behaviour is assumingly affected by the p.Y64H mutation, but to which extent, and how this causes its abnormal retention in the ER and subsequent UPR remains to study. The mutation may not only affect amelogenin self-assembly, but also affect amelogenin interaction with other proteins during its intracellular transit, as explained below in Section 1.3.4.3.2.

1.3.4.3.2 Co-transfection of p.Y46H mutant amelogenin with WT ameloblastin in COS-7 cells increased apoptosis

In cellulo, ER stress was observed in cells co-transfected with p.Y64H amelogenin and WT ameloblastin, but not with amelogenin p.Y64H alone (Brookes et al., 2014). The ameloblastin may be overexpressed in this case compared with its expression in ameloblasts *in vivo*, which may have increased further the protein load in the ER. It provides a hint to consider, though, that the p.Y64H mutation in amelogenin may cause ER stress by affecting amelogenin binding beviour with itself and/or with other proteins. Amelogenin self-assembles and may interact with other proteins including ER components and ameloblastin, as reviewed in Section 1.2.5.2.1 (pp. 45 – 47). The p.Y64H mutation impairing amelogenin binding behaviour may thereby impair the balance of protein-protein interactions necessary for amelogenin trafficking and secretion. The situation is illustrated below in Figure 19 below.

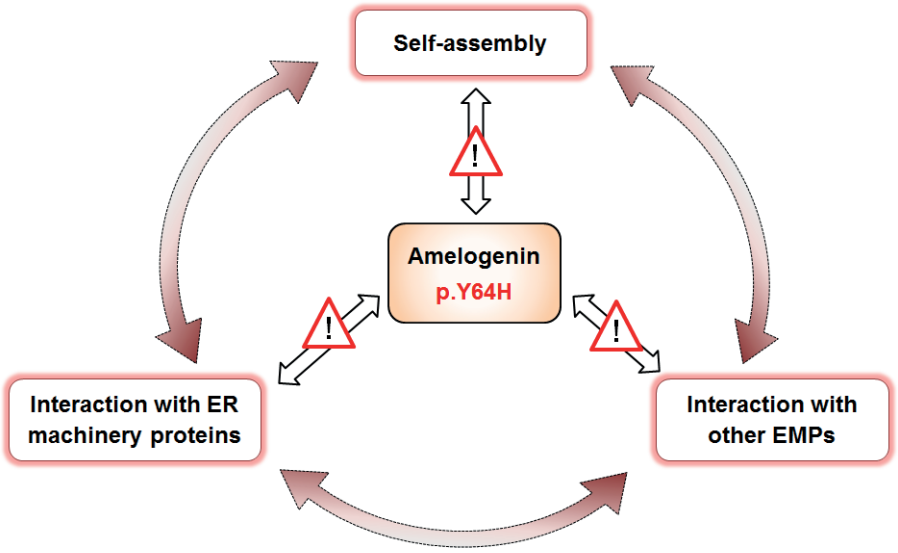


Figure 19 Amelogenin p.Y64H mutation affects amelogenin binding behaviour and therefore may perturb the balance of protein-protein interactions involving amelogenin: hypothesis.

Bearing in mind that p.Y64H mutation may affect the balance of amelogenin interactions with itself or other potential binding partners during its intracellular transit, the first step should be to analyse how p.Y64H mutation affects amelogenin self-assembly. To this end, large amounts of WT and p.Y64H mutant r-amelogenins were needed. The next Section reviews the challenges met with production and purification of amelogenins (native and recombinants).

1.3.5 Need for relevant expression and purification system for future amelogenin protein-binding assays

To gain better insight of molecular mechanisms driving AI and to identify therapeutic targets, *in vitro* studies are a necessary initial step prior to *in cellulo* and *in vivo* studies. This thesis focuses on the possible pathogenesis mechanism(s) of p.Y64H-associated AI in a mouse model which phenocopies human AI. *In vitro* functional studies designed to characterise intracellular amelogenin-amelogenin interactions would first require a reliable and plentiful source of purified wild type and p.Y64H amelogenins. Purifying native proteins from developing enamel is challenging when using teeth from large animals, such as the pig but the amounts of secretory stage enamel protein obtained from existing mice models expressing p.Y64H amelogenin would simply be too small to be of any use in protein binding studies. The cost of genetically modifying a large animal such as a pig to express p.Y64H amelogenin and the associated animal husbandry costs would be prohibitive. In addition, since p.Y64H amelogenin is not secreted and does not accumulate in the enamel matrix, the only biological source of p.Y64H amelogenin would be from the affected ameloblast monolayer which would still only return minuscule amounts of amelogenin.
The only viable source of mutated amelogenin (or for that matter WT rodent amelogenin) is therefore via recombinant technology.

Recombinant proteins are widely used as therapeutic agents and as tools to study structure-function relationships, protein interactions with other molecules and as antigens for antibody production. *E. coli* based expression systems are the most widely used methodology for producing r-amelogenin even though post translational phosphorylation of Ser^{16} will be absent. Baculo virus (Taylor et al., 2006, Xu et al., 2006) and Leishmania (Yadegari et al., 2015) expression systems, having the potential to carry out post translational phosphorylation, have been used to produce r-amelogenin but as yet do not appear to have been widely adopted, perhaps due to uncertainty as to whether the amelogenin was phosphorylated or low yield. In contrast, a yeast-based expression system has been reported to generate correctly phosphorylated r-amelogenin (Cheng et al., 2012) but again has not been widely used. Efforts have been made previously to purify r-amelogenin. One of the most frequently used purification methods, published for rM179 starts with cell lysis under denaturing conditions (6 M guanidine hydrochloride), centrifugation, precipitation with ammonium sulphate, and preparative C4 reverse phase column chromatography in 0.1% trifluoric acid/acetonitrile (Simmer et al., 1994). This method has been widely used for *in vitro* studies e.g. characterising amelogenin assembly into nanospheres (Moradian-Oldak et

al., 1994) but does not provide amelogenin at single band purity. Moradian-Oldak et al (1994) noted recombinant mouse amelogenins were co-chromatographed with contaminating *E. coli* proteins. Nickel-nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography was also used to purify His-tagged r-amelogenins for functional studies (Moradian-Oldak et al., 2000).

1.3.6 Protein-protein interactions and existing analytical techniques

As detailed in Section 1.2, the developing enamel matrix is comprised of several gene products whose function appears to be associated with protein-protein interactions – the presence of amelogenin nanospheres is the prime example of this. Understanding these protein interactions is key to understanding normal amelogenesis as well as the aetiology of AIs driven by perturbed protein interactions. One aim of this thesis was to develop a simple method for studying protein interactions that also provided a high throughput with a view to developing a screening tool that might be used in the future to assess the impact of drugs on pathological enamel protein-protein interactions. The method ought to mimic the *in vivo* environment as necessary to properly characterise protein-protein interactions.

1.3.6.1 Mathematical description of protein-protein interactions

Protein recognition and binding is driven by hydrophobic bonding (Young et al., 1994), van der Waals forces (Roth et al., 1996), electrostatic bonds (Sheinerman et al., 2000), hydrogen bonds (Jiang and Lai, 2002) or covalent bonds. Whether these interactions can occur depends on the conformations and the accessible surface area of the proteins involved (Jones and Thornton, 1996). These 'intrinsic' factors themselves also depend on the environment in which the protein-protein interactions occur such as pH, ionic strength, nature of the solvent, temperature or redox conditions.

Protein-protein interactions are kinetic phenomena that can be described mathematically. The protein binding partners associate at a rate ka and dissociate at a rate kd as illustrated below in Figure 20. The ratio (ka/kd) corresponds to the binding affinity KA, and the inverse measure is the dissociation constant KD (= kd/ka). These parameters can be measured experimentally and provide a tool to compare the binding behaviours of mouse WT and mutant p.Y64H r-amelogenins.

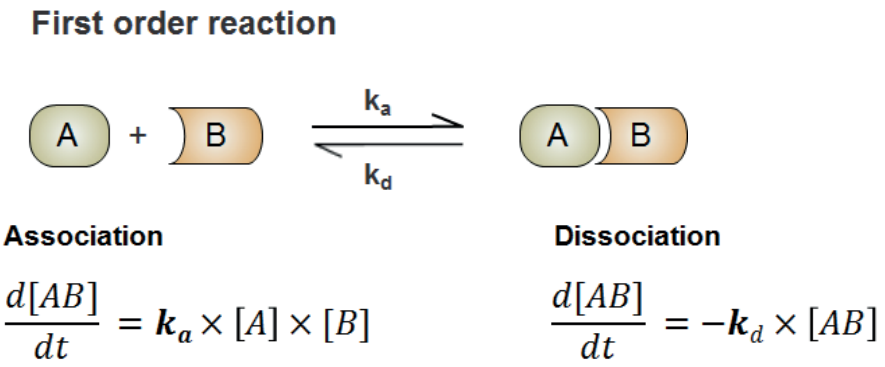


Figure 20 Protein-protein interactions as an equilibrium. In the case of a first order reaction (simplified) the two binding partners associate at rate k_a and dissociate at rate k_d. The affinity or dissociation constants can be measured at equilibrium, where ka.[A].[B] = kd. [AB].

To obtain kinetics parameters that are as reliable as possible, the experimental conditions should be set up so as to reflect *in vivo* environment as accurately as possible. There exists a plethora of protein-protein interaction studies *in vivo/in cellulo* and *in vitro*. The most popular methods are addressed below in Sections 1.3.6.2 – 1.3.6.4 (with further detail on their principles, pros and cons, provided in Appendix E), where the principles of the techniques are described in more detail).

1.3.6.2 Protein-protein interaction assays *in cellulo* or *in vivo*

To date, various in vivo or in cellulo protein-protein interaction studies have been carried out; Y2H assay (Fields and Song, 1989) protein fragment complementation assay (PCA) (Johnsson and Varshavsky, 1994, Hu et al., 2002, Galarneau et al., 2002), fluorescence/bioluminescence resonance energy transfer (FRET/BRET) (Stryer, 1978, Xu et al., 1999) and cross-linking (Wong, 1991). Other techniques including immunofluorescence microscopy, which previously identified amelogenin binding partners (see Section 1.2.5.2, pp. 45 - 46) are not detailed here.

The Y2H assay is commonly used, allowing high-throughput screening for protein-protein interactions, and has contributed to the creation of maps of interaction networks (Ito et al., 2001). It has been used to characterise binding partners and domains for amelogenin (Paine and Snead, 1997). The protein-protein interactions associated with the Y2H system occur in the yeast nucleus, which may not reflect the natural environment in which the protein-protein interactions actually occur (Piehler, 2005). The PCA assays, whose principle is similar to that of Y2H, in that the two proteins of interest, the bait and prey, are each attached to two separate fragments of a reporter protein that are brought together only when the two proteins of interest interact and then able to generate reported signal. The interaction between the 'bait' and 'prey' proteins can occur anywhere in the cell (Johnsson and Varshavsky, 1994).

Other techniques, FRET and BRET, where fluorophores or bioluminescent labels are attached to the proteins of interest, allow direct and real-time detection of protein-protein interactions in cells. The signal is generated actually by the close proximity of the proteins of interest.
While these techniques provide in vivo data and identify protein-protein interactions in real-time, they can be technically demanding and involve protein-labelling or protein fusion, which per se may affect the actual interactions of interest.
Chemical cross-linking studies allow the direct study of protein-protein interactions without any tag or label. It has been used to analyse amelogenin self-assembly in the enamel matrix (Brookes et al., 2000) and in the ameloblast (Brookes et al., 2006). Crosslinking captures protein-protein-interactions and freezes them in time allowing for crosslinked complexes to be identified but precludes the generation of kinetic data. To date, label-free techniques that characterise the kinetic parameters of protein-protein binding are mostly available in vitro. They are addressed in the next Section.

1.3.6.3 Protein-protein interaction assays in vitro

A plethora of in vitro methods exist to characterise protein-protein interactions. Among popular methods, SPR, ITC, or QCM allow direct measurements in real-time of protein binding, generating kinetics parameters such as kd, ka , or affinity/dissociation constants KA, KD.
SPR allows the detection of real-time kinetics of protein binding with a versatile assay format. QCM (Marx, 2003) can analyse protein adsorption and antibody-antigen recognition with similar sensitivity as SPR (Köβlinger et al., 1995) and both methods have been used in parallel in protein-protein interaction reports (Köβlinger et al., 1995). With both SPR and QCM one binding partner is immobilised while the other binding partner is solubilised and binds the immobilised partner from solution. This is a difficulty to consider while studying the binding behaviour of amelogenin which exhibits poor solubility and aggregates at physiological pH and temperature (Tan et al., 1998). Unlike SPR or QCM, ITC avoids immobilisation of either binding partner on a surface as interactions are studied in solution (Brown, 2009). In studying amelogenin binding behaviour, ITC indicated that amelogenin self-assembly, using rP172, is driven by hydrophobic interactions and the formation of a hydrophobic core with water removal (Lakshminarayanan et al., 2009).
AFM and DLS have been widely used to study amelogenin interactions and assemblies (see Sections 1.2.2.2.4 p. 24 and 1.3.3.2.2, p. 63). AFM can be used to visualise the structure of protein complexes (e.g. to visualise selective adsorption of amelogenin nanospheres onto specific faces of enamel crystals (Wallwork et al., 2001) and to directly measure the forces involved in the protein-protein interactions (e.g.

interactions between amelogenin and cell surface receptors (Kirkham et al., 2006) and also to analyse protein folding (Fisher et al., 1999, Yang et al., 2003). DLS can provide high-throughput data, and possibly quantitative analyses of reaction stoichiometry including equilibrium dissociation constants of protein-protein interactions occurring in solution (Hanlon et al., 2010). It has been used to characterise amelogenin self-assembly, as reviewed in Section 1.2.2.2.4 (pp. 24 - 25).

While most of these techniques can generate binding data such as association or dissociation constants or binding strength; they require large amounts of proteins and specialised and costly equipment. As discussed in the next section, microplate-based binding assays in contrast require no specialised equipment, are simple to carry out and are low cost.

In order to compare protein-protein binding behaviours of r-amelogenin WT and p.Y64H mutant, pilot tests were carried out to determine optimum working conditions. They used EMD as surrogate protein model since EMD is available in large quantities, unlike mouse WT and mutant p.Y64H r-amelogenins.

1.3.6.4 Use of microplate-based solid-phase protein-binding assays to study the effect of the p.Y64H mutation on amelogenin self-interactions

In this thesis, microplate binding assays were chosen because they are cost-effective, flexible and easily carried out in house with standard laboratory equipment which facilitates method development.

A high-throughput solid-phase microplate protein binding assay, derived from ELISA methodology, was developed, allowing the characterisation of the interactions between two different proteins, one of which was non-covalently immobilised on the microwells of a microplate, while the other was soluble in solution and able to bind the immobilised protein (Biesiadecki and Jin, 2011). After washing steps, Biesiadecki and Jin (2011) measured the amount of soluble protein bound to its immobilised partner by probing with a specific antibody to the soluble binding partner in conjunction with a chromogenic detection step – as used in standard ELISA assays. This method proved to be cost and labour- effective, providing for flexible conditions and allowing the study of both protein association and disassociation kinetics. However, the aim in this thesis was to characterise the interaction between two identical proteins – the interaction between r-amelogenin molecules. In this case an antibody detection system would be of no use since the antibody would not be able to distinguish between the immobilised r-amelogenin and r-amelogenin bound to the immobilised r-amelogenin.

To get round this problem it was initially decided to measure directly the amount of fluorescently-labelled free amelogenin bound to non-labelled immobilised amelogenin.

Aims and Objectives

This thesis aims at obtaining purified WT and mutant p.Y64H r-amelogenins for *in vitro* binding assays with the aim of investigating whether the p.Y64H point mutation has an effect on amelogenin binding behaviour. Such studies would test the hypothesis that the p.Y64H point mutation causes amelogenesis imperfecta via ER stress triggered by inappropriate aggregation in the secretory pathway.

Aim 1: Production and purification of r-amelogenins

As addressed in Section 1.3.5, the purification of r-amelogenin has often proved challenging using classic chromatographic methods, as amelogenin co-eluted with contaminating host cell proteins (even when His-tagged recombinants were subjected to nickel column chromatography; the gold standard method for purifying recombinant proteins). Contaminating proteins could be problematic for use in downstream *in vitro* studies due to the unpredictable way in which they might modify r-amelogenin interactions. Recently, an extraction method based on amelogenin preferential solubility in acetic acid allowed a significant enrichment of r-amelogenin from *E. coli* which could be potentially useful in purifying r-amelogenin from bacterial host proteins.

Specifically the objectives were to:

1) Optimise the acetic acid extraction method to enrich mouse r-amelogenins in large quantities.
2) Evaluate the use of the acetic acid extraction method as an initial clean-up step for His-tagged r-amelogenins and optimise a range secondary purification technologies: nickel column chromatography based purification, size exclusion chromatography and preparative SDS PAGE.

The importance of obtaining r-amelogenins at maximum purity (e.g. single band on analytical SDS PAGE with silver staining) is that it avoids any interference due to contaminants in functional studies. This is a major consideration for studying the effect of the p.Y64H single point mutation.

Aim 2: Studying the effect of the p.Y64H mutation

As addressed in Section 1.3.4, the p.Y64H mutation in mouse amelogenin was associated with abnormal retention of amelogenin in the ameloblast secretory pathway, causing ER stress and subsequent pro-apoptotic UPR. Further understanding of the underlying mechanisms is required to identify therapeutic targets. A first route of investigation to consider is the effect of the p.Y64H mutation on amelogenin self-assembly and binding behaviour.

Amelogenin protein-protein interactions have been investigated in multiple studies, as described in Sections 1.2.2.2 and 1.3.3.2. With the methods available to analyse protein-protein interactions (see Section 1.3.6), binding assays remain technically challenging due to WT amelogenin's intrinsic propensity to aggregate.

This thesis aims at developing a cost-effective and simple method for studying protein-protein interactions with a view to compare the binding behaviours of WT and p.Y64H mutant r-amelogenins *in* vitro and develop a screening tool for future drug design methodologies.

Specifically the objectives were to:

1) Develop a simple, high throughput and cost-effective microplate-based binding assay, which is flexible enough to use solute conditions similar to those *in* vivo in order to confidently compare with the binding behaviours of WT and mutant p.Y64H r-amelogenins.
2) Attempt to address the question related to AI research, which is whether the p.Y64H mutation affects amelogenin binding behaviour.

The importance of such methodologies is to define working *in vitro* conditions that reflect physiological environment so as to analyse, as accurately as possible, the effect of amelogenin p.Y64H mutation.

Chapter 2 Materials and Methods

2.1 Production and purification of r-amelogenins

2.1.1 General methods

2.1.1.1 Expression of recombinant WT and mutant p.Y64H amelogenin in *E. coli*

Recombinant His-tagged amelogenins either WT (WT^{+His}) or carrying the p.Y64H mutation (Mut^{+His}), were expressed by *E. coli* Rosetta DE3 competent cells previously transfected with vector pET28/*Amelx*-WT^{+His} or *Amelx*-MutY64H^{+His}. The plasmid production and transfection were performed by a commercial company (Novagen, Merck Chemicals Ltd) (Barron et al., 2010, Gabe et al., 2017). Vector pET28 carried the T7 promoter, WT His-tagged amelogenin gene (*Amelx*-WT^{+His}) or mutant His-tagged amelogenin gene (*Amelx*-Mut^{+His}), and chloramphenicol and kanamycin antibiotic resistance genes to allow for the selective growth of transfected cells. The predicted amino acid sequences of the r-amelogenins are shown below in Table 3 (the His-tag cleavage site is underlined and the actual peptide bond cleaved is indicated by a "#" character).

Table 3 Primary sequences or r-amelogenins WT^{+His} and Mut^{+His}. The cleavage site is underlined and indicated by a "#" character. The position of p.Y64H point mutation is indicated in red.

r-amelogenin	Primary sequence
WT^{+His}	MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSL EVLFQ#GPGSMPLPPHPGSPGYINLSYEVLTPLKWYQS MIRQPYPSYGYEPMGGWLHHQIIPVLSQQHPPSHTLQP HHHLPVVPAQQPVAPQQPMMPVPGHHSMTPTQHHQP NIPPSAQQPFQQPFQPQAIPPQSHQPMQPQSPLHPMQ PLAPQPPLPPLFSMQPLSPILPELPLEAWPATDKTKREE VD
Mut^{+His}	MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSL EVLFQ#GPGSMPLPPHPGSPGYINLSYEVLTPLKWYQS MIRQPHPSYGYEPMGGWLHHQIIPVLSQQHPPSHTLQP HHHLPVVPAQQPVAPQQPMMPVPGHHSMTPTQHHQP NIPPSAQQPFQQPFQPQAIPPQSHQPMQPQSPLHPMQ PLAPQPPLPPLFSMQPLSPILPELPLEAWPATDKTKREE VD

The *E. coli* cells were stored on beads at -80 °C until required. Their growth was initiated on the first day of culture by spreading cells on selective medium agar plates and incubating them overnight at 37 °C under conditions of 5% CO_2. The agar medium was made up of 20 g/L LB Lennox Broth (alfa Aesar H26760, Johnson Matthey Co., Heysham, UK) in distilled water mixed with 12 g/L Technical Agar No. 3 (Thermo Fisher Scientific Oxoid Ltd, Basingstoke, UK). The mixture was autoclaved at 121 °C for 15-20 minutes, before the addition of kanamycin A (LKT Laboratories, St. Paul, MN) at the concentration of 61.2 µM to favour the growth of transfected cells carrying the genes that provide resistance to this antibiotic.

The resulting colonies were picked under sterile conditions and used to inoculate 8 mL selective liquid growth medium. The liquid growth medium was comprised of 20 g/L LB/Broth (alfa Aesar H26760, Johnson Matthey Co., Heysham, UK) in distilled water and had been sterilised by autoclaving at 121°C for 15-20 minutes. Before each inoculation, the medium was brought to 37 °C and kanamycin A (LKT Laboratories, St. Paul, MN) and chloramphenicol (Sigma-Aldrich, St Louis, MO) were added at concentrations of 61.2 µM and 76.4 µM respectively. The inoculated tubes were incubated at 37°C (ISF-1-W Kuhner AG, Basel, Switzerland) with shaking at 200 rpm. The next day the bacterial culture was diluted 100 times in fresh liquid growth medium and bacterial growth monitored by spectrophotometry at a wavelength of 600 nm (OD_{600nm}) using a Jenway 6305 spectrophotometer (Bibby Scientific Limited, Stone, UK) until the optical density reached a value comprised between 0.6 and 1.2 AU. After the OD_{600nm} reached a value between 0.6 and 1.2 AU, isopropyl 1-thio-β-D galactopyranoside (IPTG) (Generon Ltd, Maidenhead, UK) was added to the medium at a concentration of 1 mM to induce the expression of r-amelogenin. Following the addition of IPTG, the culture flasks were left overnight on the shaking incubator at 37 °C. The following day the mixture was centrifuged at 15,008 g for 10 minutes at 4°C (Avanti J26-XP, Beckman Coulter). Supernatants were discarded and cellular pellets were retained and stored at -20 °C; due to their physical aspect, the cellular pellets harvested after *E. coli* culture and expression of r-amelogenin will be referred to as “paste” in the thesis. To monitor the expression of r-amelogenin in the *E. coli*, a 100 µL aliquot was taken from the bacterial growth flask before induction with IPTG and kept at -20°C. After addition of IPTG to the remaining culture and subsequent overnight incubation, a 100 µL aliquot was taken. The protein contents of *E.coli* in both pre- and post-induction aliquots were analysed by analytical SDS PAGE and western blot. The methods for analytical SDS PAGE and western blotting are described below in Section 2.1.1.4 (pp 100 - 103). The preparation and loading of samples for analytical SDS PAGE and western blot analyses is detailed in the supplementary data (CD).

2.1.1.2 Optimisation of 3% acetic acid amelogenin extraction procedure

The first step in isolating r-amelogenin from the bacterial paste (obtained from the cultures described above) from the plethora of contaminating bacterial proteins involved extraction in 3% acetic acid. This method, based on the apparent preferential solubility of amelogenin at acidic pH, was first published using non-His-tagged human r-amelogenin (Svensson Bonde and Bulow, 2012). Svensson Bonde and Bulow (2012) washed the *E. coli* cells in 150 mM NaCl and resuspended them in 0.5 – 5% acetic acid. The cell suspension was heated to 60-80°C in a water bath for 20 minutes and then subjected to centrifugation at 20,000 g for 20 minutes. The authors reported that higher concentrations of acetic acid gave optimum results and that at 3% acetic acid, within 'weight of *E. coli* to volume of acetic acid' ratios of 0.035 – 0.35 g of *E. coli* paste per mL of acetic acid, the higher volume of acetic acid provided a higher yield but they did not elaborate further in terms of the optimum volume of acid to use for a given mass of cells (Svensson Bonde and Bulow, 2012).

Based on this data, WT^{+His} and Mut^{+His} mouse r-amelogenins were extracted from *E. coli* in 3 % acetic acid. However, due to the presence of the His-tag (absent in the original work reported by Svensson Bonde and Bulow (2012)) and the fact that mouse and human amelogenins are not 100 % conserved, the volume of 3 % acetic acid used to extract r-amelogenin from a given weight of *E. coli* cells and the cell lysis conditions for mixing and heating were optimised for this study.

2.1.1.2.1 Determining the volume of 3 % acetic acid to *E. coli* to give optimum extraction efficiency

To identify the ratio of *E. coli* to volume of acetic acid required to give the optimum yield of r-amelogenin in a single extraction step without over dilution, the extraction of mouse r-amelogenin in 3 % acetic acid was tested in ranges of 0.0033 - 0.33 g/mL of *E. coli* paste (obtained as described in Section 2.1.1.1 pp. 92 - 93) in acetic acid. Cells expressing WT^{+His} r-amelogenin were washed to remove extracellular proteins by resuspension in 150 mM NaCl at 0.033 g/mL *E. coli* paste in NaCl solution and centrifuged at 3220 g for 20 minutes at room temperature. After centrifugation, the supernatant was discarded and the pelleted cells were resuspended in 3 % acetic acid at either 0.33, 0.17, 0.033 or 0.0033 g/mL by manual mixing and then heated at 75°C for 20 minutes in a water bath and centrifuged at 3220 g for 15 minutes. The protein content in the supernatant was analysed by SDS PAGE (see Section 2.1.1.4, pp. 100 - 102). The preparation of the samples for analytical SDS PAGE is detailed in the supplementary data (CD).

2.1.1.2.2 Optimisation of the mixing regimen and extraction temperature on the yield of r-amelogenin extracted from *E. coli*

Svensson Bonde and Bulow (2012) extracted r-amelogenin by subjecting the cells (resuspended in acetic acid) to either sonication (8x10 seconds), followed by heat treatment or simply by a 20-minute heat treatment at 60-80 °C in a water bath. They obtained r-amelogenin at a greater purity by using direct heat treatment rather than sonication. However, it was not clear what form of sonication was used; was it relatively mild using a sonic water bath or more extreme using an ultrasonic probe designed specifically for tissue homogenisation?

For *E. coli* expressing Mut^{+His}, resuspending the cells in acetic acid manually was inefficient, being both time and effort-consuming. The extraction regimen was therefore optimised in an attempt to make it quicker and easier and to improve the yield. An ultrasonic probe was trialled as an alternative technique that simultaneously resuspended the cell pellet and ruptured the cells.

The optimisation of the mixing regimen including use of an ultrasonic probe is illustrated in Figure 21A p.92. After resuspension in 150 mM NaCl, the *E. coli* cells expressing Mut^{+His} were pelleted by centrifugation 3220 g for 20 minutes. The pellets were crudely resuspended by hand in 3 % acetic acid at 0.033 g/mL cells in acetic acid. The mixture was sonicated for 30 - 60 seconds, using an ultrasonic liquid processor (Ultrasonix processor XL sonicator, Misonix) operated at maximum permissible power (power setting 4 (arbitrary units)) followed by heating at 75°C in a water bath for 20 minutes. The mixture was then allowed to cool to room temperature,

followed by centrifugation at 3220 g for 20 minutes. The protein content of the mixture was analysed by SDS PAGE with Coomassie Blue staining (see Section 2.1.1.4, pp. 100 - 102). The preparation of the samples for analytical SDS PAGE is detailed in the supplementary data (CD).

To investigate the efficacy of the heating step on yield, the extraction procedure was also carried out at room temperature (as illustrated in Figure 21B below). The contents of the supernatant and pellet fractions were analysed by analytical SDS PAGE with Coomassie Blue staining (see Section 2.1.1.4, pp. 100 - 102). The preparation of the samples for analytical SDS PAGE is detailed in the supplementary data (CD). Gel densitometry was used to compare the extraction yield with and without heating.

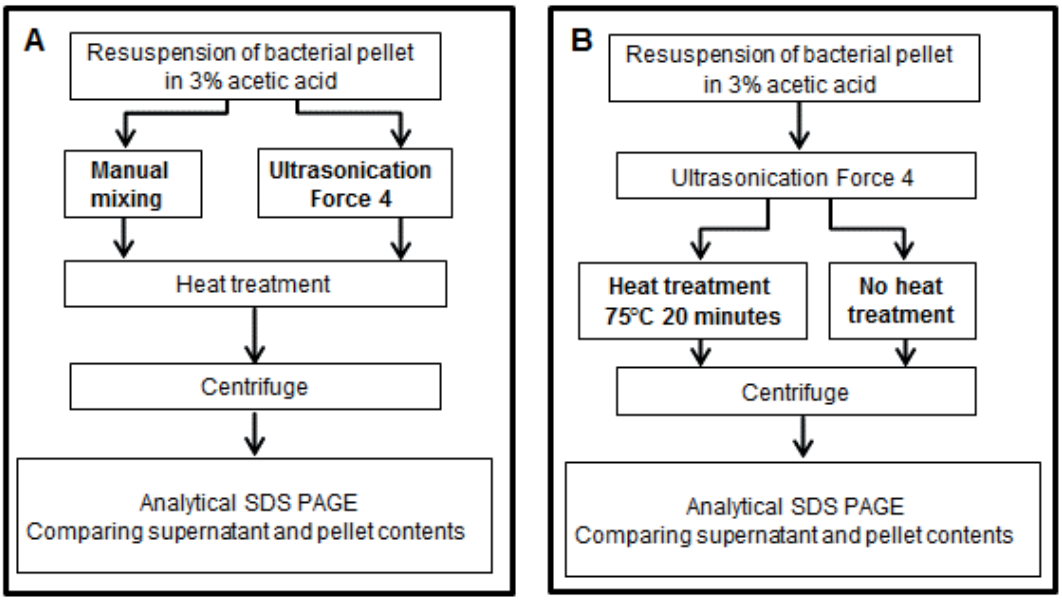


Figure 21 Panel (A) summarises optimisation of the mixing regimen used during the acetic acid extraction of r-amelogenin. The effect of ultrasonication and manual mixing on the resulting acetic acid extract were investigated by comparing the protein contents of the final supernatant and pellet by analytical SDS PAGE. Panel (B) describes the procedure carried out to investigate the effect of the heat treatment on the extraction yield. The effect of heating treatment was analysed by comparing the contents of final supernatants and pellet by analytical SDS PAGE.

2.1.1.2.3 Large scale acetic acid extraction of r-amelogenin using the optimised extraction methodology

Following the optimisation described above, it was decided that an extraction regimen that included ultrasonication and heating to 75°C gave the best results (see Section 3.1.1.2.2 pp. 140 - 141). The finalised extraction procedure of r-amelogenin in 3 % acetic acid was scaled up and is summarised in Figure 22. The cells harvested after induction of protein expression were washed in 150 mM NaCl, at a ratio of 0.033 g/mL *E. coli* paste in NaCl solution, then centrifuged 3220 g for 20 minutes. The supernatant was discarded and the cells were resuspended to 0.033 g/mL wet weight cells in 3 % acetic acid, then sonicated until the mixture appeared homogeneous and heated for 20-25 minutes at 75°C in a water bath. After removal from the water bath, the lysate was left to cool to room temperature and was then centrifuged again as above. The supernatant was frozen at -80°C, lyophilised and desalted (see Section 2.1.1.3, p. 99 for desalting procedure). The desalted extracts were subjected to mass spectrometry analyses (see Section 2.1.1.5, p. 104, for mass spectrometry analyses).

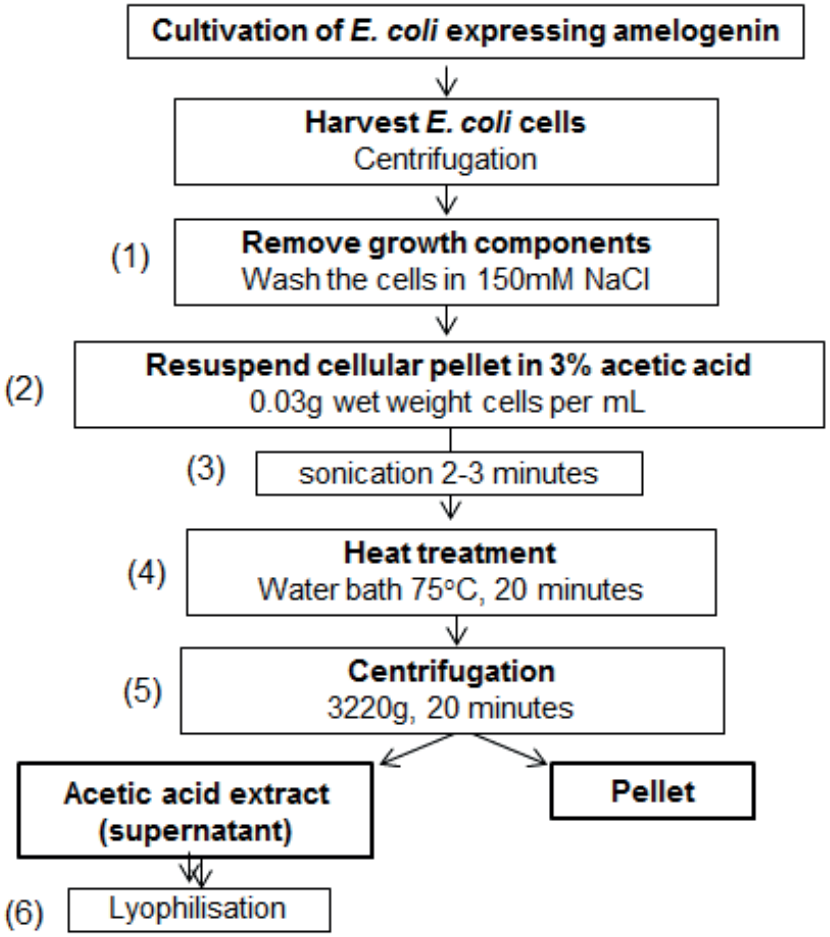


Figure 22 Flow-diagram of the optimised extraction procedure. (1) *E. coli* were washed in 150 mM NaCl and (2) resuspended in 3 % acetic acid at 0.033 g cell paste per mL acetic acid. The suspension was (3) mixed by ultrasonication, and (4) heated at 75°C for 20 minutes. (5) The fraction of proteins solubilised was isolated in supernatant by centrifugation at 3220 g for 20 minutes and (6) was lyophilised.

2.1.1.3 Desalting procedure

Samples containing r-amelogenins were desalted against degassed 125 mM formic acid by size exclusion chromatography (SEC). The larger r-amelogenin molecules are excluded from the size exclusion beads and washed through the column quickly in the mobile phase. Smaller molecules (ions, salts, etc.) enter the beads and their passage through the column is retarded, meaning they elute after the protein. Degassed formic acid at 125 mM was prepared by dissolving formic acid (cat. no. 33015, Sigma-Aldrich, St Louis, MO) to 125 mM in ultrapure water[1] and then degassing. It was used as the mobile phase because it provides the hypothetical ionic strength to reduce undesirable interactions between the sample proteins and the gel filtration medium. The pKa formic acid is 3.75 (Riddick et al., 1986), so the pH of 125 mM formic acid is low enough (2.34) to solubilise r-amelogenin (Tan et al., 1998) without causing long term damage to the column medium. In addition, formic acid is volatile and can be easily removed from the desalted protein by lyophilisation.

The lyophilised samples containing amelogenin (either r-amelogenin or EMD components- see later, detailed in Section 2.2.1 p. 120) were dissolved in a minimum volume of degassed 125 mM formic acid. Any precipitate/insoluble contaminants were removed by filtering the mixture through a 0.45 µm filter. Then, the solution was loaded onto the HiPrep 26/10 column (GE Healthcare, Buckinghamshire, UK) previously equilibrated with degassed 125 mM formic acid. The elution was carried out at room temperature and a flow rate of 3.5 mL/minute using an AKTA FPLC chromatography system (GE Healthcare, Buckinghamshire, UK). The column eluate was monitored at 280 nm and the desalted protein, eluting in the first peak, was frozen at -80 °C, after which the volatile formic acid was removed by lyophilisation.

Some procedures required desalting of r-amelogenin that had been previously lyophilised from 0.1 M Na_2CO_3 (see Sections 2.1.2.3.2 p. 118 -119 and 2.2.1.1 p. 120). Because of the presence of Na_2CO_3 in the lyophillisate, the addition of 125 mM formic acid failed to acidify the solution sufficiently to dissolve the r-amelogenin without massively over-diluting the sample. Instead, increasing volumes of 100 % formic acid were added to the mixture until it appeared clear. It was then subjected to filtration and subsequently desalted as described above.

[1] In the thesis, ultrapure water (18.2 MΩ.cm) was supplied by Purelab Option-Q system (ELGA LabWater, Buckinghamshire, UK).

2.1.1.4 SDS PAGE analyses and western blotting

SDS PAGE analyses were based on the method of Laemmli, 1970 (Laemmli, 1970) except that the resolving gel also contained 20 % (v/v) glycerol, to increase its density so that the stacking gel could be cast directly on top of the resolving gel without having to wait for it to polymerise. Proteins were resolved "using a 12 % resolving gel at pH 8.8 and a 4 % stacking gel at pH 6.8" (Gabe et al., 2017). Gels were of 1.0 mm thickness, cast using a Mini-PROTEAN III electrophoresis system with 15 sample wells (Bio-Rad Laboratories Ltd., Hertfordshire, UK). The composition of the gels and buffers is detailed in Table 4[2]. The reagents for SDS PAGE were supplied by National Diagnostics (National Diagnostic, Nottingham, UK). Gels were polymerised using ammonium persulphate solution (stock 10%) and N,N,N',N'-tetramethylethylenediamine (TEMED) (Sigma-Aldrich, St Louis, MO) immediately prior to casting. After the gels were cast and polymerised, samples containing the proteins to be separated were loaded in to the sample wells. The preparation of samples to load on SDS PAGE was carried out in one of two ways depending upon whether the samples were in solution or lyophilisates:

- Protein in solution was prepared by mixing with concentrated 4 times (4X) Laemmli non-reducing sample buffer (composition detailed in Table 4) at a ratio of 3:1 (v:v) sample volume: sample buffer. The mixture was heated for 1.5 minutes at 90-100°C, cooled to room temperature and then loaded on the gel.
- Lyophilised samples were resuspended at their original concentration in concentrated 1 time (1X) Laemmli non-reducing buffer (composition detailed in Table 4). The mixture was heated for 1.5 minutes at 90-100°C, cooled to room temperature and then loaded on the gel.

The sample preparation for each case is specified in the supplementary data (CD). Loading for each gel was optimised in each case (further details about the volumes loaded are provided in the supplementary data (CD)). Electrophoresis was carried out at a constant 200 V until the bromophenol blue tracker dye had reached the bottom of the gel. After electrophoresis, the proteins were detected using Coomassie Blue (Expedeon Ltd., Cambridgeshire, UK) or a silver staining (kit # 24612, ThermoFischer Scientific, Leicestershire, UK) or by UV illumination in ultrapure water where fluorescence detection was required.

For molecular weight analysis and characterisation of the bands observed on analytical SDS PAGE, molecular weight markers were ran on the analytical SDS PAGE in separate lanes. The markers used were either "Precision Plus Protein All

[2] The composition of the gels and buffers, detailed in Table 4, was the same for analytical and preparative SDS PAGE (the latter, detailed later in Section 2.1.2.3)

Blue Standards" (cat no. 161-0373, Bio-Rad Laboratories Ltd., Hertfordshire, UK) or "Prestained SDS PAGE Standards, Broad Range" (cat no. 161-0318, Bio-Rad Laboratories Ltd., Hertfordshire, UK).

Western Blotting

To confirm the presence and identity of r-amelogenin in the extracts, the proteins were electroblotted onto 0.2 µm nitrocellulose membranes for western blotting analysis (Bio-Rad Laboratories Ltd., Hertfordshire, UK) during which the membranes were probed with antibodies specific for amelogenin. The proteins resolved on the gels were transferred onto the membrane by electroblotting for 1 hour at 60 V using the 'sandwich' procedure (Towbin et al., 1979). The membrane was placed in blocking buffer for at least 1 hour at room temperature or overnight at 4°C in order to block non-specific binding of antibodies to the membrane. The blocking buffer was comprised of 5 % (m/v) non-fat dry milk in Tris buffered saline (TBS). After blocking, the membrane was washed 2 x 7 minutes in TBS containing 0.05 % Tween 20 (TTBS). Then the membrane was incubated for 1 - 2 hours with affinity purified rabbit IgGs raised against a peptide corresponding to the amelogenin hydrophilic C-terminal (telopeptide) (custom made by Eurogentec, Southampton, UK) diluted in blocking buffer or TTBS (see Supplementary data (CD) for details on dilutions and buffers, as these are specific of the western blot analysis). It was then washed 2 x 7 minutes in TTBS before incubation with goat anti-rabbit secondary antibodies conjugated to peroxidase (Sigma-Aldrich, St Louis, MO) diluted in blocking buffer or TTBS (see Supplementary data (CD) for details on dilutions and buffers, as these are specific of the western blot analysis). This was followed by 5 x 7 minute washes with TTBS, and the cross reactivity was visualised by developing the blot with metal-enhanced 3,3'-diaminobenzidine tetrahydrochloride (Sigma-Aldrich, Dorsetshire, UK) according to the manufacturer's instructions. The membrane was rinsed with distilled water as soon as bands appeared; development took between a few seconds to 7 minutes (the duration is specified for each analysis in the supplementary data (CD)). Then, the membrane was dried overnight between weighted paper towels.

The reagents used in western blotting are listed in Table 5. Since all western blotting procedures carried out in this work used the same antibodies (rabbit anti-telopeptide IgGs and goat-anti-rabbit IgG peroxidase conjugates), whose specificity for r-amelogenin was confirmed by controls, western blotting in this thesis will be referred to as 'anti-amelogenin western-blotting'.

Table 4 List of reagents used for SDS-PAGE. This table identifies the reagents (with manufacturers, for specific products) used in stacking and resolving gels, sample buffer and running buffer. All concentrations of reagents listed are written as percentages or concentrations in ultrapure water.

Buffer Solution Composition	Component Reagents
Stacking gel 4 % acrylamide; 0.125 M Tris-HCl; 0.1 % (w/v) SDS; 0.05 % (w/v) ammonium persulphate and 0.1 % (v/v) TEMED, pH=6.8	13 % (v/v) stock solution polyacrylamide (EC-890 National Diagnostics, Nottingham, UK) 25 % (v/v) stock solution stacking buffer (EC-893 National Diagnostics, Nottingham, UK) 0.05 % (w/v) ammonium persulphate (AGTC Bioproducts, Hessle, UK), 0.1 % (v/v) TEMED (EC-503, National Diagnostics, Nottingham, UK).
Resolving gel 12 % acrylamide, 0.39 M Tris-HCl; 0.1 % (w/v) SDS; 20 % (v/v) glycerol; 0.05 % (w/v) ammonium persulphate and 0.1 % (v/v) TEMED, pH=8.8	40 % (v/v) stock solution polyacrylamide (EC-890 National Diagnostics, Nottingham, UK) 25 % (v/v) stock solution resolving buffer (EC-892 National Diagnostics, Nottingham, UK) 20 % v/v glycerol 0.05 % (w/v) ammonium persulphate (AGTC Bioproducts, Hessle, UK), 0.1 % (v/v) TEMED (EC-503, National Diagnostics, Nottingham, UK)
Sample buffer (4X stock) 104 mM Tris-HCl; 10 % (v/v) glycerol, 1.42 % (w/v) SDS, 0.04 % (w/v) bromophenol blue)	52 % (v/v) stock solution stacking buffer (EC-893 National Diagnostics, Nottingham, UK) 41.7 % (v/v) glycerol 0.04 % (w/v) bromophenol blue (Fischer Scientific, Loughborough, UK) 1.25 % (w/v) SDS
Running buffer 25 mM Tris base; 0.192 mM glycine, 0.1 % (w/w) SDS	10 % (v/v) stock Tris-Glycine-SDS PAGE Buffer (EC-870 National Diagnostics, Nottingham, UK)

Table 5 List of reagents used for western blotting. This table identifies the reagents (with manufacturers, for specific products) used. All concentrations of reagents listed are written as percentages or concentrations in pure water.

Buffer Solution / consumable	**Reagents**
Membrane	• Nitrocellulose membrane 0.2 µm (Bio-Rad Laboratories Ltd., Hertfordshire, UK)
Transfer buffer 20 % methanol, 25 mM Tris, 192 mM glycine	20 % methanol, 25 mM Tris, 192 mM glycine
TBS 20 m M Tris, 500 mM NaCl, pH 7.4	10 % (v/v) stock Tris buffered saline (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK)
Blocking solution TBS, 5 % (w/v) non-fat dry milk	10 % (v/v) stock Tris buffered saline (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK) 5 % (m/v) non-fat dry milk (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK)
TTBS TBS+ 0.05 % (v/v) Tween	10 % (v/v) stock Tris buffered saline (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK) 0.05 % (v/v)Tween 20 (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK)
Antibodies and development	• **Primary antibodies** *(diluted in TBS or in Blocking solution)* - Rabbit anti-telopeptide IgG (custom made by Eurogentec, Southampton, UK) stock solution concentrated 2.3 mg/mL • **Secondary antibodies** (*diluted in TBS or in Blocking solution)* - Goat anti-rabbit IgG peroxidase conjugate (cat no.A0545) (Sigma-Aldrich, St Louis, MO) stock solution concentrated 4.0 - 11.0 mg/mL • **Development:** SIGMAFAST™ DAB (3,3'-diaminobenzidine tetrahydrochloride) and metal enhancer tablet, kit D0426 (Sigma-Aldrich, St Louis, MO)

2.1.1.5 Mass spectrometry analyses

After desalting and lyophilisation, samples (obtained after acetic acid extraction, see Figure 22 p. 98 or after preparative SDS PAGE, see Figure 26 p.117) were characterised by mass spectrometry, a technique that separates molecules based on their charge-to-mass ratio and allows for the accurate determination of molecular mass of each species present. An accurate mass determination helps confirm the identity of proteins present in the samples as well as their degree of purity.

Electron spray mass spectroscopy (TOF MS ES+) was performed by the mass spectrometry facility (Astbury Centre for Structural Molecular Biology, Astbury Building, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK).

2.1.2 Strategies for secondary purification to obtain purified r-amelogenin for binding studies

Following the process optimisations carried out as described above, the extraction procedure of r-amelogenins in 3% acetic acid was scaled up as summarised in Figure 22 p. 98. As shown in the Results chapter (Sections 3.1.1.2.1-2, pp. 138 - 141) this process resulted in a significant enrichment of r-amelogenin. However, acid-soluble contaminating bacterial proteins were still visible between 10 – 75kDa on SDS PAGE with Coomassie Blue staining. Therefore, it was deemed necessary to carry out secondary purification to isolate r-amelogenin at single band purity on SDS PAGE for future use in binding assays. A panel of standard protein purification techniques were tested and eventually optimised (Sections 2.1.2.1, 2.1.2.2, 2.1.2.3). The range of purification strategies explored is summarised below in Figure 23.

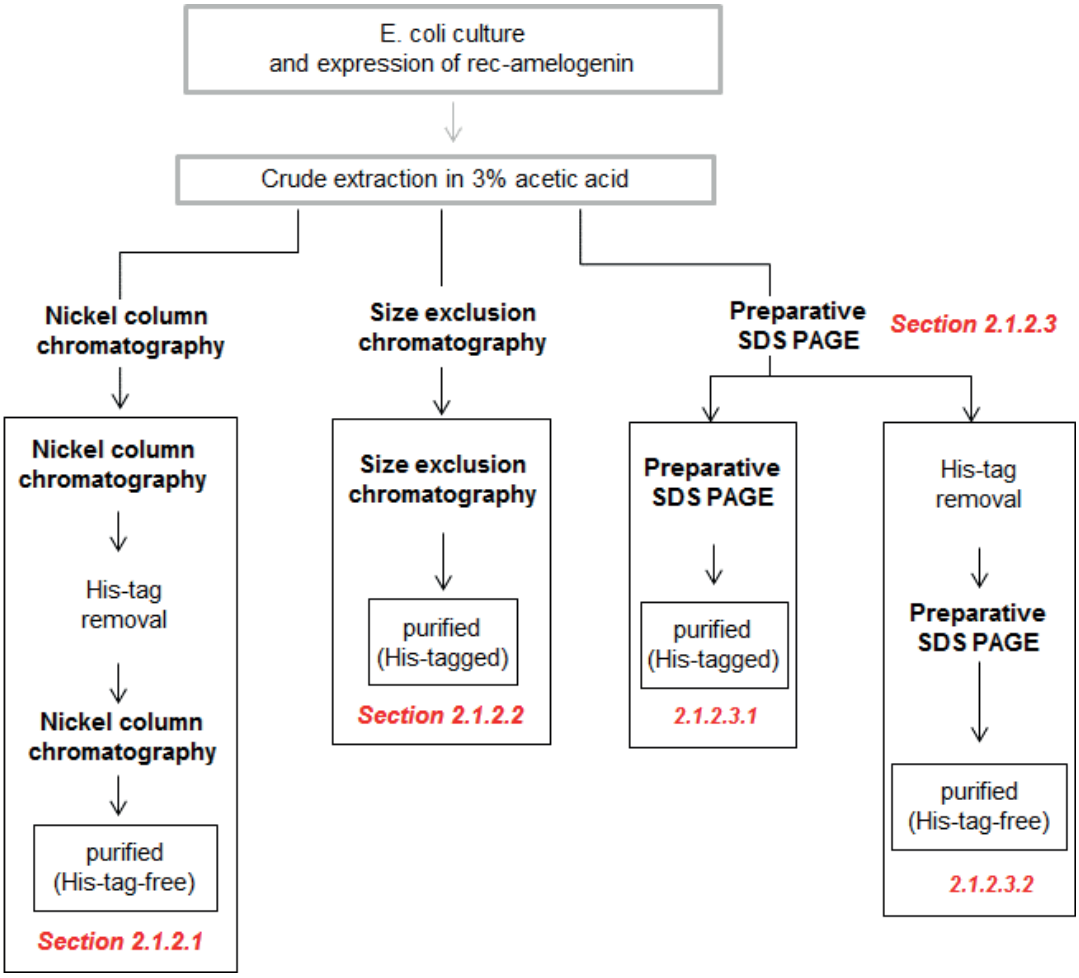


Figure 23 Secondary purification of r-amelogenin from 3% acetic acid extracts. Summary of the strategy developed to determine the optimum method providing with r-amelogenin at a suitable purity for future binding studies.

2.1.2.1 Purification of r-amelogenin using nickel column chromatography

His-tag affinity chromatography using nickel affinity chromatography column (referred to as “nickel column” chromatography) is a purification method commonly used to isolate a protein of interest from a whole protein mixture (Hochuli et al., 1988). Recombinant proteins (including those described earlier in this thesis) are frequently initially engineered to carry six consecutive histidine residues present in an N-terminal tag (see Table 3 p. 92) which have a high affinity for nickel. In theory, when a mixture of proteins is pumped through the nickel column in loading buffer, proteins without the His-tag are not retained and pass through the column, leaving the His-tagged protein of interest bound to the column. After the non-His-tagged contaminants have washed through the column, the His-tagged protein can then be eluted from the column by increasing the concentration of imidazole in the buffer, as imidazole has a high affinity for nickel and competitively displaces the His-tagged recombinant protein. Once the protein is collected and the elution buffer salts removed, the recombinant protein can either be used directly in functionality studies, or undergo further steps of preparation that include removal of the His-tag with the aim of obtaining a recombinant protein as close as possible to the native form. The His-tag can be removed enzymatically by a protease and the cleavage mixture is subject to a second round of nickel column chromatography to isolate the His-tag-free recombinant protein which, now unable to bind, washes straight through the column whilst uncleaved recombinant protein and the cleaved His-tag are retained.

Nickel column chromatography was used in this thesis to attempt to purify His-tagged r-amelogenins before and after cleavage of the His-tag as summarised above. This preparation process included desalting and buffer exchange steps. The equipment used is listed in Table 6 below. The procedure is summarised below in Figure 24, with specific methodological details provided in the Sections 2.1.2.1.1 – 2.1.2.1.3 below..

Table 6 List of reagents used in nickel column chromatography purification. All buffer solutions were made with ultrapure water and degassed.

Nickel column chromatography (Both rounds) Sections 2.1.2.1.1 and 2.1.2.1.3	**Chromatography equipment** • Ni-NTA column: HisTrap FF 5mL(GE healthcare bioscience, Buckinghamshire, UK) used in conjunction with AKTA FPLC system (GE healthcare bioscience, Buckinghamshire, UK) **Nickel column chromatography buffer** Loading buffer: 4 M urea, 50 mM Tris, 400 mM NaCl, 20 mM imidazole, HCl, to pH = 7.6 Elution buffer: Loading buffer plus imidazole (60 - 200 mM)
His-tag cleavage Section 2.1.2.1.2	• Recombinant type 14 3C protease from human rhinovirus HRV 3C (Merck Millipore, Darmstadt, Germany) • Buffer: 50 mM Tris, HCl to pH = 8
Buffer exchange Included in Section 2.1.2.1.2	**Chromatography equipment** • HiPrep 26/10 column (GE healthcare bioscience, Buckinghamshire, UK) **Buffer** • 4 M urea, 50 mM Tris, 400 mM NaCl**,** 20 mM imidazole, HCl to pH = 7.6

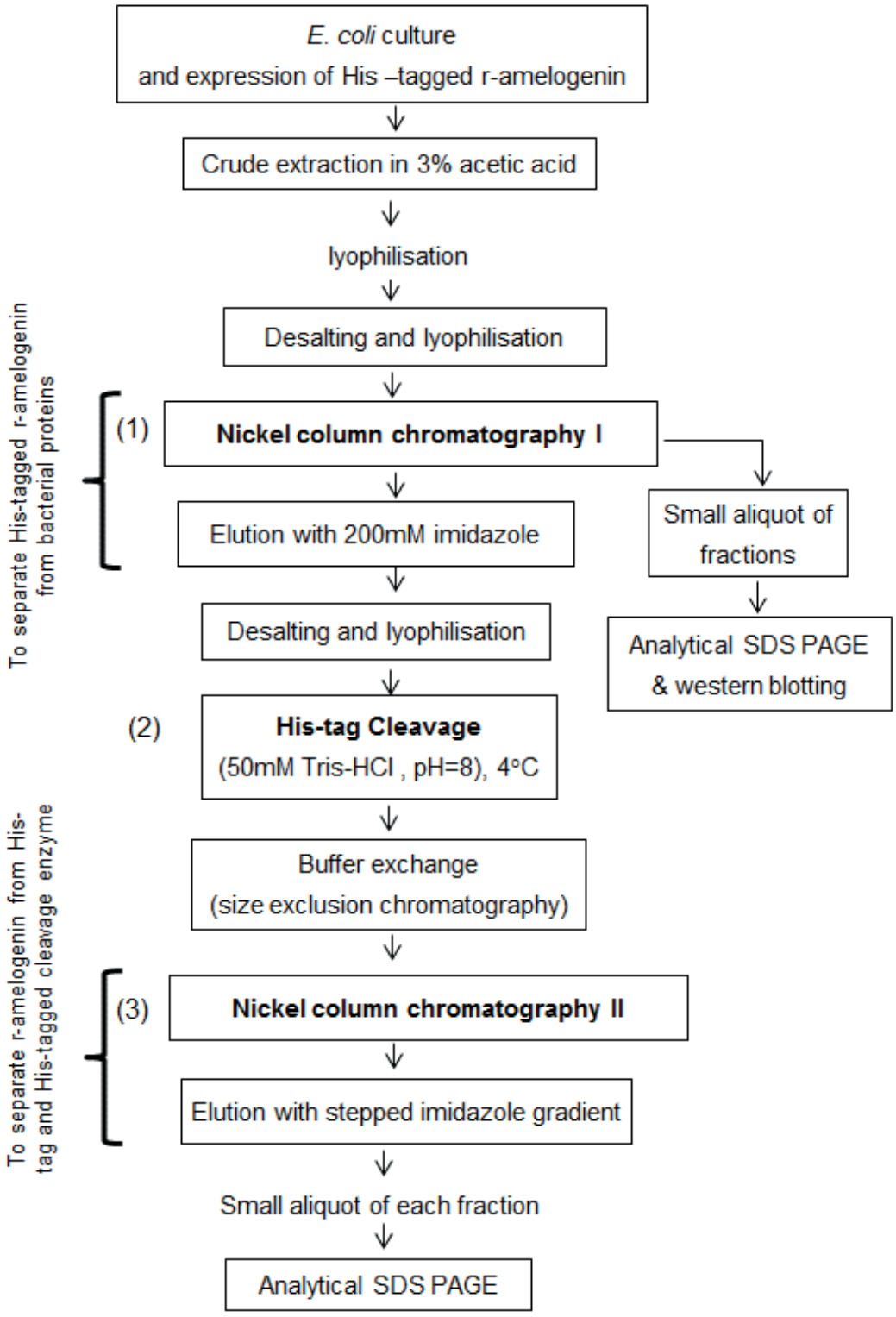


Figure 24 Overview summary of purification of r-amelogenin using two rounds of nickel column chromatography. The details of steps (1), (2) and (3) are described respectively in the following Sections (2.1.2.1.1, 2.1.2.1.2, 2.1.2.1.3).

2.1.2.1.1 Step 1: First round of nickel column purification of His-tagged r-amelogenins extracted in acetic acid

Following the method used by Gabe et al (2017), desalted and lyophilised crude extracts of r-amelogenin WT “were dissolved in a minimum volume of nickel column binding buffer (20 mM imidazole, 4 M urea, 50 mM Tris-HCl, 400 mM NaCl, pH = 7.6), loaded on to a nickel column up to a maximum protein load of 40 mg and eluted at a flow rate of 4 mL/minute. Unbound proteins (column flow through) were collected. Bound r-amelogenin was eluted by increasing the imidazole concentration to 200 mM”. “Fractions were analysed by analytical SDS PAGE and anti-amelogenin western blotting” (Gabe et al., 2017). SDS PAGE procedure is described in Section 2.1.1.4 (pp. 100 - 103). Details for sample preparation, loading and blotting conditions are provided in the supplementary data (CD). The fractions comprising r-amelogenin were desalted by SEC against 125 mM formic acid, as described in Section 2.1.1.3 (p. 99) and lyophilised.

2.1.2.1.2 Step 2: His-tag cleavage

Similar to what was done by Gabe et al (2017) the lyophilised r-amelogenin obtained from the first round of nickel column purification (see previous Section), “was dissolved at 2 mg/mL in 50 mM Tris-HCl, pH = 8 […]. The His-tag was removed enzymatically by adding recombinant restriction grade protease HRV3C (Merck Millipore, Watford, UK) at a ratio of 3 µL enzyme solution per mg recombinant protein”. The cleavage was performed over 24 hours at 4°C and the protein contents of the solutions were analysed by analytical SDS PAGE (see Section 2.1.1.4 , pp. 100 - 102, for the procedure). The sample preparation and loading is described in the supplementary data (CD). “The resulting cleavage reaction mixture comprising cleaved r-amelogenin, any uncleaved r-amelogenin, free His-tag and other contaminants was finally buffer-exchanged into nickel column binding buffer by SEC (HiPrep 26/10 column, GE Healthcare) ready for a second round of nickel column chromatography” (Gabe et al., 2017).

2.1.2.1.3 Step 3: Second round of nickel column purification to remove uncleaved r-amelogenins, free His-tag, and cleavage enzyme

Similar to what was done by Gabe et al (2017), “a second round of nickel column chromatography was employed to remove cleaved His-tag and other contaminants present in the cleavage reaction mixture including any uncleaved recombinant. Cleavage products in binding buffer were loaded on to the nickel column” as described in Section 2.1.2.1.1. “Unbound proteins (expected to contain the cleaved r-amelogenin) were collected in the column flow through. Bound proteins (expected to

contain any uncleaved r-amelogenin, His-tag, and His-tagged HRV3C) were eluted using a stepped imidazole gradient" (Gabe et al., 2017). The imidazole concentrations were 60 and 200 mM. Fractions were analysed by analytical SDS PAGE (see Section 2.1.1.4, pp. 100 - 102 and Supplementary data (CD)).

2.1.2.2 Purification of r-amelogenin using size exclusion chromatography

Size exclusion chromatography (SEC) allows for separation of proteins within a mixture according to their apparent molecular weights. The protein sample is dissolved in a buffer (mobile phase[1]) which is pumped through a column matrix of porous micro beads. Larger molecules are excluded from the column matrix and quickly pass through the column carried along in the mobile phase and are eluted first. Lower molecular weight components with a smaller hydrodynamic radius can diffuse from the mobile phase into the beads and thus their progress through the column is slower and depends on their rate of diffusion in and out of the beads (Grubisic et al., 1967). Therefore, lower molecular weight protein elutes later.

SEC was explored as a possible purification strategy when seeking to isolate r-amelogenins in the current study. The effect of operational variables such as matrix pore size and column length were investigated on the resolution of separation. The reagents and equipment used are listed in Table 7 below; the development of a size-exclusion-based purification method is summarised below in Figure 25.

Table 7 Equipment used in SEC[2]: column models and matrices.

	Column characteristics	**Matrix**
Bio-gel P-30 (35 cm) Section 2.1.2.2.1	Column XK 16/40 (GE Healthcare bioscience. Buckinghamshire, UK)	Bio-gel P-30, cat. no. 1504154 (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK)
Bio-gel P-10 (35 cm) Section 2.1.2.2.2	Column XK 16/40 (GE Healthcare bioscience. Buckinghamshire, UK)	Bio-gel P-10, cat. no. 1504144 (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK)
Bio-gel P-10 (95 cm) Section 2.1.2.2.3	Column XK 16/100 (GE Healthcare bioscience. Buckinghamshire, UK)	Bio-gel P-10, cat. no. 1504144 (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK)

[1] The mobile phase employed in this thesis was degassed 125 mM formic acid, prepared as described in Section 2.1.1.3 (p .95).

[2] The size exclusion chromatography used the equipment AKTA FPLC system (GE healthcare bioscience, Buckinghamshire, UK).

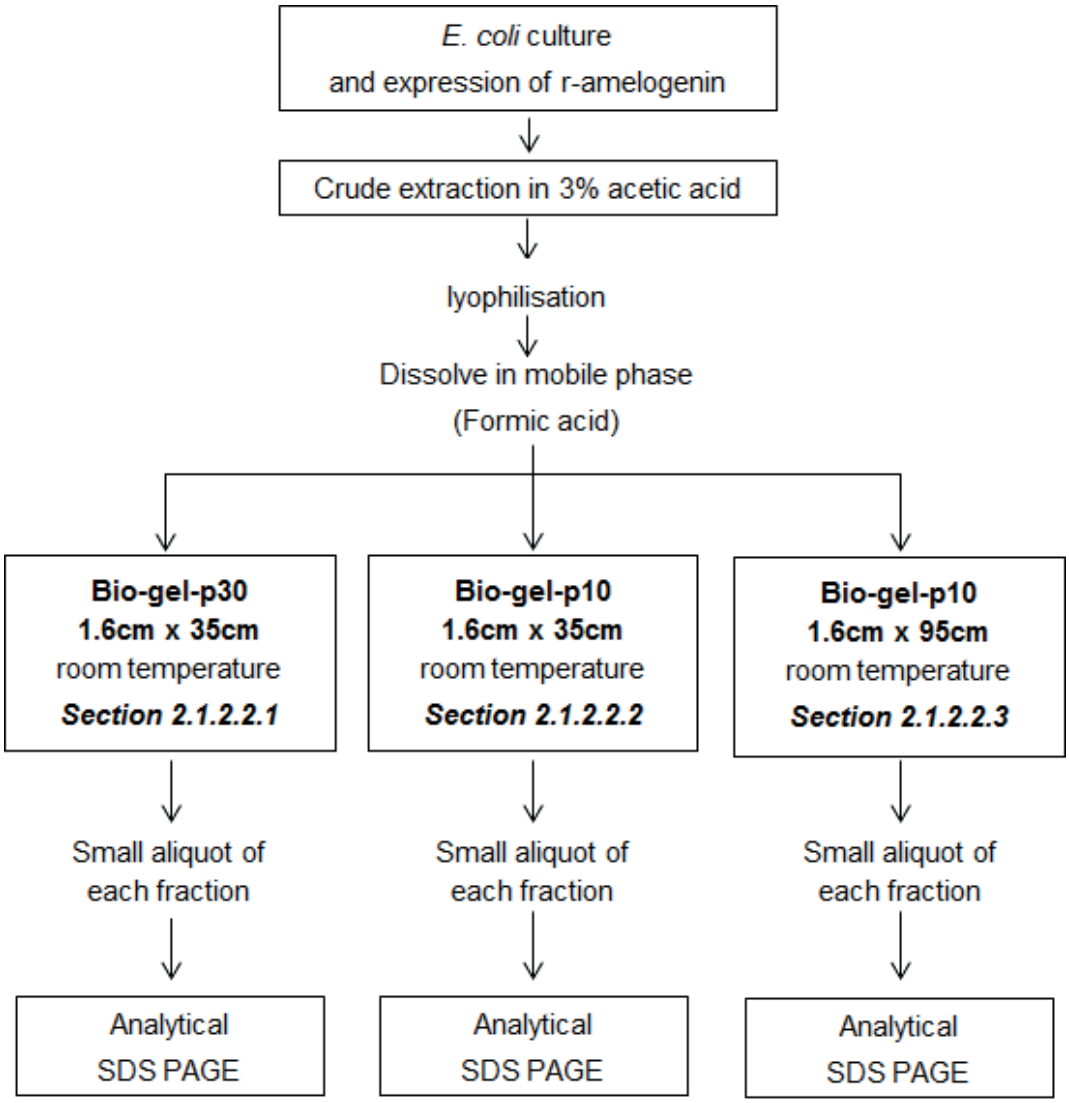


Figure 25 Purification of r-amelogenin using SEC: optimisation of column length and pore size. Once lyophilised, the acetic acid extract containing r-amelogenin was directly subjected to SEC. For a column length of 35 cm, two matrices with different pore sizes were tested: Bio-gel P-30 matrix (detailed in Section 2.1.2.2.1) and Bio-gel P-10 matrix (detailed in Section 2.1.2.2.2). The second column length tested was 95 cm, using Bio-gel P-10 matrix (detailed in Section 2.1.2.2.3). Small aliquots from the fractions collected were taken for SDS PAGE analyses.

2.1.2.2.1 Size exclusion chromatography using Bio-gel P-30 matrix, 35 cm bed height

SEC was carried out on the AKTA FPLC system using 40 x 1.6 cm GE healthcare XK column (GE healthcare bioscience). The column was packed with a polyacrylamide matrix Bio-Gel P-30 (fractionation range of 2.5 - 40 kDa) (Bio-Rad, cat no 150-4154) according to the manufacturer's instructions to a bed height of 35 cm. The sample was prepared by dissolving 15 mg of acetic acid extract of r-amelogenin WT (prepared as in Figure 22 p. 98, lyophilised, not desalted) in degassed 125 mM formic acid to 10 mg/mL and 1.44 mL of the sample were loaded and injected onto the column that had been previously equilibrated with degassed 125 mM formic acid. The separation was carried out at room temperature using a mobile phase comprised of degassed 125 mM formic acid at a flow-rate of 0.25 mL/minute. The elution was carried out using 1.5 column volumes of degassed 125 mM formic acid and 7 fractions were collected manually as indicated on the chromatograph (Figure 46, p. 163). The protein contents of the fractions 1-7 were analysed by analytical SDS PAGE (see Section 2.1.1.4, pp. 100 - 102 and Supplementary data (CD)).

2.1.2.2.2 Size exclusion chromatography using Bio-gel P-10 matrix, 35 cm bed height

Bio-gel is supplied with a range of pore sizes. Altering the pore size changes the exclusion limit of a column, which in turn changes the fractionation range of the column. Bio-gel P-10 has small pores and a fractionation range of 1.5 - 20 kDa. It was thought that this might therefore provide a better fractionation of r-amelogenin than Bio-gel P-30 (fractionation range 2.5 – 40 kDa). The matrix was packed according to the manufacturer's instructions (Bio-Rad, Ltd. Hertfordshire, UK) into a 40 cm long column (XK 16/40) to a bed height of 35 cm. Eleven milligrams of acetic acid extract of r-amelogenin WT (prepared as in Figure 22 p. 98, lyophilised, not desalted) were resuspended in 1.1 mL of degassed 125 mM formic acid and clarified by centrifugation at 20,800 g at 20°C for 10 minutes. The supernatant was collected and 1 mL was loaded and injected on to the column that had been previously equilibrated with degassed 125 mM formic acid. The separation was carried out at room temperature with a flow-rate of 0.35 mL/minute. The elution was carried out using 1.5 column volumes of degassed 125 mM formic acid and 10 fractions of 1 mL were collected. The protein contents of the fractions 1, 2, 3, 4, 10 were analysed by analytical SDS PAGE (see Section 2.1.1.4, pp. 100 - 102 and Supplementary data (CD)).

2.1.2.2.3 Size exclusion chromatography using Bio-gel P-10 matrix, 95 cm bed height

An attempt to further improve the resolving power of Bio-gel P-10 matrix was carried out by increasing the bed height to 95 cm. The matrix Bio-Gel P-10 was packed according to the manufacturer's instructions into a 100-cm-long column (XK 16/100) to a bed height of 95 cm. The sample was prepared by dissolving 11 mg of acetic acid extract of r-amelogenin WT (prepared as in Figure 22 p. 98, lyophilised, not desalted) in degassed 125 mM formic acid to 10 mg/mL. The mixture was then centrifuged at 20,800 g at 20°C for 10 minutes. The supernatant was collected and 1 mL was loaded and injected on to the column that had been previously equilibrated with 125 mM formic acid. The separation was carried out at room temperature with a flow-rate of 0.40 mL/minute. The elution was carried out using 1.5 column volumes of degassed 125 mM formic acid and 10 fractions of 1 mL were collected. The protein contents of the fractions 1-10 were analysed by analytical SDS PAGE (see Section 2.1.1.4, pp. 100 - 102 and Supplementary data (CD)).

2.1.2.3 Purification of r-amelogenins using preparative SDS PAGE

An alternative purification strategy used preparative SDS PAGE to isolate r-amelogenin based upon apparent molecular weight. This was trialled with the aim of:
(1) Isolating His-tagged r-amelogenin from acetic acid extracts prior to cleavage and nickel column chromatography (Section 2.1.2.3.1, summarised in Figure 26 p. 117), and
(2) Isolating cleaved r-amelogenin following His-tag cleavage without the need to use nickel column chromatography (Section 2.1.2.3.2, summarised in Figure 27 p. 119).
The reagents for the gels and buffer used in preparative SDS PAGE were the same as those used for analytical SDS PAGE (Section 2.1.1.4, pp. 100 - 102, the reagents are listed in Table 4 p. 102).
Analytical SDS PAGE was capable of resolving r-amelogenin from bacterial contaminants and in theory could be scaled up to for preparative use. The Model 491 Prep Cell (Bio-Rad Laboratories Ltd., Hertfordshire, UK) was trialled here to assess the potential of preparative SDS PAGE in purifying r-amelogenin. Briefly, the sample was electrophoresed down a cylindrical gel that has the capacity to resolve tens of mg of protein (compared to tens of µg in the case of analytical SDS PAGE). Rather than stopping the electrophoresis and staining the gel as it is the case in analytical SDS PAGE to visualise the resolved protein bands, in preparative SDS PAGE, separated proteins were deliberately electrophoresed off the bottom of the separating gel into an elution chamber from where they were continually removed by a flow of buffer to a fraction collector.

2.1.2.3.1 Isolation of His-tagged r-amelogenin from crude acetic acid extracts by preparative SDS PAGE

After desalting and lyophilisation, acetic acid extracts (containing WT^{+His} r-amelogenin) were solubilised in 1X non-reducing SDS PAGE sample buffer (see Section 2.1.1.4, p.100) at a concentration of 1 mg/mL. Ten millilitres of the starting solution were subjected to preparative SDS PAGE. Preparative electrophoresis was carried out using a Model 491 Prep Cell (Bio-Rad Laboratories Ltd., Hertfordshire, UK). As carried out by Gabe et al (2017), a 12 % resolving gel (composition detailed in Table 4, p. 102) “was cast in the 28 mm internal diameter gel tube to a height of 9.5 cm with a 2-cm 4% stacking gel” (composition detailed in Table 4, p. 102). “The gel was run at a constant 12 W power at room temperature using the circulating cooling pump as recommended by the manufacturer. Immediately after the tracker dye began to run out of the gel, 2.5 mL fractions were collected at a flow rate of 0.8 mL/minute” (Gabe et al., 2017). Every third fraction collected was subjected to analytical SDS PAGE with Coomassie Blue staining to identify approximately which of the fractions contained

WT^{+His} r-amelogenin (further details provided in Section 2.1.1.4, pp. 100 - 102 and Supplementary data (CD) for analytical SDS PAGE). Once the range of fractions containing the WT^{+His} r-amelogenin had been identified, all fractions in that range were subjected to analytical SDS PAGE with silver staining and anti-amelogenin western blotting to accurately determine whether the fractions contained WT^{+His} r-amelogenin at single band purity on silver staining. The fractions comprising WT^{+His} r-amelogenin at single band purity were pooled together, lyophilised, desalted then lyophilised again. The final amount of purified WT^{+His} r-amelogenin obtained by preparative SDS PAGE was weighed to obtain the yield and further characterised by mass spectrometry (See Section 2.1.1.5 p. 104). The purification process is summarised in Figure 26 below. The plan was to subject the lyophilisate to His-tag cleavage and then nickel column chromatography to obtain a purified ‘cleaved’ r-amelogenin (Dashed box in Figure 26).

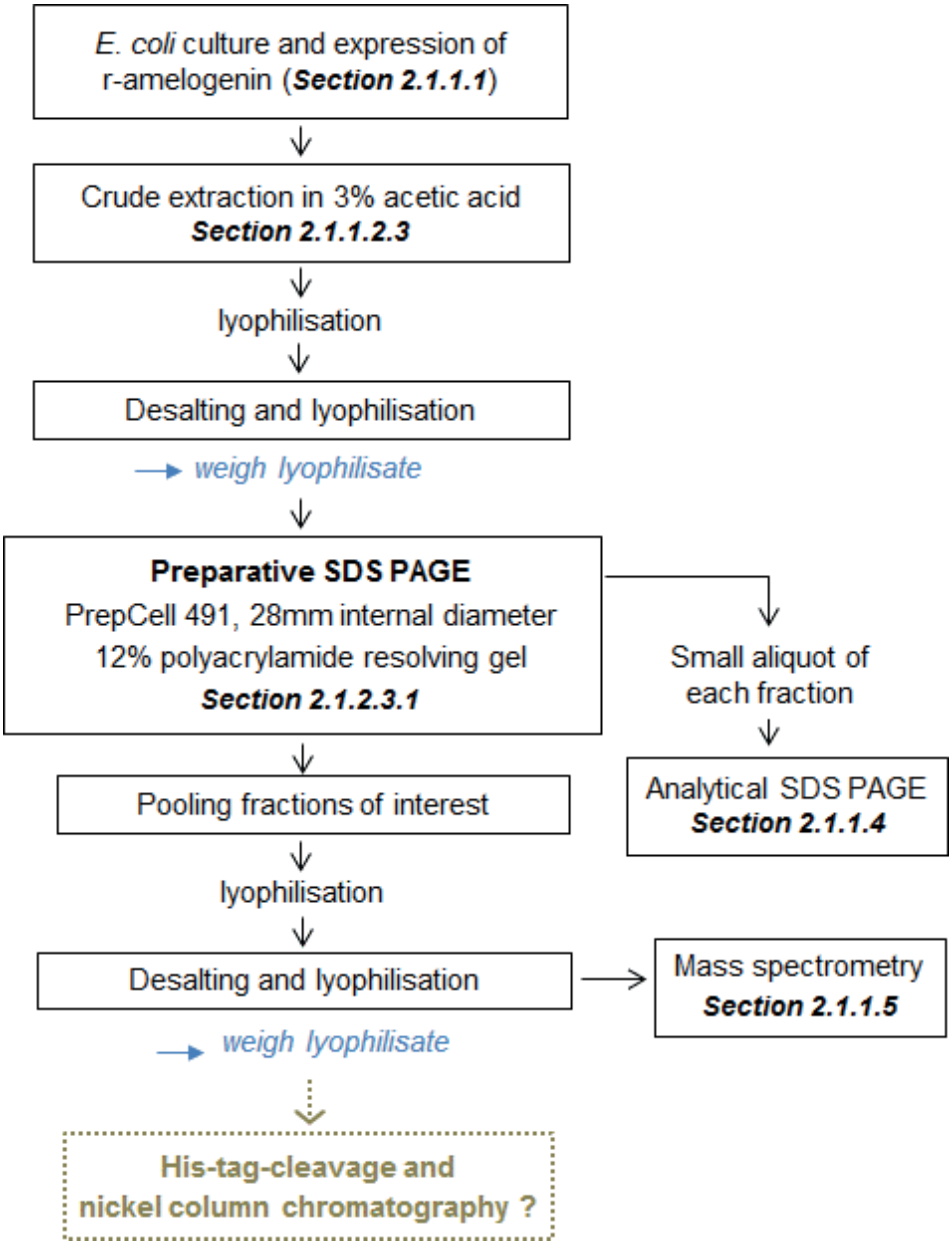


Figure 26 Purification strategy of WT^{+His} r-amelogenin from acetic acid extracts using preparative SDS PAGE. The acetic acid extract was desalted, lyophilised and resuspended into SDS PAGE sample buffer 1X and heated, to solubilise the proteins. The mixture was loaded on to the preparative SDS PAGE gel to separate its components according to their apparent molecular weights. From the fractions collected, small aliquots were taken and their content was analysed by analytical SDS PAGE with Coomassie blue staining, silver staining and western blotting to identify the fractions containing r-amelogenin at single band purity. These fractions were pooled together, desalted, lyophilised and subjected to mass spectrometric analysis.

The purification of the sample achieved with preparative SDS PAGE (shown later in Section 3.1.2.3.1) was such that nickel column chromatography was not actually needed anymore to isolate the cleaved r-amelogenin. Instead, Preparative SDS PAGE was tested to isolate the cleaved r-amelogenin directly after His-tag cleavage of r-amelogenin in acetic acid extract. This second method is described in Section 2.1.2.3.2 below.

2.1.2.3.2 Isolation of cleaved r-amelogenin by preparative SDS PAGE

This Section describes the attempt to isolate cleaved (His-tag-free) r-amelogenin by preparative SDS PAGE after His-tag cleavage of the His-tagged r-amelogenin without any involvement of nickel column chromatography. The process carried out is summarised in Figure 27.

The crude acetic acid extract containing WT^{+His} r-amelogenin was desalted against 125 mM formic acid and lyophilised. The lyophilisate was then dissolved in 0.1 M Na_2CO_3 , pH = 9 at 2 mg/mL and incubated with HRV3C protease (Merck Millipore, Watford, UK) for 48 hours, at 2.5 µL enzyme stock solution per mg protein. The mixture was then lyophilised and desalted against 125 mM formic acid (Section 2.1.1.3, p. 99), then lyophilised again. The lyophilisate was dissolved in 1X non-reducing SDS PAGE sample buffer (as described in Section 2.1.1.4, pp. 100 - 102) at 2 mg/mL and 5mL were subjected to preparative SDS PAGE as described in Section 2.1.2.3.1, p. 115. Every third fraction collected was subjected to analytical SDS PAGE with Coomassie Blue staining to identify approximately which of the fractions contained the cleaved r-amelogenin (further details provided in Section 2.1.1.4, pp. 100 - 102, and in the supplementary data (CD) for analytical SDS PAGE). Once the range of fractions containing the cleaved r-amelogenin had been identified, all fractions in that range were subjected to analytical SDS PAGE with silver staining (further details provided in the supplementary data (CD)). Fractions containing cleaved r-amelogenin at single band purity on silver staining were pooled together, desalted and lyophilised. The final amount of the purified cleaved r-amelogenin was estimated by spectrophotometry at 280 nm.

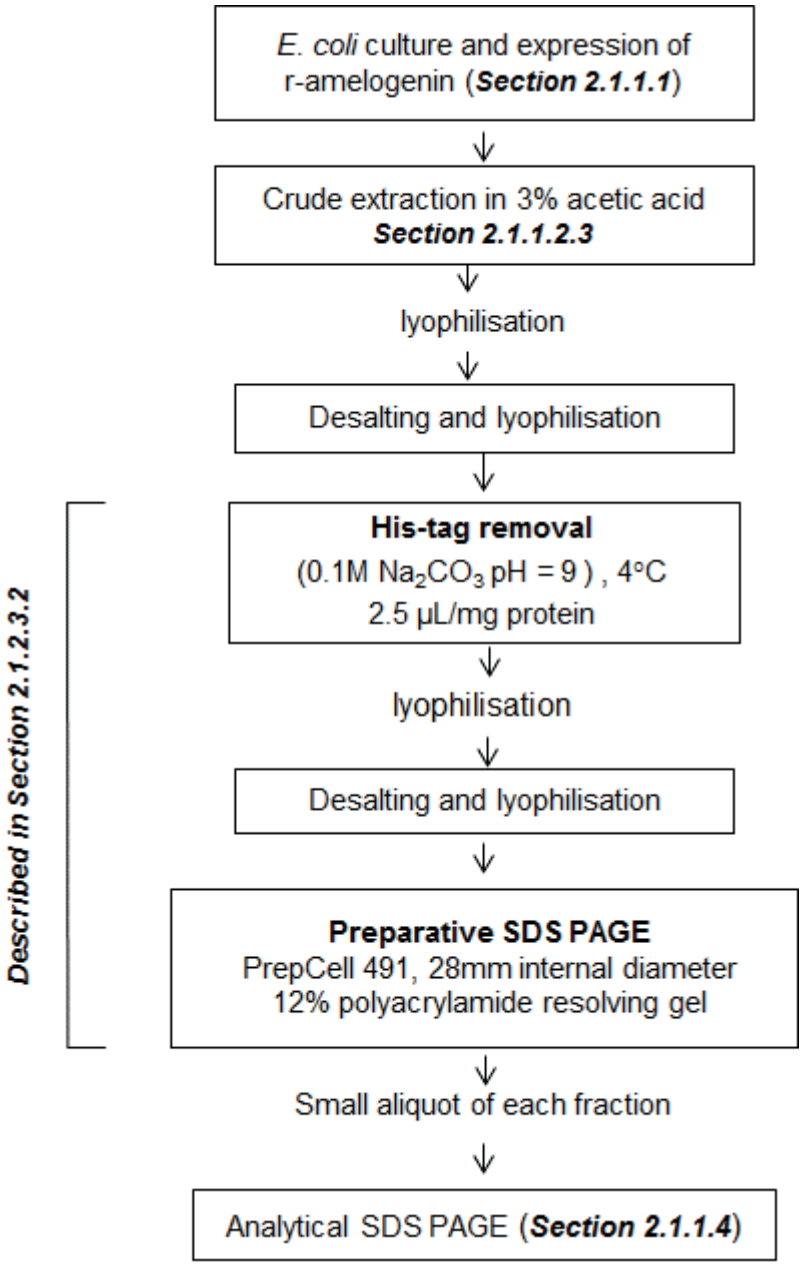


Figure 27 One-step purification of His-tag free 'cleaved' r-amelogenin from acetic acid extract using preparative SDS PAGE. The methodology carried out was the same as that described in Section 2.1.2.3.1 (see Figure 26 p. 117), except that an additional step of His-tag cleavage (with subsequent lyophilisation, desalting and lyophilisation) was added. The fractions obtained from preparative SDS PAGE purification were analysed by analytical SDS PAGE.

.

2.2 Development of protein-binding assays

Two binding assays were developed in order to study amelogenin-amelogenin binding and ultimately to investigate the effect, if any, of the p.Y64H mutation.

2.2.1 Method 1: Fluorescence-based binding assay

The basic strategy for method 1 involved immobilising amelogenin (bait protein) to microplate well surfaces to which was added fluorescein isothiocyanate (FITC)-labelled amelogenin (the free ligand in solution). After incubation, wells were washed and any amelogenin-amelogenin binding was determined by measuring the fluorescence remaining bound to the wells. To conserve r-amelogenin, the 20 kDa amelogenin, abundant in enamel matrix derivative (EMD – the non-commercial form of Emdogain) (Maycock et al., 2002) was purified by preparative SDS PAGE to be used as a surrogate for initial method development.

The EMD was kindly provided by Dr. Petter Lyngstadaas, University of Oslo, NO. The FITC was cat no.F7250 (Sigma-Aldrich, St Louis, MO)

The methods presented below detail

(i) FITC labelling of EMD

(ii) purification of FITC labelled 20 kDa amelogenin from EMD and

(iii) studies using FITC-20 kDa amelogenin in a microplate based binding assays

2.2.1.1 FITC labelling of EMD

The method employed to label EMD with FITC was adapted from the FITC manufacturer's instruction (F7250 - Product Information Sheet, Sigma-Aldrich, St Louis, MO). EMD was dissolved at 5.5 mg/mL in a solution of 0.1 M Na_2CO_3 at pH = 9. The mixture was left stirring at 4°C overnight. FITC was dissolved to 10 mg/mL in dimethyl sulphoxide (DMSO) and FITC solution added slowly by 10 µL increments to the EMD solution at 4°C with stirring to give a final concentration of 58.33 µg FITC per mg EMD. The labelling mixture was then stirred overnight at 4°C. The next day, the mixture was frozen to -80°C and lyophilised.

2.2.1.2 Purification of the 20 kDa amelogenin from FITC-labelled EMD using Preparative SDS PAGE

The lyophilised labelling mixture (prepared as described above, Section 2.2.1.1) was desalted against 125 mM formic acid (see Section 2.1.1.3, p .99) and lyophilised again. The lyophilisate was dissolved in 1X non-reducing SDS PAGE sample buffer

(as described in Section 2.1.1.4, pp. 100 - 102) at a concentration of 8 mg/mL. Approximately 0.8 mL of the starting solution were subjected to preparative SDS PAGE. Preparative electrophoresis was carried out as described in Section 2.1.2.3.1 (p. 115). Fractions containing the labelled 20 kDa amelogenin were pooled together, desalted against 125 mM formic acid (see Section 2.1.1.3, p .99) and lyophilised for binding assays (detailed in Section 2.2.1.3 overleaf). The 20 kDa amelogenin was identified based on its migration on analytical SDS PAGE and by the fact that it is the most abundant species present in EMD (Maycock et al., 2002). The purification process is summarised in Figure 28.

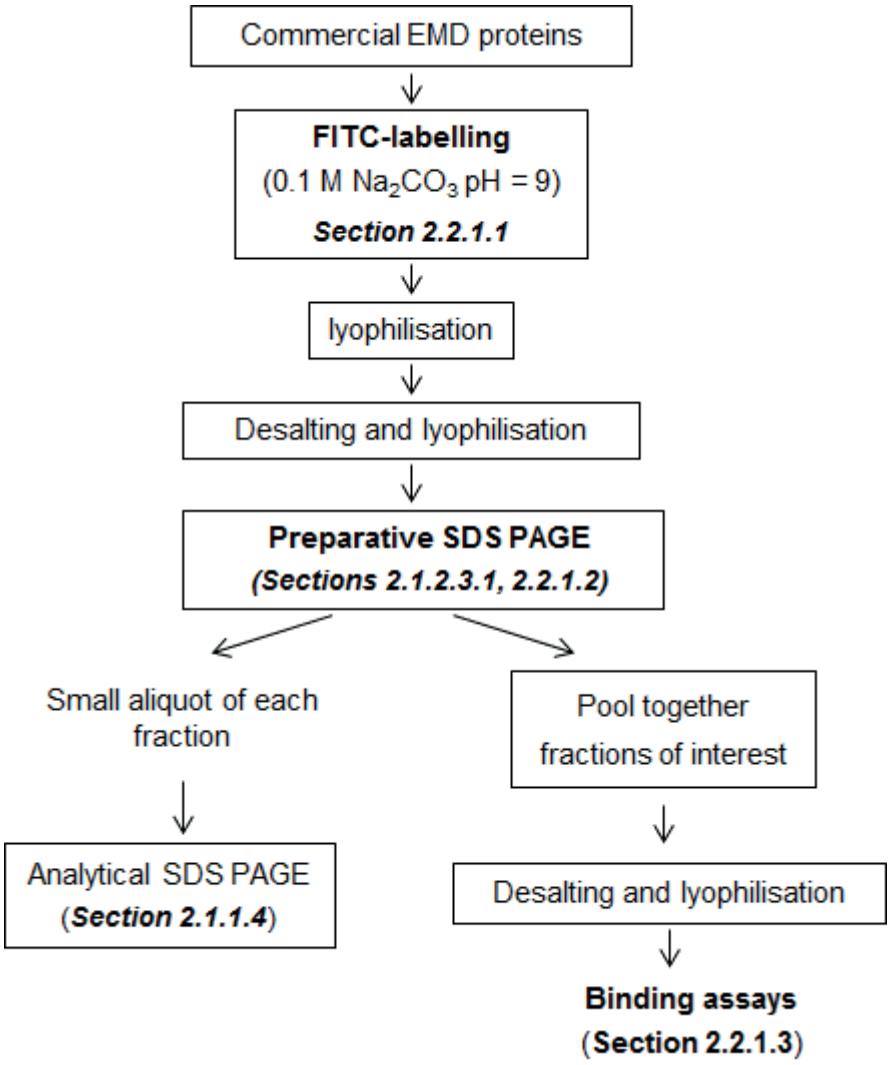


Figure 28 Diagram summarising the preparation and purification of FITC-labelled EMD proteins for fluorescence binding assays.

2.2.1.3 Measuring amelogenin binding between free FITC-labelled amelogenin and immobilised amelogenin as binding partners

To test the effect of the p.Y64H mutation on amelogenin binding behaviour, a microplate-based binding assay was developed, involving the immobilisation of unlabelled amelogenin on the microplate well surfaces to which FITC-labelled amelogenin was allowed to bind. The degree of binding could then be determined (and compared) by measuring the amount of retained fluorescence. This method was adapted from a “FITC-anti-FITC system” previously published (Harmer and Samuel, 1989). As explained earlier (p. 120), the 20 kDa amelogenin from EMD (obtained as described in Section 2.2.1.2) was used as a surrogate.

First, it was necessary to ensure that the unlabelled 20 kDa amelogenin from EMD (bait protein) saturated the polystyrene microwell surfaces to prevent false positives caused by FITC-labelled 20 kDa amelogenin (free ligand) binding to any exposed polystyrene (incompletely covered wells). In addition, it was also necessary to verify whether the binding affinity of the 20kDa amelogenin bait protein for the well surfaces was sufficiently large to resist the various washing and handling steps involved in carrying out the assay. For convenience, FITC-labelled 20 kDa amelogenin (purified from EMD using preparative SDS PAGE as described in Section 2.2.1.2, p. 121) was used as a surrogate unlabelled bait protein to investigate the binding kinetics of amelogenin to the microwell surfaces; this could be easily measured using the fluorescence microplate reader. Note: for clarification, the actual binding assays were to be carried out later, using non-labelled r-amelogenin bound to the wells as the immobilised bait protein with FITC-labelled r-amelogenin being used as the free ligand in solution.

In the optimisation experiments, FITC-labelled 20 kDa amelogenin, acting as surrogate bait protein, was dissolved at 20 µg/mL in binding buffer (see below) and serially diluted to concentrations of 1, 2, 5, 7.5, 10 and 20 µg/mL. Three different binding buffers were tested for efficacy of coating: 0.1 M Na_2CO_3 at pH=9, PBS at pH=7.4 and TBS at pH=7.2. The incubation of the plates was carried out overnight at either room temperature or 4°C.

One hundred microlitres of unlabelled 20 kDa amelogenin in binding solution were added per microwell in triplicate and left to incubate overnight with the microplates sealed with parafilm. The microplates used were black 96-well microfluor 2 plate (cat. No. 7205, ThermoFisher Scientific, Loughborough, UK). After overnight incubation, the fluorescence signal was measured using an excitation wavelength of 490 nm and an emission wavelength of 525 nm. The microwells were then washed. Washing consisted of emptying the microwell content by discarding binding solution and tap-

drying, followed by addition of 120 µL of washing buffer (either PBS + 0.05 % Tween20, or TB S+ 0.05 % Tween20) which was subsequently discarded and the plates tap-dried. After washing the microwells, 150 µL of blocking solution were added per microwell and left to incubate for 1 hour at room temperature and sealed. The blocking solution was made up of 1 % bovine serum albumin (BSA) in washing buffer (blocking serves no purpose in this experiment but was included as it would later feature in the actual binding assays; to prevent false positives caused by FITC-labelled ligands binding to the well surfaces directly rather than the immobilised bait proteins).

After blocking, the microwells were washed 3 times as described above and pure water was added prior to fluorescence detection (490 nm/525 nm). The fluorescence associated with bound 20 kDa amelogenin was then measured. The process is summarised in Figure 29.

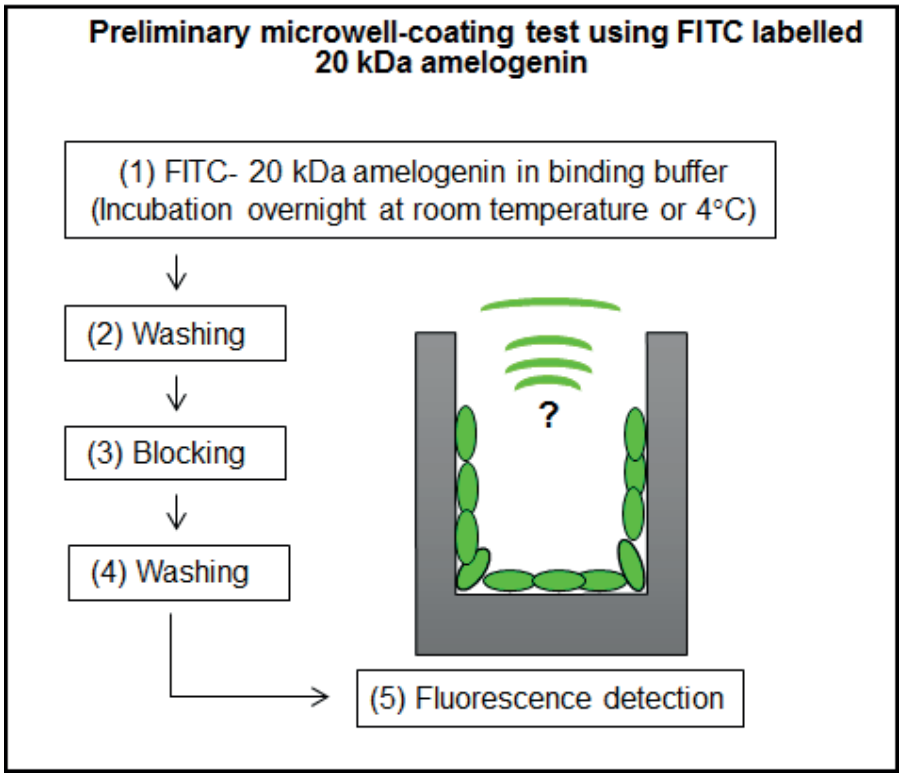


Figure 29 Design of experiment to determine whether amelogenins can adsorb to, and coat the surfaces of the microwell for later use in a fluorescence binding assay. FITC-labelled 20 kDa amelogenin at 1, 2, 5, 7.5, 10 and 20 µg/mL were adsorbed onto the microwell surface to determine the lowest concentration required to saturate the microwell surface. Saturating the microwell surface with bait protein was essential in order to reduce false positives in the later binding experiments if protein ligands free in solution were able to adsorb directly to the exposed microwell surface rather than to the immobilised bait proteins.

2.2.2 Method 2: Determining the kinetics of protein-microwell binding by monitoring the disappearance of amelogenin from solution

The second method to characterise amelogenin binding events in microwell plates involved simply incubating unlabelled EMD proteins (comprising manly amelogenin) in UV transparent microwells. The decrease in UV absorbance was monitored (illustrated in Figure 30) as proteins initially adsorbed to the microwell surfaces (protein-polystyrene binding) and free proteins bound to the immobilised proteins (protein-protein binding). Protein concentration in solution is frequently measured spectrophotometrically at 280 nm, which is the absorbance maximum of the aromatic side chains of tryptophan and tyrosine (Layne, 1957). However, these amino acids are not particularly common in amelogenin so in order to increase sensitivity, absorbance was measured at 220 nm, which is within the wavelength range at which peptide bonds maximally absorb UV (Goldfarb et al., 1951) and is within the transparency range of the UV transparent microplates used.
The microplates employed were polystyrene NUNC 96-well UV transparent microplate 8404 (Thermofischer Scientific, Loughborough, UK). Absorbance was monitored at 220 nm over time after adding unlabelled EMD protein and recorded using a Varioskan Flash microplate reader - SkanIt software (Thermoscientific, UK).

The methods presented in this Section detail the development of the microplate binding assay with regards to:

(i) The effect of the initial EMD concentration on the EMD-polystyrene binding equilibrium (Section 2.2.2.1),

(ii) Optimisation of the methodology to investigate the kinetics of EMD protein-protein interactions (Section 2.2.2.2),

(iii) The effect of EMD proteins adsorbed to the bottom of the microwell and attenuating the UV before it passes through the free proteins still in solution potentially confounding the absorbance readings (Section 2.2.2.3)

(iv) Attempts to block EMD adsorption to the bottom of the microwells to circumvent the issue of UV attenuation by EMD adsorbing to the bottom of the micro wells (Section 2.2.2.4).

(v) A use of the assay to study p.Y64H mutant and WT r-amelogenin interactions (Section 2.2.2.5).

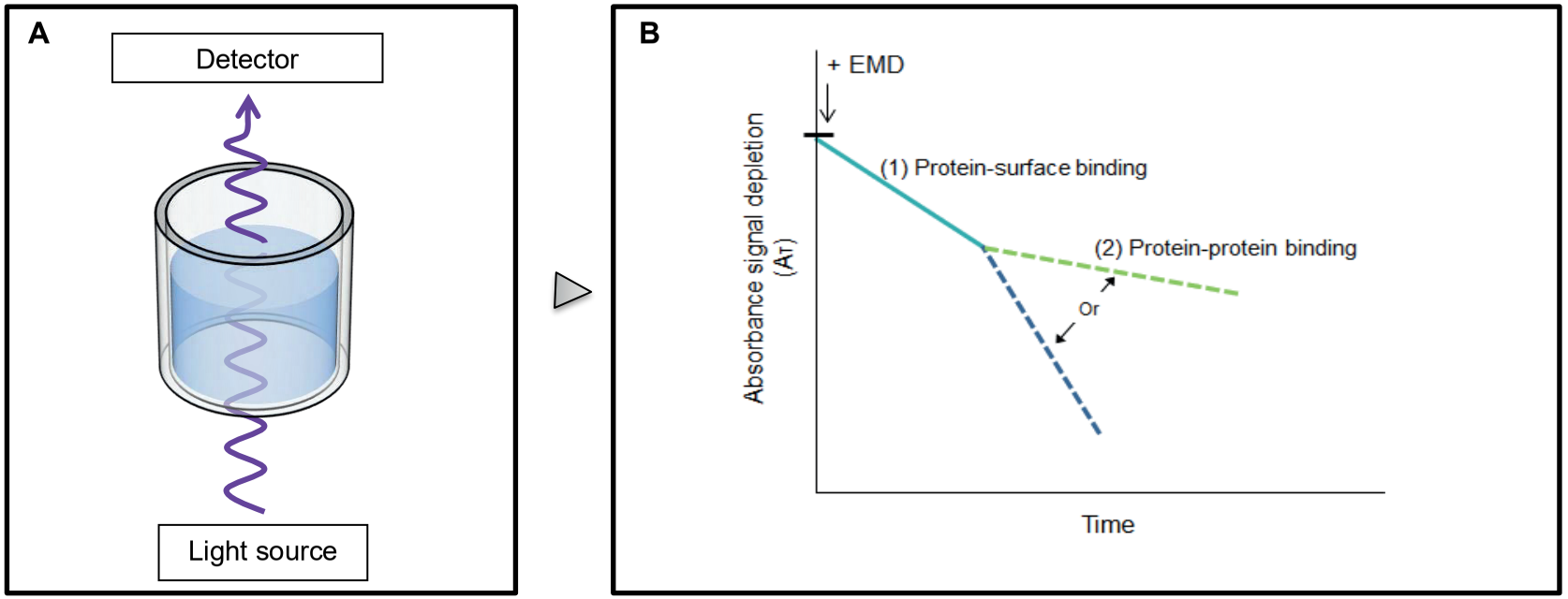


Figure 30 Principles of protein binding kinetics based on protein depletion from solution as detected using UV absorbance. (A) Photometric measurements are based on UV light transmission through a sample contained in a microwell. A UV light source located below the microplate illuminates the protein sample. The light beam passes up through the sample solution and the absorbance is measured by a detector. (B) By Beer-Lambert's law, absorbance will fall as solubilised proteins adsorb to the microwell surface and are taken out of solution (step (1), solid turquoise line). The hypothesis is that a point will be reached when the microwell surfaces become saturated. At this time, proteins would continue to be removed from solution by interaction with the protein already immobilised to the microwell surfaces. This second phase of protein removal from solution may occur at a different rate, resulting in an inflection point (step (2) dashed lines). If protein-protein binding is faster than protein-surface binding, the signal will resemble the dark blue dashed line. If protein-protein binding is slower than protein-surface binding, the signal obtained will resemble the green dashed line.

2.2.2.1 Effect of the initial EMD concentration on binding equilibrium: determination of an EMD concentration required to saturate the microwells surfaces.

The minimum concentration of EMD proteins required to saturate the polystyrene surface of the microwells was optimised by measurement of loss of protein from solution using spectrophotometry in the UV at 220 nm.

At room temperature, EMD was dissolved in 1 % acetic acid to 48,308 µg/mL to prepare a stock solution. The stock was then diluted 1:290 with PBS to give a final EMD concentration of 166 µg/mL. This solution was then further diluted with PBS (containing 1 % acetic acid at 290:1) to give solutions containing EMD at 166, 99.6, 66.4 and 33.2 µg/mL. Immediately after preparing the EMD solutions, aliquots were taken for SDS PAGE analyses. Then, 200 µL of each solution were added into empty polystyrene UV-transparent microwells (n=6 replicates) and left to incubate for 24 hours in the microplate reader[1]. Absorbance measurements at 220 nm were recorded every hour during the incubation period. The averages of the 6 replicates were plotted using Excel software, after subtracting the blank values and the standard deviations calculated. At the end of the incubation, 15 µL aliquots were taken from each of the replicates, pooled, frozen to -80°C and lyophilised. The lyophilisates were re-dissolved to their original volume in 90 µL non-reducing sample buffer for analytical SDS PAGE (see Section 2.1.1.4, pp. 100 - 102 and Supplementary data (CD))

2.2.2.2 Optimisation of the methodology to analyse the kinetics of protein-protein interactions

In the Section above (2.2.2.1), the conditions required to saturate the surface of the microwells with EMD were identified. The next stage in developing the binding assay was to determine the effects of adding a fresh solution of EMD to the already saturated microwell surface to investigate whether subsequent protein – protein interactions occurring between immobilised EMD (bait protein) and the freshly added EMD free in solution could be followed by monitoring the disappearance of free EMD from solution by UV absorbance spectroscopy. Figure 31 below shows a hypothetical result illustrating the principle of the method. Any decrease of absorbance (protein concentration in solution), during the second incubation period would be assumed to reflect protein-protein binding rather than the loss of protein due to a combination of

[1] Note: The absorbance measurements started 10 minutes after the EMD solution was prepared. As 10 minutes represent 0.7% of the 24-hour incubation period, these 10 minutes were considered as negligible and the starting time of absorbance measurement was considered to be the starting time of experiment.

protein-surface + protein-protein interactions occurring during the first incubation period during which the microwell surface becomes saturated.

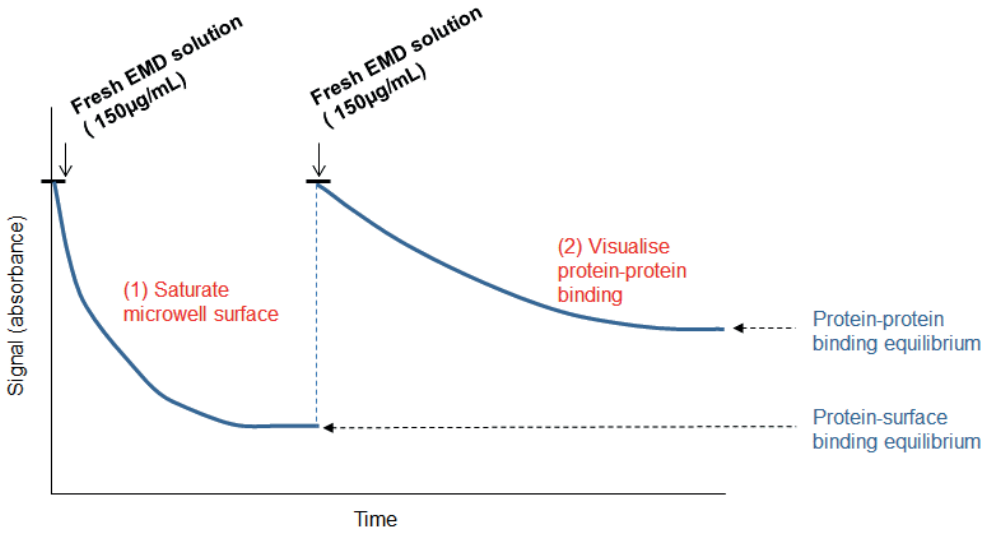


Figure 31 Hypothetical signal pattern obtained while monitoring initial EMD-polystyrene interactions that saturate the microwell surface and subsequent EMD-EMD interactions occurring after the addition of fresh protein solution. The absorbance decrease in this second incubation period is assumed to be due to depletion of the solution due to the freshly added EMD interacting with the EMD already immobilised to the polystyrene surface during the initial saturation step.

The procedure is described below and summarised in Figure 32 (p. 129).
EMD was accurately dissolved in 1 % acetic to 42,735 µg/mL to provide a stock solution. The stock was then diluted 1:284 with PBS (pre-warmed to 37°C) to give a final EMD concentration of 150 µg/mL. Immediately after preparing the solutions, aliquots were taken for SDS PAGE analyses. A blank solution comprising of 1 % acetic acid diluted with PBS (pre-warmed to 37°C) at a ratio 1:284 was also prepared. Then, 200 µL of each solution were added into empty polystyrene UV-transparent microwells (n=6 replicates), and subjected to an initial incubation period of 24 hours at 37°C during which time the well surfaces became saturated with EMD protein ("Phase 1"). The microplate was covered with cling film to prevent evaporation and the process was monitored by recording absorbance measurements at 30-minute intervals throughout the incubation period. The average absorbance of the 6 replicates were plotted with after the subtracting the blank at each time point and standard deviations calculated. At the end of the incubation period, 15 µL aliquots were taken from the solutions in each microwell replicate, pooled, frozen to -80°C and lyophilised. The lyophilisates were redissolved to their original volume (90µL) in non-reducing sample

buffer (see Section 2.1.1.4, p.100) for analytical SDS PAGE. Further details regarding the SDS PAGE analyses are provided in Section 2.1.1.4 (p. 100 - 102) and in the supplementary data (CD)).

The microplates were then inverted and tapped forcefully to remove the residual solutions from the microwells.

Aliquots of a fresh solution of the EMD (at concentration 150 µg/mL as described for phase 1 above) were taken for analytical SDS PAGE and a second incubation period was then immediately initiated ("Phase 2") by adding 200 µL of the fresh EMD solution to the recently emptied microwells (the resulting protein interactions are referred to as "EMD- EMD" i.e. free EMD was presumed to be interacting with EMD already immobilised to the surface). Two-hundred microliters of the fresh EMD solution were also added into microwells previously incubated with blank solutions as a control (this condition is referred to as "EMD-Polystyrene" i.e. free EMD is just interacting with the polystyrene surface and is essentially a repeat of phase 1 but it is concomitant with phase 2). The microplate was covered with cling film to prevent evaporation and the second incubation lasted 24 hours at 37°C. UV absorbance at 220 nm was again recorded every 30 minutes throughout. The average absorbance readings of the 6 replicates were plotted, readings for the blanks subtracted and standard deviations calculated. After the incubation had ended, aliquots were taken from the solutions in the microwells for SDS PAGE as described above for phase 1. Further details regarding the SDS PAGE analyses are provided in Section 2.1.1.4, pp. 100 - 102 and in the supplementary data (CD)).

The data points obtained during the second incubation period (describing EMD- EMD interactions and EMD-polystyrene interactions (control)), were replotted to represent the gain of protein by the polystyrene surface rather than the loss of protein from solution. The points were replotted as A_0-A_T (where A_0 = initial absorbance at the start of phase 2 and A_T = absorbance measured at time T between the start and the end of phase 2). They were then subjected to curve-fitting on OriginPro 9.1 (OriginLab corporation, Northampton, MA), to identify the equation that best fitted the data which would allow kinetics descriptors such as half-time $T_{1/2}$ and the maximum ordinate value (i.e. the final amount of protein adsorbed to the polystyrene surface at equilibrium) to be calculated.

OriginPro comprises over 150 functions for curve-fitting, grouped as "function selection" sets (e.g.: exponential, logarithm, polynomial). Within each set, OriginPro can perform curve-fitting with model ranking for a given dataset. For the "gain of protein by polystyrene surface" data plots, 70 functions fitted the data to some degree, with 21 yielding an r-square value of between 0.99 and 1. The Hill equation model (Hill, 1910) was one of these and fitted both plots for "EMD-EMD" and "EMD-

Polystyrene" reactions. It was therefore chosen as the best fitting model and was used to predict $T_{1/2}$ and the maximum ordinate value reached.

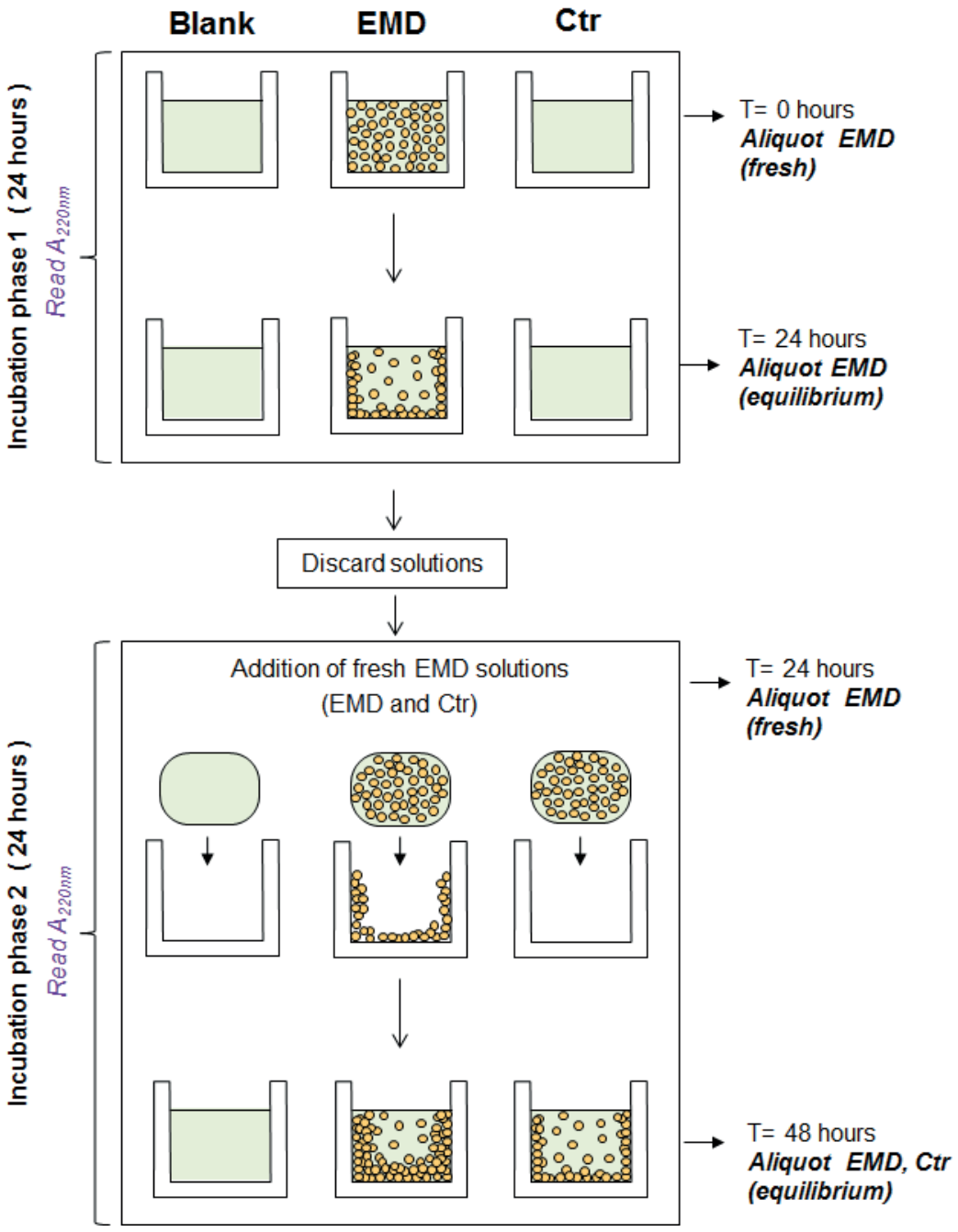


Figure 32 Summary of the methodology used to investigate EMD-EMD interactions using a microplate based assay. A solution of EMD proteins (150 µg/mL) was left to incubate for 24 hours in microwells (phase 1). Aliquots of the solution were taken for SDS PAGE analyses before and after incubation (at the time points T= 0 hours, 24 hours, indicated). Then the microwells were emptied and fresh EMD solution (150 µg/mL) was added again in the test microwells (phase 2). Aliquots of the solution were taken for SDS PAGE analyses before and after incubation (at the time points T= 24 hours, 48 hours, indicated). The absorbance was monitored at 220nm.

2.2.2.3 Effect of proteins adsorbed on the bottom of the microwells on the absorbance value: a possible confounder

During the course of these binding experiments, it was an inevitability that proteins would bind to the bottom of the microwells in addition to the microwell sides. The aim of this experiment was to investigate how the microplate absorbance measurements were affected by proteins binding to the bottom of the microwells. The methodology was again based on determination of the depletion of free protein from solution. Proteins coating the bottom of the microwells are lost from solution but they may still contribute to the UV absorbance measured, as they will still be in the light path. The question was therefore whether the layer of proteins adsorbed to the bottom of the microwell exhibited significant UV absorbance compared to the amount in solution, potentially confounding the data. To determine the significance of this potential confounder a solution of EMD (at concentration 150 µg/mL) was freshly prepared as described above (Section 2.2.2.2) and 200 µL were incubated for 24 hours at 37°C in the microwells (n=6) along with blank microwells containing a solution comprised of PBS:1 % acetic acid at a ratio 284:1. The microplate was covered with cling film to prevent evaporation. At the end of incubation, the solutions were discarded from the microwells, immediately after which the absorbance was read again. The average absorbance measurements were recorded[1] and the blank values subtracted. Any absorbance remaining was assumed to be associated with proteins adsorbed to the microwell bottom.

The data obtained (in Section 3.2.2.3) suggested that proteins adsorbing to the bottom of the microwells (either directly to the surface or via immobilised proteins already present on the surface) could be reliably detected using UV absorbance spectroscopy. This could therefore provide an ‘end point’-based microplate assay for monitoring protein-protein interactions (as opposed to the method described above based on using UV absorbance spectroscopy to monitor the depletion of protein from solution as it interacts with immobilised proteins already on the well surfaces). However, as described in the next Section an attempt was made to address the issue of protein adsorption to the bottom of the microwells by pre-blocking the bottom of the microwells with standard blocking proteins used in numerous assays such as western blotting and enzyme linked immunosorbent assays.

---

[1] Note: The absorbance measurements started 4.5 minutes after the EMD solution was prepared. As 4.5 minutes represent only 0.3% of the 24-hour incubation period, the starting time of absorbance measurement was considered to be the starting time of experiment.

2.2.2.4 Blocking the bottom of the microwells to inhibit adsorption of test proteins

To prevent any EMD proteins from binding to the bottom of the microwells and confounding the determination of the absorbance values of the solutions remaining in the microwells, an attempt was made to block EMD binding to the bottom of the microwells. This required an effective blocker that would prevent EMD proteins from binding to the polystyrene surface. Microplate-based immunological assays employ various blocking regimes to abolish non-specific binding of detection antibodies to exposed microwell surfaces once the sample containing the antigenic target of interest has been adsorbed (Gibbs, 2001). Gold standard blockers bovine serum albumin (BSA) and non-fat dry milk proteins (NFDM) were tested therefore to investigate their ability to prevent the unwanted adsorption of EMD on to the microwell bottom. Solutions of BSA and NFDM were added into microwells and the proteins allowed to bind (block) the microwell surfaces. The microwells were then washed and the ability of the blocked surfaces to bind EMD were tested by adding EMD in solution and monitoring its depletion from solution using UV absorbance spectroscopy. As shown in Figure 33 below, effective blocking was expected to result in the absorbance remaining constant over the incubation period whereas the absence of blocking (control) would cause the absorbance to fall over time as the EMD was removed from solution by binding to the well surfaces.

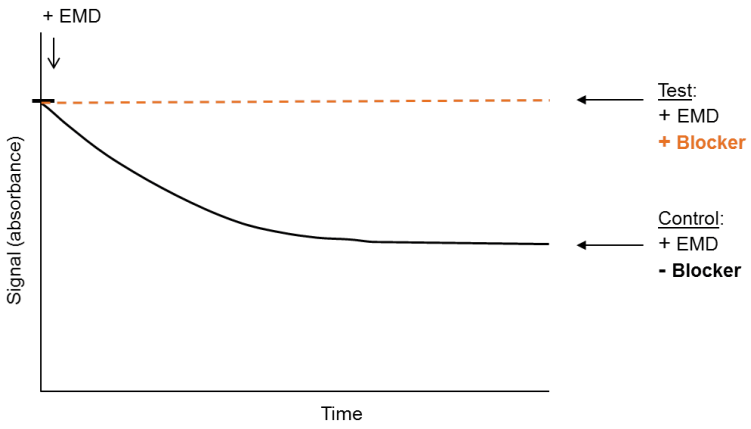


Figure 33 Hypothetical signal expected over incubation time in microwells effectively blocked (orange dashed line) compared to control (solid black line) with no blocking . In case of effective blocking the concentration of EMD in solution would remain constant, causing a constant absorbance signal overtime. In contrast, in the control without blocking, absorbance would fall with time as the concentration of EMD in solution falls as it adsorbs the well surfaces.

Blocking solutions were prepared by dissolving BSA cat no. 05479 (Sigma-Aldrich, St Louis, MO) and NFDM cat no. 170-6404 (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK) at 1 % in a mixture of PBS and 1 % acetic acid (290:1). Two hundred microliters of each blocking solution were added to microwells (n=6) and the microplate was sealed with parafilm and incubated overnight at 4°C. The next day, the microwells were washed 3 times with PBS and tapped dry. EMD dissolved at 166 µg/mL in a mixture of PBS and 1 % acetic acid (290:1) was prepared and an aliquot was taken immediately for SDS PAGE analyses later. Two hundred microliters of the EMD solution were then added to the blocked microwells. As a control, 200 µL of EMD were also added to non-blocked wells (n=6). The absorbance was recorded[1] at 220 nm at hourly intervals for 24 hours at room temperature. After incubation, 15 µL aliquots were taken from the microwells, pooled, frozen to -80°C and lyophilised. The lyophilisates were re-dissolved to their original volume (90 µL) in non-reducing sample buffer for analytical SDS PAGE (see Section 2.1.1.4 pp. 100 - 102 and Supplementary data (CD)). The experiment is summarised in Figure 34 below.

[1] Note: The absorbance measurements started 10 minutes after the EMD solution was prepared. As 10 minutes represent only 0.7% of the 24-hour incubation period, the starting time of absorbance measurement was considered to be the starting time of experiment.

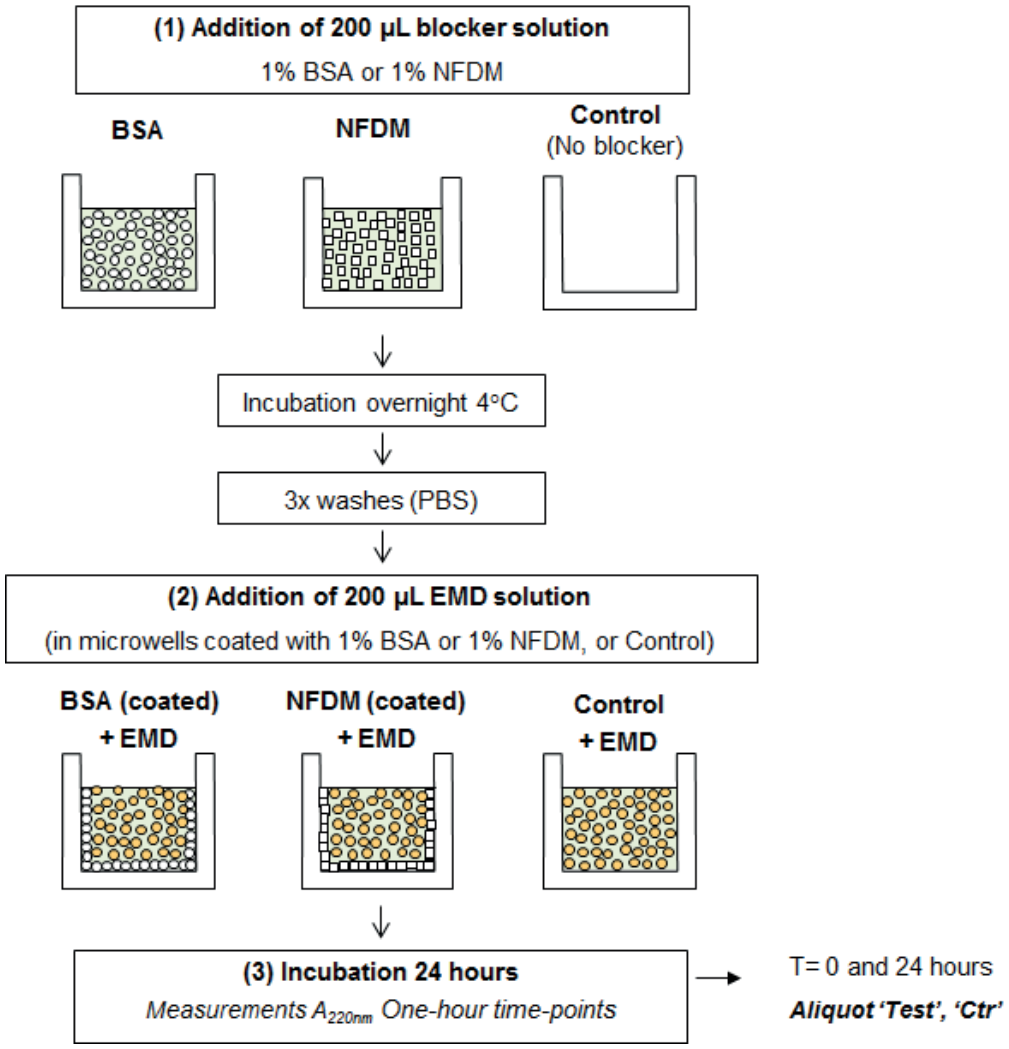


Figure 34 Summary of microwells blocking trial. (1) A solution of 1% BSA or 1% NFDM was left to incubate overnight in microwells at 4°C. At the end of incubation the microwells were washed and tapped dry. (2) Then, fresh EMD solution (166 µg/mL). The control condition consisted of incubating fresh EMD solution in microwells that were not previously incubated with BSA or NFDM. Aliquots of the solution were taken for SDS PAGE analyses before and after incubation with EMD (at the time points T= 0 hours, 24 hours, indicated). The absorbance was monitored at 220nm.

2.2.2.5 Behaviour of WT and mutant pY64H r-amelogenins in the microplate-based assay

The r-amelogenins used throughout the thesis carried a His-tag for purification by nickel column chromatography (see Figure 24, p. 108), which was the initial purification strategy and a standard method for purifying recombinant proteins. The initial aim was then to obtain cleaved and purified WT and p.Y64H mutant r-amelogenins for use in protein binding studies. However, nickel column chromatography did not completely purify r-amelogenins successfully (see Section 3.1.2.1.3), while preparative SDS PAGE provided highly purified fractions (see Section 3.1.2.3.1). As the initial acetic acid extraction and subsequent preparative SDS PAGE could reach a high degree of purification, it became clear that nickel column chromatography and the associated His-tag were redundant. A further advantage with using His-tag-free r-amelogenins is that it bypasses the time-consuming and costly His-tag cleavage step. To this end, His-tag-free WT and p.Y64H mutant r-amelogenins (respectively "WT^{-His}" and "Mut^{-His}" r-amelogenins) were expressed by *E. coli* BL21 DE3 competent cells using plasmids generated by a commercial company (Novoprotein Scientific, NJ, USA) during the final write up period of the thesis (earlier attempts by the company to produce these plasmids were unsuccessful due to 'technical problems'). Details regarding the plasmids and sequences are provided in Appendix A. WT^{-His} and Mut^{-His} r-amelogenins were initially purified using the acetic acid extraction procedure (kindly performed by Dr Sarah Myers and Mr. Matthew Percival). Time precluded further purification by preparative SDS PAGE and as described below, these were used directly along with WT^{+His} and Mut^{+His} r-amelogenins (also extracted in 3% acetic acid) in the microplate assay which simply measured the absorbance of protein accumulating directly on the bottom of the wells (see Section 2.2.2.3, p. 130). The acetic acid extracts containing r-amelogenins were lyophilised, desalted against 125 mM formic acid (see Section 2.1.1.3, p. 99*, for details on desalting procedure) and lyophilised. The protein content of the extracts was analysed by mass spectrometry (see Appendix B for His-tag-free r-amelogenins and Figure 39 (p. 145) for His-tagged r-amelogenins). The lyophilisates were then dissolved in 1 % acetic acid to provide a 1.62 mM stock solution. The stock solutions were diluted 217.4 times in PBS pre-warmed to 37°C, to give a final r-amelogenin concentration of 7.46 µM. Considering that r-amelogenins were by far the major proteins in the acetic acid extracts, the concentration values disregarded any contaminants present. Immediately after preparing the test solutions, aliquots were taken for analytical SDS PAGE. Then, 200 µL of each solution were added into empty polystyrene UV-transparent microwells (n=6) and left to incubate for 24 hours at 37°C. The microplate was covered with cling

film to prevent evaporation. The depletion of protein from solution was monitored by UV absorbance spectroscopy at A_{220nm}. Absorbance was recorded every minute during the first 4 hours of incubation, then every 30 minutes until the end of the incubation period. The average absorbance of the 6 replicates for each of the 4 recombinant proteins were plotted[1] after subtracting the blank values using Excel software, with the standard deviations calculated. The blank solution was comprised of PBS:1 % acetic acid at a ratio 216.4:1 (PBS: acetic acid). At the end of the incubation, 20 µL aliquots were taken from each of the microwell replicates, pooled, frozen to -80°C and lyophilised. The lyophilisates were re-dissolved to volume (120 µL) in 1X non-reducing sample buffer for analytical SDS PAGE. Further details regarding the SDS PAGE analyses are provided in Section 2.1.1.4 pp. 100 - 102 and in the supplementary data (CD)). Meanwhile the microplates were inverted and tapped forcefully to remove the solutions from the microwells, and the UV absorbance at 220 nm recorded and attributed to proteins present on the bottoms of the wells.

[1] Note: The absorbance measurements started 7 minutes after the EMD solution was prepared. As 7minutes represent less than 0.5% of the 24-hour incubation period, the starting time of absorbance measurement was considered to be the starting time of experiment.

Chapter 3 Results

3.1 Production and purification of r-amelogenins

3.1.1 Expression of recombinant WT and p.Y64H amelogenins in *E. coli* and their extraction in acetic acid.

3.1.1.1 Expression of recombinant WT and mutant p.Y64H amelogenins in *E. coli*

R-amelogenins WT^{+His} and Mut^{+His} were produced in large quantities using an *E.coli* expression system. The expression of WT^{+His} or Mut^{+His} r-amelogenin was induced with IPTG when the cell density was sufficient, that is, when the culture density OD_{600nm} was above 0.6 AU (see Section 2.1.1.1, pp. 92 - 93).

From a starting value below 0.1 AU, which corresponds to the starting time of the large scale growth of *E. coli*, the OD_{600nm} reached the value 0.67 AU after 2.5 hours growth. At this time, the expression of r-amelogenin was induced by adding IPTG and the cultures were incubated overnight. Figure 35 (p. 137) shows the SDS PAGE of total protein present in the *E. coli* . Cultures induced overnight with IPTG expressed a prominent 27 kDa protein that was absent in non-induced cultures. This band was confirmed to be amelogenin by western blotting using anti-amelogenin specific antibodies (Figure 35B). Prominent cross-reactively was present at 27kDa with less intensely stained bands at 54 kDa, suggesting the presence of an amelogenin dimer. A cross-reactive band at 30 kDa was present in samples from both non-induced and induced *E. coli* cultures. Clearly, some bacterial protein of unknown identity was cross-reacting with the anti-amelogenin antibody. The r-amelogenin was seen to be running 3-5 kDa higher than the positive control due to the presence of a His-tag in the case of the recombinant. Also, cross-reactive bands at 15 and 20 kDa were present, at relatively low concentration in samples from induced *E. coli* cultures. These may be breakdown products of the r-amelogenin.

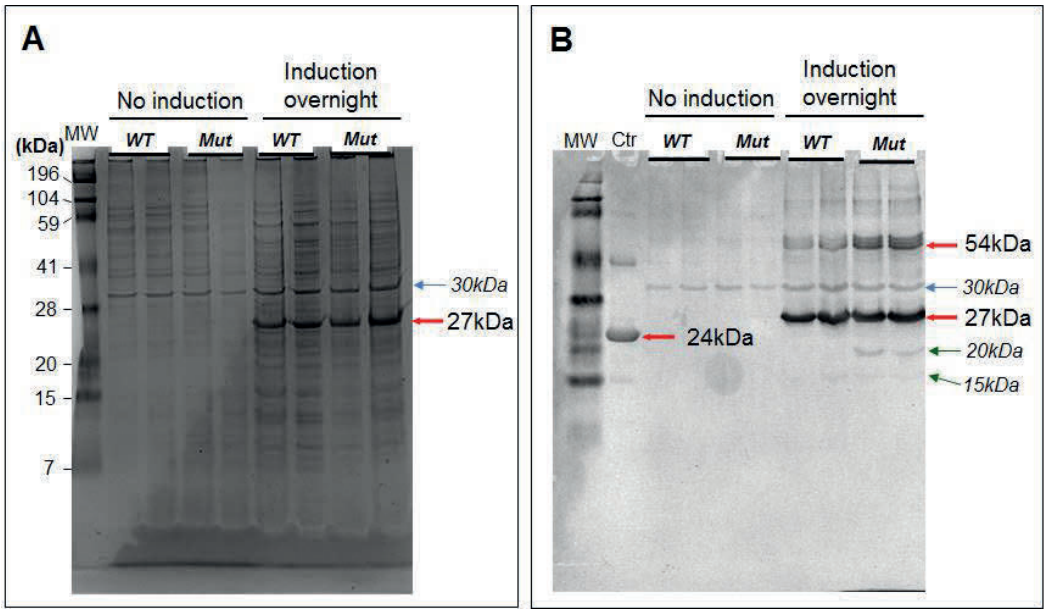


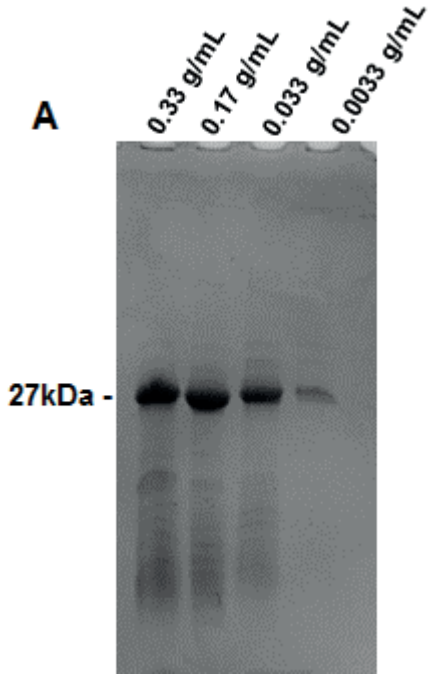
Figure 35 SDS PAGE and western blotting of expressed *E. coli* proteins. R-amelogenin was expressed by *E. coli* cells in large quantities, as WT^{+His} or Mut^{+His} (carrying a His-tag). (A) SDS PAGE showing that overnight incubation after induction with IPTG resulted in the expression of a prominent protein at 27 kDa. (B) Western blot showing that the 27 kDa protein cross-reacted with anti-amelogenin antibodies. with less intense cross-reactivity at 53 kDa.
(MW = molecular weight markers; Ctr = r-amelogenin (minus His-tag) at 24 kDa).

3.1.1.2 Optimisation of 3% acetic acid amelogenin extraction procedure

3.1.1.2.1 Optimisation of *E. coli* extraction in terms of 'weight of *E. coli* to volume of acetic acid' used in the extraction procedure

To identify the ratio of weight of *E. coli* to volume of acetic acid that gave the maximum yield of r-amelogenin in a single extraction step but without over dilution, *E. coli* paste was resuspended in 3% acetic acid at 0.33, 0.17, 0.033 and 0.0033 g/mL (see Section 2.1.1.2.1, p. 95). Figure 36A below shows that easily detectable amounts of r-amelogenin were present in the 0.33, 0.17 and 0.033 g/mL extracts whereas r-amelogenin was less easily detected in the 0.0033 g/mL extract (when equal volumes of sample were loaded on the gels). The obvious reason for this is that the r-amelogenin was diluted by a factor of 100 in the 0.0033 g/mL extract compared to the 0.33 g/mL extract. To account for the different dilution factors, the staining intensity associated with each band was quantified by gel densitometry (Figure 36B). From the table it is clear that extracting r-amelogenin from *E.coli* paste at 0.0033 g/mL provided the greatest yield of r-amelogenin. However, large volumes are more difficult to handle in downstream processing steps (e.g. lyophilisation) so as a compromise, extraction at 0.033 g/mL was chosen as the optimum 'weight of *E. coli* to volume of acetic acid' ratio for extraction (giving a yield of r-amelogenin about 7 times greater than extracting at 0.33 g/mL).

A



B

Lane number (weight *E. coli* to volume acetic acid)	**Dilution factor relative to lane 1**	**Band intensity** (arbitrary units)	**Band intensity corrected** (arbitrary units)
Lane 1 (0.33 g/mL)	1x	29.4	29.4
Lane 2 (0.17 g/mL)	2x	29.7	59.3
Lane 3 (0.033 g/mL)	10x	20.5	205.0
Lane 4 (0.0033 g/mL)	100x	5.7	570.0

Figure 36 Optimisation of the weight of *E. coli* paste to volume of acetic acid in the extraction process. (A) The contents of the supernatants obtained are shown on analytical SDS PAGE with Coomassie Blue staining. (B) The inset table compares the intensities of the 27 kDa band on the gels read by gel densitometry. When corrected for the dilution factor, the intensities obtained were clearly higher for the larger volumes of extraction (or lower ratios 'weight of *E. coli* paste to volume of acetic acid').

3.1.1.2.2 Optimisation of the mixing regimen and temperature on the yield of WT r-amelogenin extraction

The results reported above suggested that using 30 mL of acetic acid to extract WT^{+His} r-amelogenin from 1 g of *E. coli* (equivalent to 0.033 g/mL extract, see Section 3.1.1.2.1 above) was efficient and demanded relatively little effort to extract the bulk of the WT^{+His} r-amelogenin. However, for *E. coli* expressing Mut^{+His} , the process was not so efficient, as it was more difficult to resuspend the cells in acetic acid, albeit not impossible. The extraction regimen was therefore optimised in an attempt to make it quicker and improve the yield, and to ease the procedure. The extraction process was carried out as summarised in Figure 37A, using either manual mixing or a 0.5-1 minute ultrasonication step to resuspend the cells.
To verify the effect of the heating treatment on the extraction procedure, the extraction, using ultrasonication, was carried out with or without the heating step, as summarised in Figure 37B p. 141.

SDS PAGE analysis in Figure 37A shows that Mut^{+His} r-amelogenin (27kDa) was successfully extracted in acetic acid. The similarity of the protein profiles observed and the amount of 27 kDa protein obtained (as indicated by similar band staining intensity) indicated that ultrasonication did not obviously alter the protein content, nor the yield of r-amelogenin extraction. These observations contrast with those of Svensson Bonde and Bulow (2012) who suggested that sonication may not be “very feasible in large scale” and may reduce the purity of the amelogenin extract obtained (Svensson Bonde and Bulow, 2012). In this study, ultrasonication provided a quicker means of resuspending the cells without affecting the quality of r-amelogenin extraction and therefore it was adopted as the mixing regimen.
Analytical SDS PAGE, shown in Figure 37B, showed the protein profiles obtained when fresh cells were resuspended in acetic acid, mixed and either heated to 75°C or kept at room temperature for 20 minutes. Gel densitometry analysis indicated that with the heat treatment, the 27 kDa band intensity was 33% stronger in the supernatant and 66% weaker in the pellet obtained compared respectively with the 27 kDa band intensities in the supernatant and pellets obtained by the procedure without heating. This indicates that the heating step significantly increased the yield of r-amelogenin extraction. However, analytical SDS PAGE also indicated that heating may increase the extraction of other contaminants as well, although not significantly. This was considered to be non-problematic, since the r-amelogenin extract would be subjected to further purification using a panel of secondary purification techniques (strategy summarised in Figure 40, p. 146) as described in the next Sections.

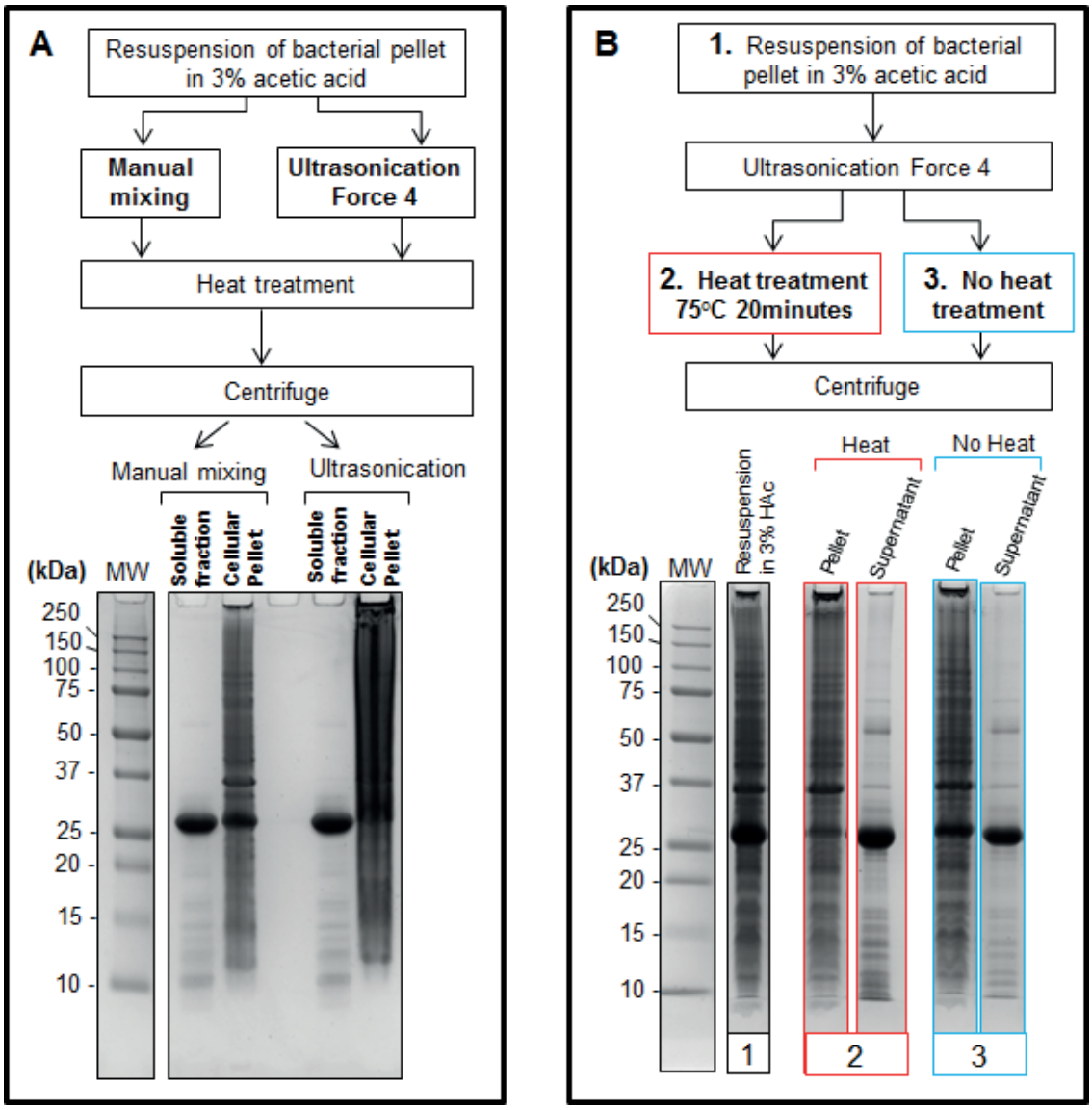


Figure 37 Panel (A) summarises the optimisation of the mixing regimen for acetic acid extraction of r-amelogenin. The effects of ultrasonication and manual mixing on the resulting acetic acid extract were compared. SDS PAGE showed similar protein profiles in the acid soluble fractions for both mixing regimens used, indicating that ultrasonication did not affect the yield nor the quality of extraction of r-amelogenin. Panel (B) shows the effect of heat treatment on the yield. SDS PAGE showed the contents of the initial cell resuspension in acetic acid (lane labelled 1, black frame), of the final pellets and acetic acid-soluble fractions obtained with the heating step ('Heat', lanes labelled 2, red frames) or without heating ('No Heat', lanes labelled 3, blue frames). This experiment showed that heating the mixture to 75°C significantly increased the yield of extraction of r-amelogenin in acetic acid.

3.1.1.2.3 Large scale acetic acid extraction of r-amelogenin using the optimised methodology

The extraction process using 3% acetic acid was optimised specifically for the mouse WT^{+His} and Mut^{+His} r-amelogenins. This series of optimisations provided a method for large scale preparation, summarised in Figure 38 below. These findings are mostly consistent with those of Svensson Bonde and Bulow (2012). For scaling up, longer sonication time (2-3 minutes) was needed in order to obtain a homogeneous mixture. The acetic acid extract obtained was lyophilised, desalted, lyophilised again and analysed by mass spectrometry (Figure 39, p. 145).

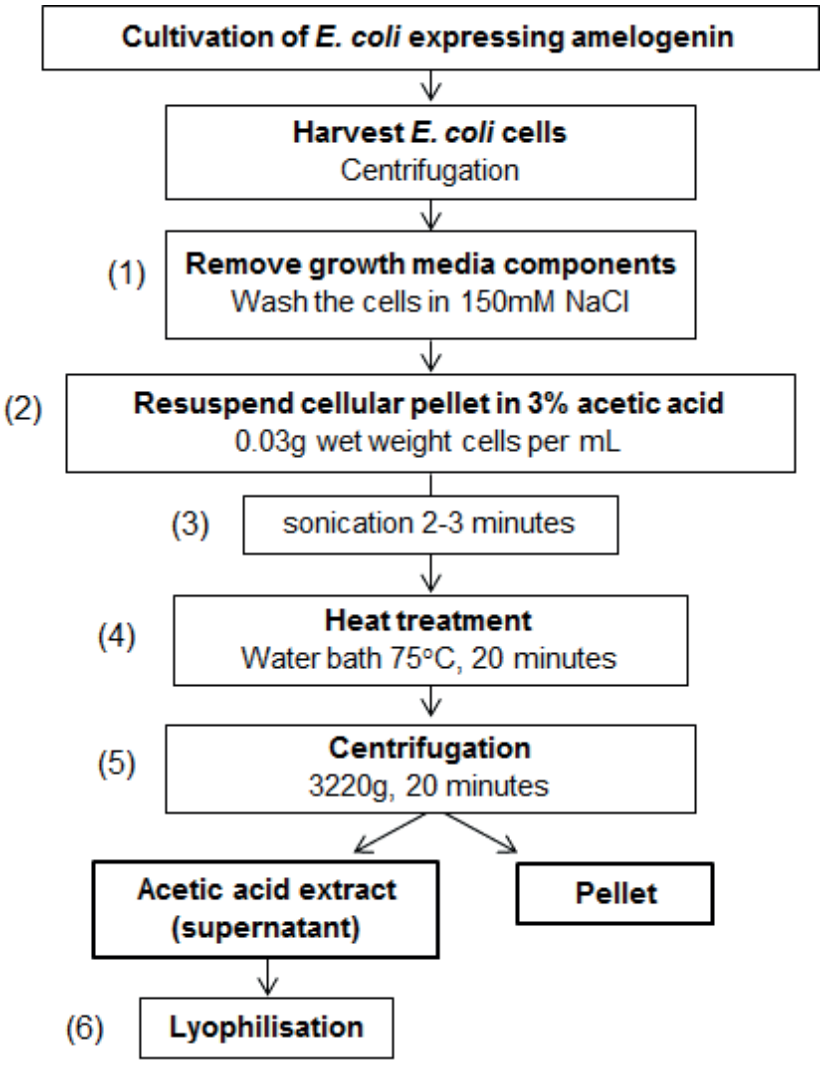


Figure 38 Flow-diagram of the optimised extraction procedure. (1) *E. coli* were washed in 150 mM NaCl and (2) resuspended in 3% acetic acid at 30 mL/g wet weight of cells. The suspension was (3) mixed by sonication and (4) heated at 75°C for 20 minutes. (5) The acid soluble proteins were separated from insoluble residues by centrifugation at 3220 g for 20 minutes and (6) the supernatant containing the acid soluble proteins was lyophilised.

The spectrum obtained (Figure 39, p. 145) for the lyophilisate (desalted) showed peaks at 24732.99, 24748.33, 24764.22, 24779.83, 24795.63, 24812.20 Da. The peak at 24732.99 Da possibly corresponds to the WT^{+His} r-amelogenin which has lost a methionine; its MW predicted on ProtParam (Ahirwar et al., 2015) is 24864.42 Da. The methionine lost is likely Met^1 as *E. coli* endogeneous methionine aminopeptidases can remove N-terminal methionine (Gibbs, 2001). The presence of the other peaks at 24748.33, 24764.22, 24779.83, 24795.63, 24812.20 Da indicate additions of 16 Da, which can be caused by oxidations at various degrees (from 1 to 5 additions of 16 Da). For clarity, the hypothesis is summarised in Table 8.

Table 8 Interpretation of mass spectrometry spectrum (peaks observed, Figure 39) obtained from the desalted lyophilisate). In the hypothesis that additions of 16 Da correspond to oxidations, the number of oxidations are referred to as "+ … Ox."

r-amelogenin state	Predicted MW	Total MW observed on spectrum
Full length	m_0 = 24864.42 Da	None
Full length –Met^1	m_1 = 24733.23 Da	24732.99 Da
[Full length –Met^1] +1x Ox.	= m_1 + 16 Da	24748.33 Da
[Full length –Met^1] +2x Ox.	= m_1 + 32 Da	24764.22 Da
[Full length –Met^1] +3x Ox.	= m_1 + 48 Da	24779.83 Da
[Full length –Met^1] +3x Ox.	= m_1 + 64 Da	24795.63 Da
[Full length –Met^1] +5x Ox.	= m_1 + 80 Da	24812.20 Da

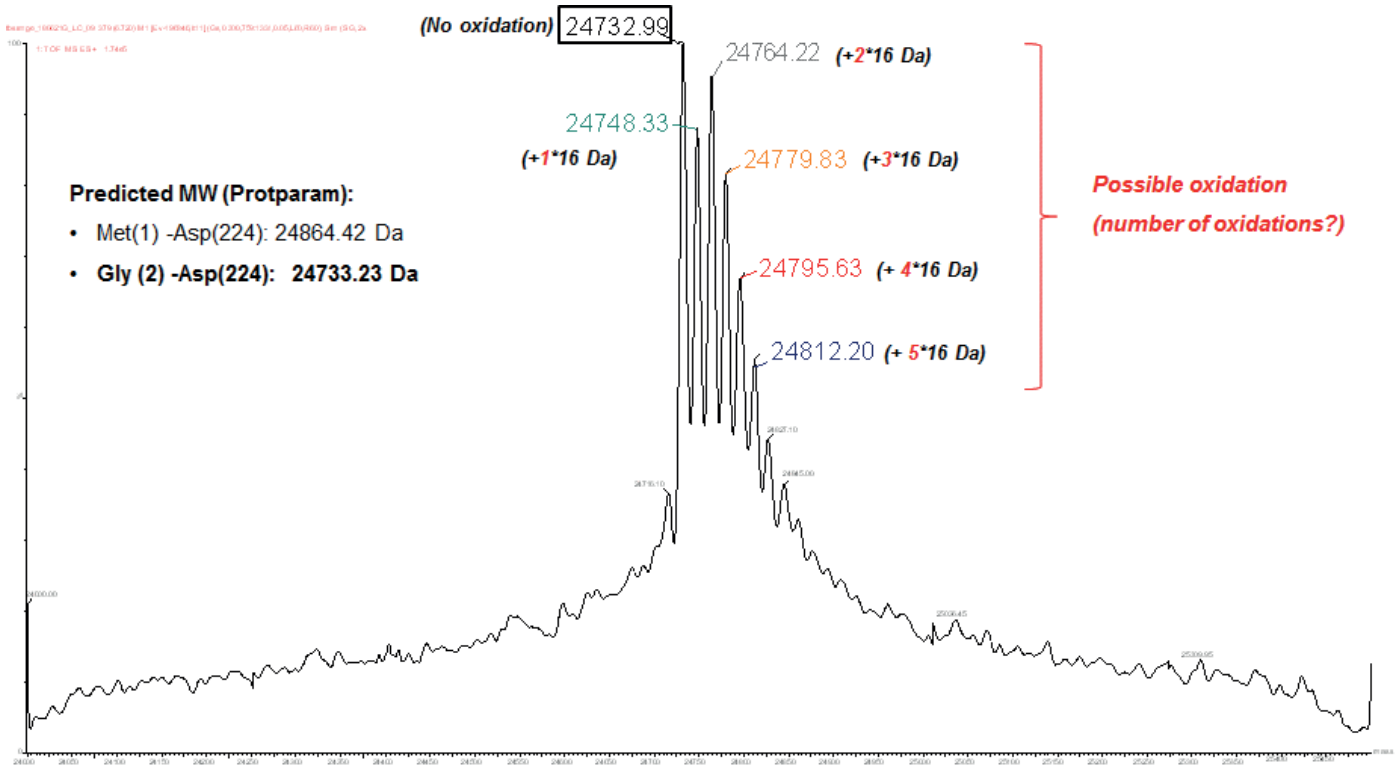


Figure 39 Mass spectrometry analysis of the acetic acid extract of r-amelogenin (WT^{+His}) obtained from the optimised extraction procedure (summarised in Figure 38, p. 143). See text above (p. 144) for further description of the spectrum. The coloured numbers indicate the peaks corresponding to the different masses (in Da) detected. These masses differ by 16 Da, suggesting that there may be different degrees of oxidation. ProtParam tool[1] was used to calculate the theoretical mass of r-amelogenin WT^{+His} minus Met[1].

[1] ProtParam (Gasteiger, 2005) is a protein analysis tool accessible on the ExPASy server (https://web.expasy.org/cgi-bin/protparam/protparam)

3.1.2 Strategies for secondary purification to obtain purified r-amelogenin for binding studies

Following optimisations as described above, the procedure for r-amelogenins extraction in 3% acetic acid was scaled up as summarised in Figure 38 (p. 143). This process resulted in a significant enrichment of r-amelogenin. However, acid-soluble bacterial proteins were still visible (with contaminants of 10 – 75k Da visible on SDS PAGE with Coomassie Blue staining, see Figure 37). Secondary purification was therefore needed to remove these and to isolate r-amelogenin at a suitable purity for future use in binding assays. A panel of standard protein purification techniques was tested and eventually optimised (Sections 3.1.2.1 – 3.1.2.3). The purification strategy is summarised below in Figure 40.

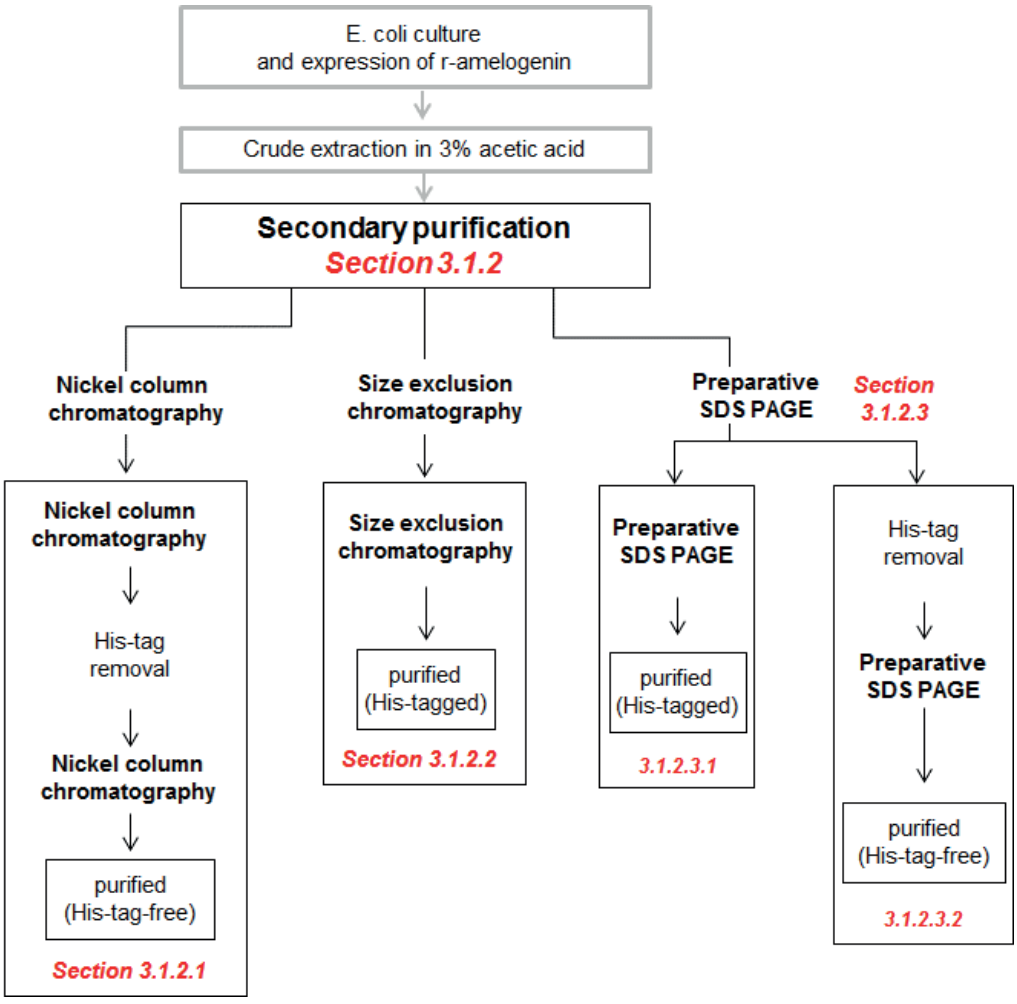


Figure 40 Secondary purification of r-amelogenin from 3% acetic acid extract. Summary of the strategy developed to determine the optimum method providing with r-amelogenin at a suitable purity for future binding studies.

3.1.2.1 Purification of r-amelogenin using nickel column chromatography

As shown previously (see Section 3.1.1.2.2, pp. 140 - 144), acetic acid extracts were enriched with r-amelogenin but still contaminated with bacterial proteins (seen on analytical SDS PAGE as bands ranging 10 - 75 kDa; Figure 37). Further purification steps were therefore trialled to reduce these bacterial contaminants. The first method tested was that of nickel column affinity chromatography.

The results presented below detail the purification of r-amelogenin using standard two-round nickel column chromatography. The two-round nickel column chromatography procedure is summarised in Figure 41 overleaf.

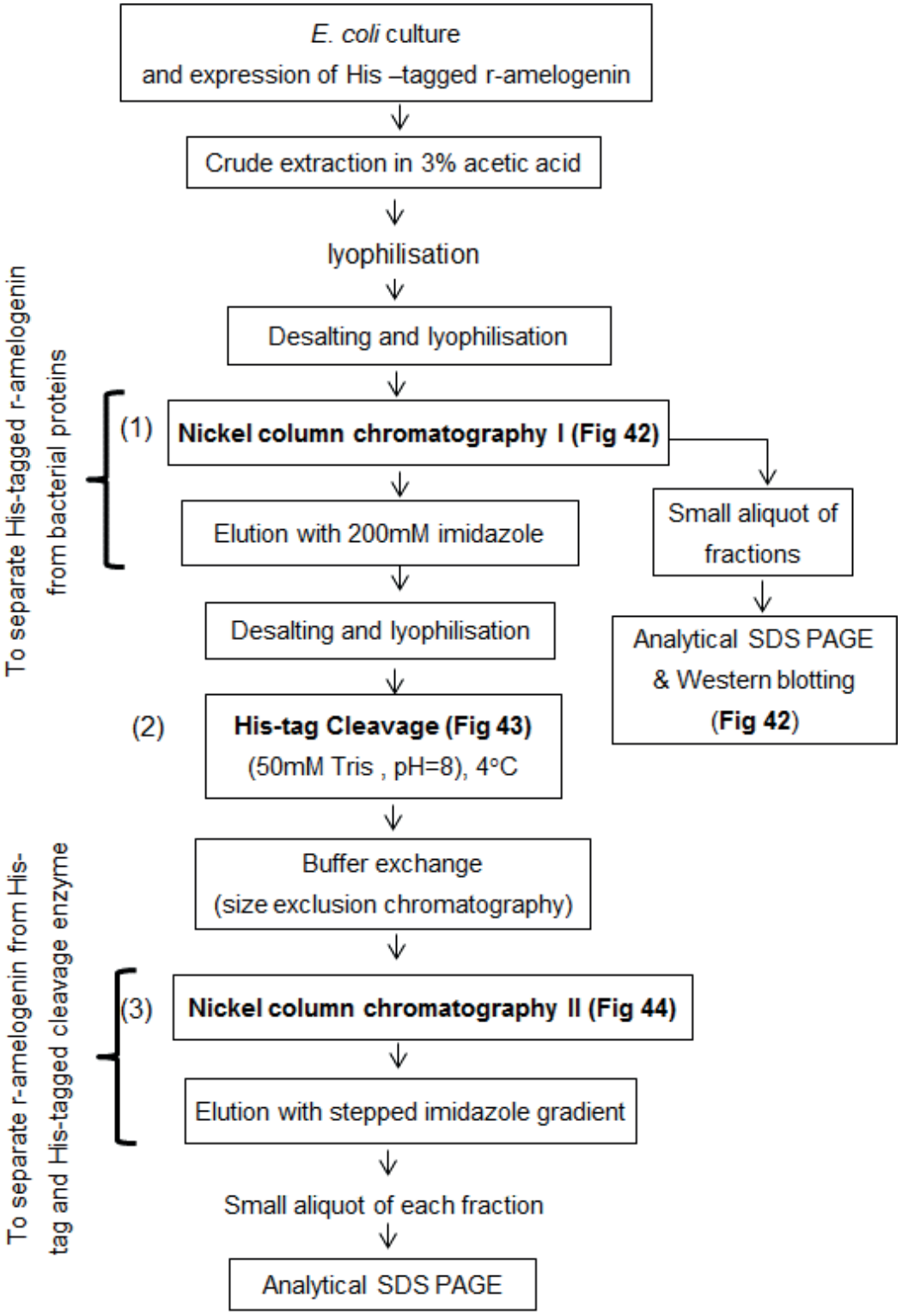


Figure 41 Purification of r-amelogenin using two rounds of nickel column chromatography. The results obtained from steps (1), (2) and (3) are shown in Figures 42, 43 and 44 as indicated on the flow chart.

3.1.2.1.1 First round of nickel column chromatography to accomplish an effective “clean-up”

The acetic acid extract (produced by the procedure summarised in Figure 38, p. 143) was desalted and subjected to a first round of nickel column chromatography to purify His-tagged r-amelogenin from any bacterial contaminants co-extracted in the acetic acid. To summarise, the technique employs immobilised nickel ions in a column through which the sample is pumped. His-tag has a high affinity for nickel (Jeyachandran et al., 2009) and His-tagged proteins (i.e. r-amelogenin) are retained on the column while other proteins are washed out in the flow-through. The His-tagged protein is subsequently recovered by increasing the concentration of imidazole in the buffer passing through the column which competitively displaces the His-tagged protein from the nickel allowing for elution and collection.

The chromatogram obtained using this method is shown in Figure 42A along with Coomassie Blue-stained SDS PAGE analysis showing the protein content of the starting material (the crude acid extract) and the protein content of the two peaks collected during chromatography. SDS PAGE of the starting material showed the clear presence of a 27 kDa component that corresponded to the molecular weight of His-tagged r-amelogenin along with a number of contaminating bacterial proteins. The first 2 fractions collected after the sample was injected onto the column (Fr 1 and Fr 2) corresponded to the flow-through of proteins that failed to bind to the column in 20 mM imidazole. SDS PAGE of these two fractions confirmed that the flow-through comprised the bulk of bacterial contaminants. On increasing the imidazole concentration to 200 mM, a second peak was eluted and collected in Fr 3, Fr 4 and Fr 5. Analytical SDS PAGE showed that these fractions (especially Fr 3) were highly enriched in r-amelogenin, migrating at 27 kDa. Figure 42B reproduces the SDS PAGE analysis described above to allow it to be compared to a corresponding anti-amelogenin western blot. The intense immuno cross-reactivity detected at 27 kDa confirmed that protein migrating at this molecular weight was r-amelogenin. The cross-reactivity above 27 kDa was presumably due to the presence of His-tagged r-amelogenin homodimers migrating at 54 kDa.

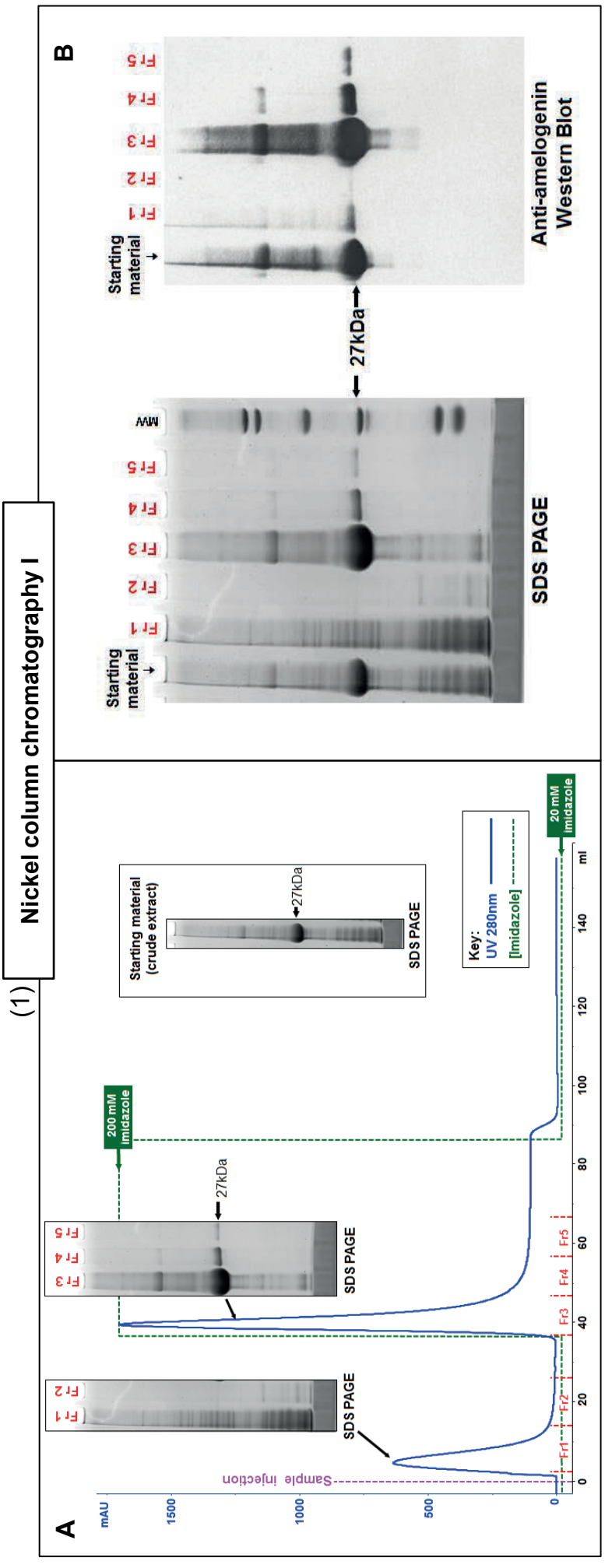

(1)
Nickel column chromatography I
A
mAU
1500
1000
500
0
Sample injection
0
20
40
60
80
100
120
140
ml
Fr1
Fr2
Fr3
Fr4
Fr5
20 mM imidazole
200 mM imidazole
Key:
UV 280nm
[Imidazole]
Starting material (crude extract)
27kDa
SDS PAGE
Fr 1
Fr 2
Fr 3
Fr 4
Fr 5
B
Starting material
MW
27kDa
SDS PAGE
Anti-amelogenin Western Blot

Figure 42 Nickel chromatography round 1: Isolation of His-tagged r-amelogenin from crude acid extracts using nickel column purification (A) SDS PAGE data (inset) showed the protein composition of the acetic acid extract obtained from the *E. coli*. The extract contained a 27 kDa protein corresponding to the molecular weight of His-tagged r-amelogenin together with a range of contaminating bacterial proteins. The chromatogram and accompanying SDS PAGE analysis of the two peaks obtained showed that the column flow through collected in Fr 1 and Fr 2 was comprised mainly of contaminating bacterial proteins that failed to bind the column in 20 mM imidazole. Increasing the imidazole concentration to 200 mM caused the immediate elution of protein that was collected in Fr 3-Fr 5. Accompanying SDS PAGE of these fractions indicated that the protein eluted was highly enriched in the 27 kDa protein that corresponds to His-tagged r-amelogenin. (B) SDS PAGE analysis reproduced from Gabe et al. (2017)[1] but shown alongside the corresponding anti-amelogenin western blot. Intense immune-staining at 27 kDa indicated that this band is comprised of His-tagged r-amelogenin. The cross-reactivity above 27 kDa is presumably due to the presence of His-tagged r-amelogenin dimers and other aggregates. The figure has been adapted from Gabe et al, 2017.

[1] Figure 42 is reproduced from Gabe et al. (2017) under the terms of the Creative Commons license CC BY (https://creativecommons.org/licenses/by/4.0/).

3.1.2.1.2 Second round of nickel column chromatography to isolate cleaved r-amelogenin

The His-tagged r-amelogenin eluted in Fr 3 (see Figure 42, p. 150) was subjected to His-tag cleavage by incubation with HRV3C protease (method described in Section 2.1.2.1.2, p. 109). The ratio of enzyme-to-substrate was previously optimised (Appendix C) and 3 µL of stock HRV3C protease per mg substrate was identified within the range of minimum amounts of enzyme resulting in a maximum yield. The cleavage was carried out in 50 mM Tris-HCl, pH=8, previously shown to be an effective cleavage buffer (Appendix D).

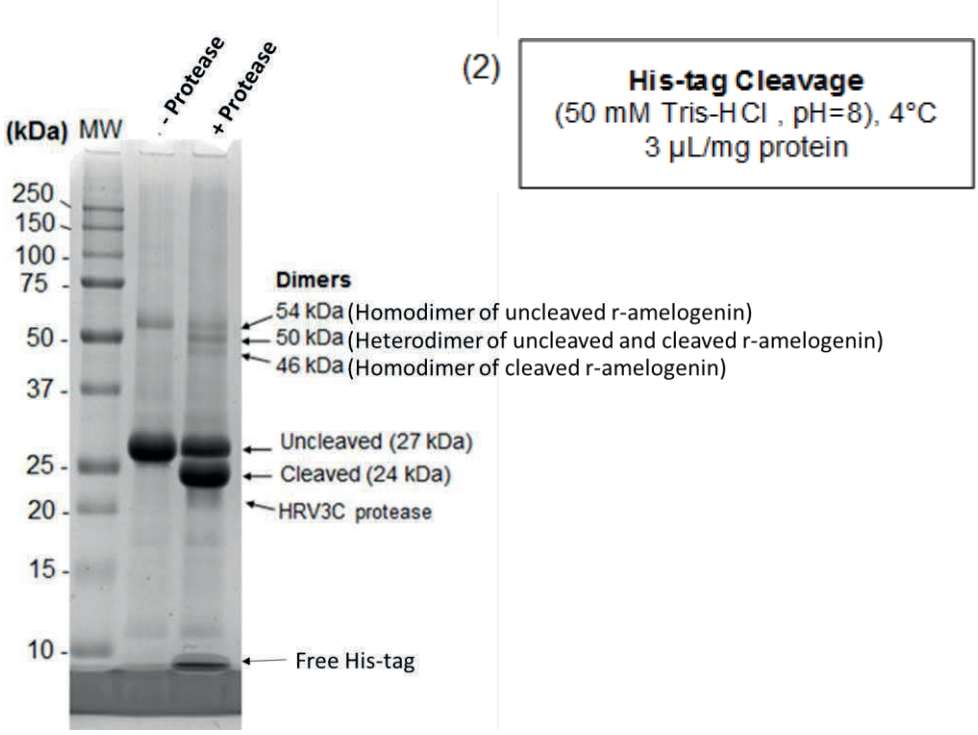


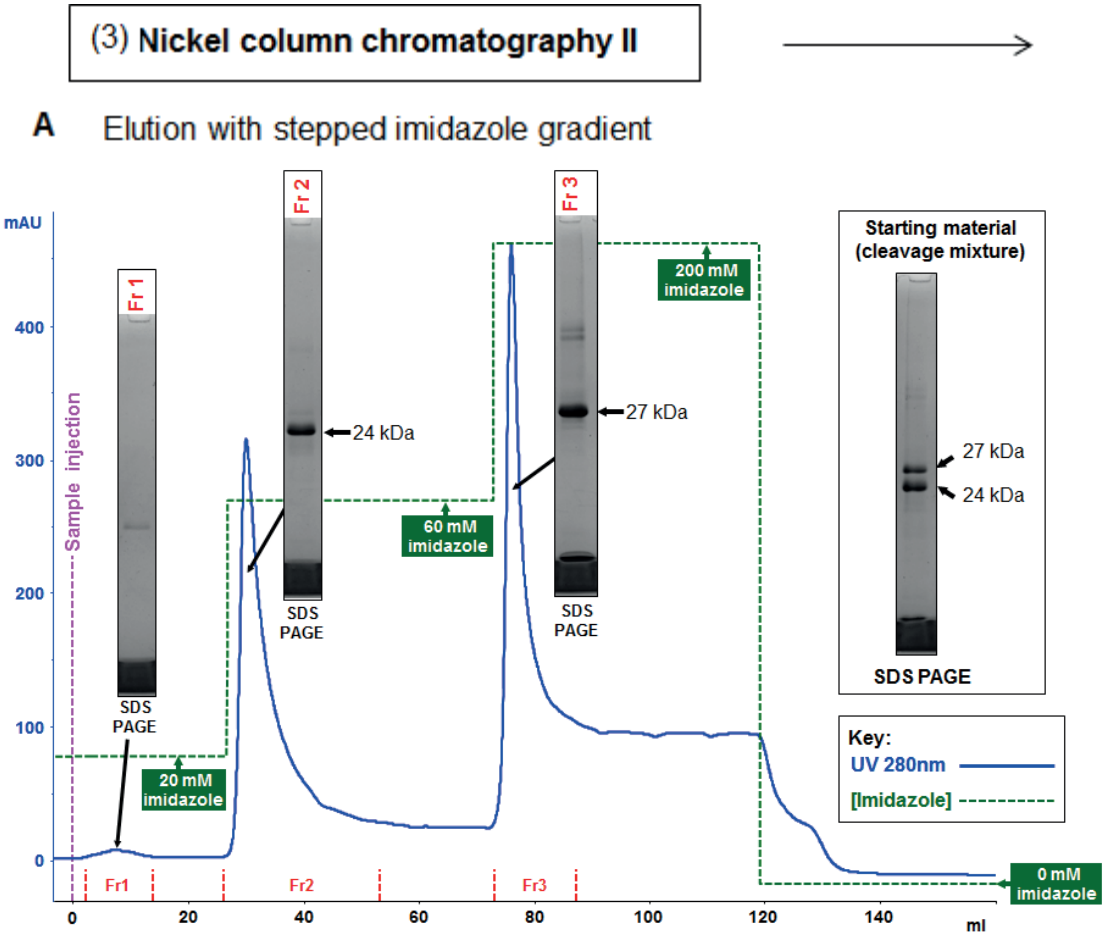
Figure 43 SDS PAGE showing the efficacy of His-tag cleavage from r-amelogenin using HRV protease over 24 hours. The 27 kDa uncleaved His-tagged r-amelogenin was cleaved with approximately 50% efficiency to a generate His-tag-free r-amelogenin cleavage product at around 24 kDa. The cleavage reaction mixture also contained an additional three bands (46, 50 and 54 kDa) possibly representing homodimers of cleaved and uncleaved r-amelogenins and heterodimers of cleaved and uncleaved r-amelogenin respectively. The Figure has been adapted from Gabe et al, 2017[1]

[1] Figure 43 is adapted from Gabe et al. (2017) under the terms of the Creative Commons license CC BY (https://creativecommons.org/licenses/by/4.0/)

Analytical SDS PAGE with Coomassie Blue staining (Figure 43) showed that the His-tag cleavage was only 50% efficient, as both 24 kDa (‘cleaved’) and 27 kDa band (‘uncleaved’) r-amelogenins were present in the cleavage mixture. The apparent homodimer migrating at 54 kDa was observed in the uncleaved sample but after cleavage, another two bands migrating at 50 and 46 kDa were evident. The 46 kDa and 50 kDa bands could correspond to homodimers of r-amelogenin and heterodimers of cleaved and uncleaved r-amelogenin respectively. The cleaved sample also contained a band migrating below 10 kDa which corresponds to the cleaved His-tag. Following incubation with HRV3C protease, the reaction mixture was buffer-exchanged into 20 mM imidazole nickel column chromatography binding buffer ready for reloading on the nickel column for a second round of nickel column purification. The second round of nickel column chromatography was carried out to isolate cleaved r-amelogenin (now freed of its His-tag) from the remaining uncleaved His-tagged r-amelogenin, the cleaved His-tag and the HRV3C protease (which itself is His-tagged). The cleaved r-amelogenin was the only protein in the mixture without a His-tag and in theory at least, should not have bound to the nickel column but instead should have eluted in the flow-through.

Figure 44A below shows the chromatogram and Coomassie Blue-stained SDS PAGE analysis of the fractions obtained. SDS PAGE of the starting material showed the presence of the cleaved and uncleaved r-amelogenin migrating at 24 and 27 kDa respectively. The initial flow-through collected in Fr 1 contained only a trace of the expected cleaved r-amelogenin – clearly the cleaved r-amelogenin was retained on the column even though the His-tag had been cleaved off (the reason for this phenomenon is explored in Section 4.1.2.2.3, p. 221). In an attempt to elute cleaved r-amelogenin, the imidazole concentration was increased to 60 mM; this was an attempt to compete off the cleaved r-amelogenin whilst leaving the uncleaved His-tagged r-amelogenin still bound to the column. Increasing the imidazole concentration in the buffer flowing through the column to 60 mM immediately resulted in the elution of protein which was collected in Fr 2. SDS PAGE of Fr 2 showed that 60 mM imidazole was able to elute the 24 kDa cleaved r-amelogenin while leaving the uncleaved r-amelogenin still bound to the nickel column. The uncleaved r-amelogenin (and the cleaved His-tags) was subsequently eluted from the column as expected by increasing the imidazole concentration to 200 mM.

Figure 44B shows each fraction presented in Figure 44A run on a single gel to make comparisons between each fractions easier to see. The analytical SDS PAGE graphically demonstrated that the methodology was able to separate cleaved (Fr 2) from uncleaved r-amelogenin (Fr 3). However, upon staining the gel with silver, a more sensitive stain than Coomassie Blue (Morrissey, 1981), revealed that the uncleaved r-amelogenin (Fr 2) was still contaminated with traces of uncleaved r-amelogenin and possibly HRV3C protease (Figure 44C). Interestingly, Fr 3 contained the bulk of 50 and 54 kDa- molecular complexes which required 200 mM imidazole for elution while Fr 2 contained the 46 kDa complex eluting with 60 mM imidazole . This agreed with the idea that the 50 and 54 kDa- molecular complexes are comprised of heterodimers of the cleaved and uncleaved r-amelogenin and homodimers of the uncleaved r-amelogenin whilst the 46kDa complex is a dimer of the cleaved r-amelogenin (since the presence of a His-tag in the complex increased its affinity for the nickel column).


(3) Nickel column chromatography II
A
Elution with stepped imidazole gradient
mAU
400
300
200
100
0
Sample injection
Fr 1
Fr 2
Fr 3
24 kDa
27 kDa
SDS PAGE
SDS PAGE
SDS PAGE
20 mM imidazole
60 mM imidazole
200 mM imidazole
0 mM imidazole
Fr1
Fr2
Fr3
0
20
40
60
80
100
120
140
ml
Starting material (cleavage mixture)
27 kDa
24 kDa
SDS PAGE
Key:
UV 280nm
[Imidazole]

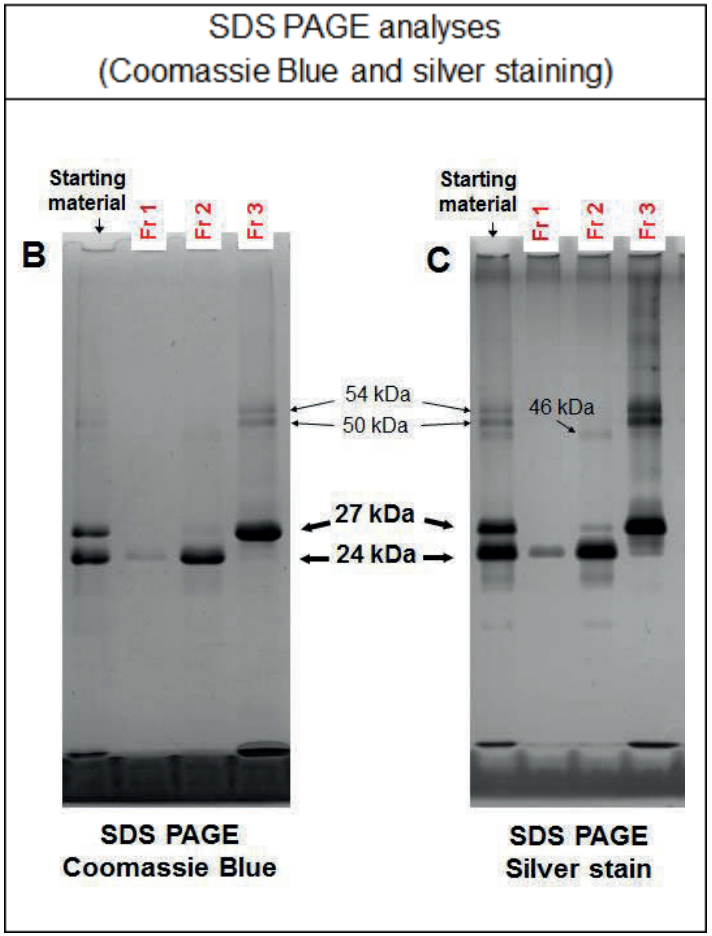

SDS PAGE analyses
(Coomassie Blue and silver staining)
B
C
Starting material
Fr 1
Fr 2
Fr 3
54 kDa
50 kDa
46 kDa
27 kDa
24 kDa
SDS PAGE
Coomassie Blue
SDS PAGE
Silver stain

Figure 44 Nickel column chromatography round 2: Isolation of cleaved His-tagged free r-amelogenin following His tag cleavage using HRV3C protease. (A) SDS PAGE (inset) showed the protein composition of the starting material (comprising of the mixture of 24 kDa cleaved and 27 kD uncleaved r-amelogenins generated by HRV3C protease (Figure 43, p. 153). The chromatogram and accompanying SDS PAGE analysis show that the column flow-through collected in Fr 1 contained very little protein whereas the expectation was that this fraction would contain the His-tag-free cleaved r-amelogenin. Instead, the 24 kDa cleaved r-amelogenin was collected in Fr 2 as a result of increasing the imidazole concentration flowing through the column from 20 to 60 mM. The 27 kDa uncleaved r-amelogenin (still exhibiting a His tag) was eluted later when the imidazole concentration was stepped up to 200 mM. (B) SDS PAGE of the various fractions presented on a single gel for ease of comparison. Silver staining showed that the cleaved 24 kDa r-amelogenin (Fr 2) was not totally pure. Note that the cleaved 24 kDa r-amelogenin fraction (Fr 2) contained an apparent dimer at 46 kDa whereas the uncleaved 27kDa r-amelogenin fraction (Fr 3) contained dimers dominated by species at 50 and 54 kDa. This supports the contention that the 46 kDa species was a homodimer of His-tag-free cleaved r-amelogenins, the 50 kDa species a heterodimer of cleaved and uncleaved r-amelogenins and the 54 kDa species a homodimer of uncleaved r-amelogenins (as the 50 and 54 kDa species will still exhibit a His-tag and bind the nickel column with high affinity). The figure has been adapted from Gabe et al, 2017[1].

[1] Figure 44 is adapted from Gabe et al. (2017) under the terms of the Creative Commons license CC BY (https://creativecommons.org/licenses/by/4.0/).

In summary, using classic two-step nickel column chromatography methodology produced a fraction highly enriched with r-amelogenin, though the fraction was not absolutely pure as judged by silver stained SDS PAGE. Separating the cleaved r-amelogenin was problematic in that it was still bound to the nickel column even without a His-tag and an elevated imidazole concentration (60 mM) was required to elute the protein; this possibly explains the presence of contaminating uncleaved protein that was also eluted in small amounts in the presence of 60 mM imidazole. In addition, the efficacy of the proteolytic cleavage step to remove the His-tag was only about 50% which would have an obvious impact on the yield obtained.

3.1.2.2 Purification of r-amelogenin using size exclusion chromatography

As purification of r-amelogenin based on His-tag affinity chromatography did not generate absolutely pure r-amelogenin, SEC was trialled as an alternative methodology. Using this methodology, separation is based on the molecular weight of the components of interest. The procedure is easily scalable and does not necessitate the use of and disposal issues associated with hazardous chemicals (unlike nickel columns, which are regenerated with highly toxic nickel II solution).

The effect of operational variables such as the column matrix pore size and the column bed height were tested to see the degree of separation achieved. The development of a size-exclusion-based purification method is summarised in Figure 45 overleaf.

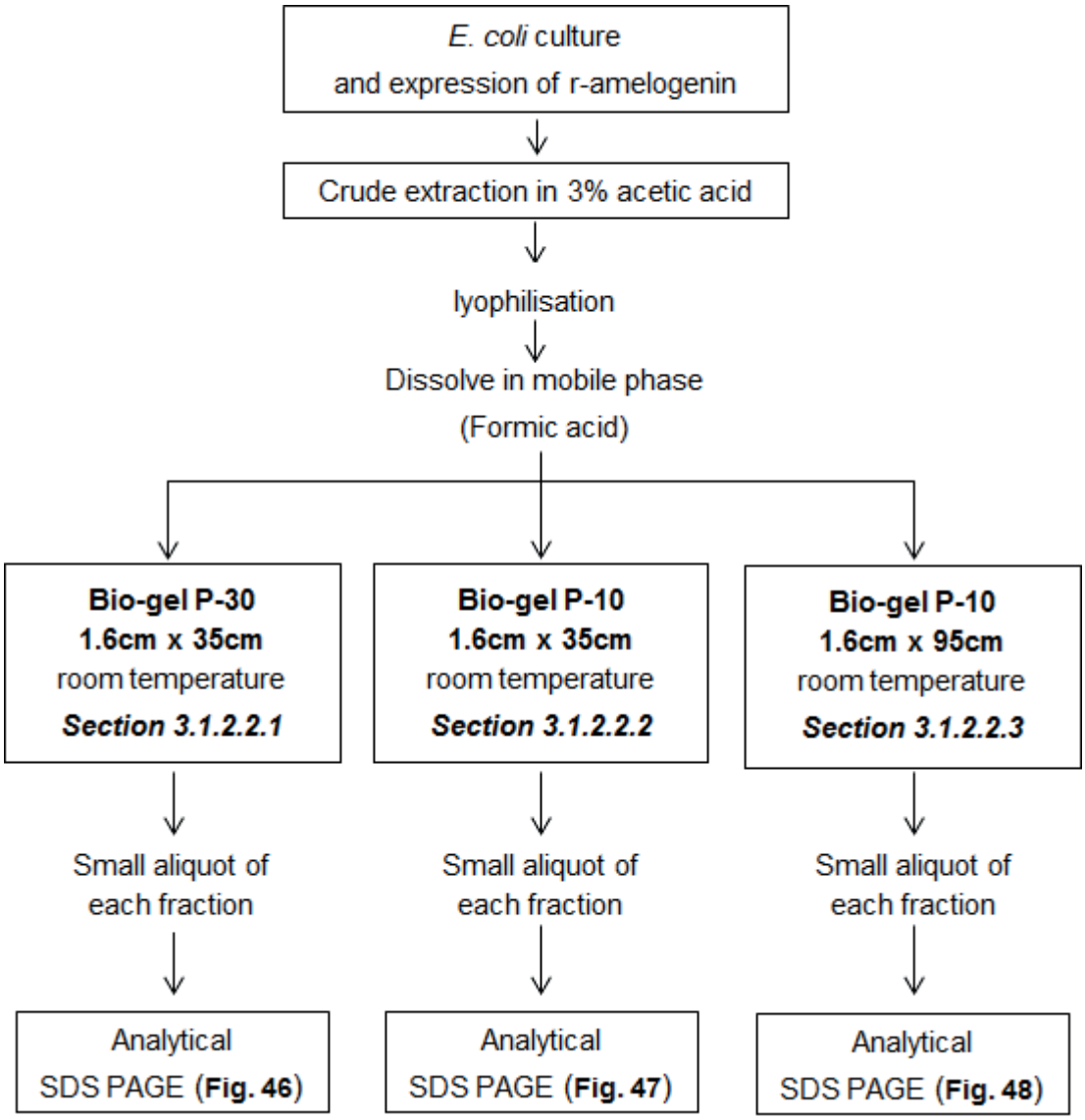


Figure 45 Purification of r-amelogenin using SEC: optimisation of bed height and pore size. Once lyophilised, the acetic acid extract was directly subjected to SEC. For a bed height of 35 cm, two matrices with different pore sizes were tested: Bio-gel P-30 matrix (detailed in Section 3.1.2.2.1) and Bio-gel P-10 matrix (detailed in Section 3.1.2.2.2). The second bed height tested was 95 cm, using Bio-gel P-10 matrix (detailed in Section 3.1.2.2.3). Small aliquots from the fractions collected were taken for analytical SDS PAGE.

3.1.2.2.1 Size exclusion chromatography using Bio-gel P-30 matrix, 35 cm bed height

An attempt to isolate r-amelogenin (WT^{+His}) from the acetic acid extract was made using a XK column (1.6 cm diameter * 35 cm bed height) packed with polyacrylamide matrix Bio-gel P-30. According to the manufacturer, this matrix has a fractionation range of 2.5 – 40 kDa, which brackets the 27 kDa r-amelogenin. Figure 46 below shows the separation of the acid extract components using 125 mM formic acid as the mobile phase. Analytical SDS PAGE with Coomassie Blue staining confirmed that the initial sample was comprised of the 27 kDa r-amelogenin and a number of contaminants (mostly migrating between 7-17 kDa). It was anticipated that the 27 kDa r-amelogenin would elute early and the first UV peak was collected into six fractions. SDS PAGE of these 6 fractions showed that the r-amelogenin at 27 kDa was found predominantly in fractions 1- 3. Fraction 1 contained the r-amelogenin at the highest purity with fractions 2-3 containing increasing amounts of the lower molecular weight contaminants, indicating that Bio-gel-p30 was not able to provide the degree of purification required. The broader peaks eluting later corresponded to small molecules (e.g. amino acids and other metabolites) present in the acetic acid extracts; they were not resolved by analytical SDS PAGE due to their small sizes (data not shown).

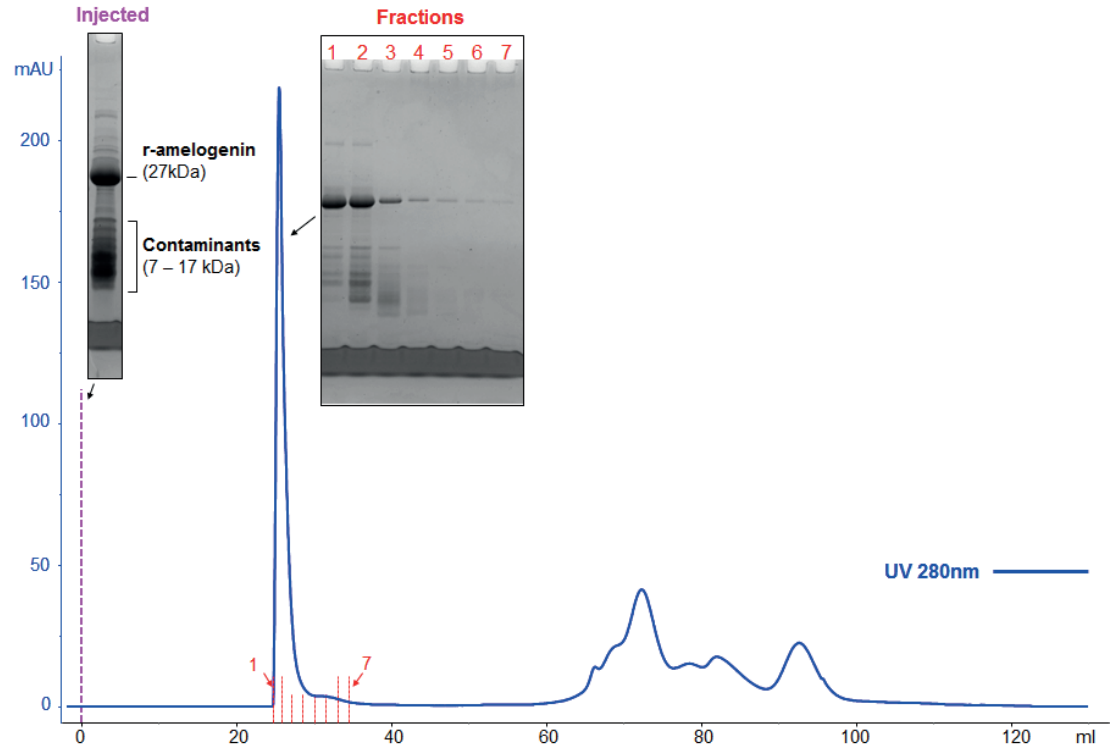


Figure 46 Separation of acetic acid extract components by SEC using Bio-gel P-30 matrix, 35 cm bed height. The chromatogram and accompanying SDS PAGE analysis showed that the r-amelogenin at 27 kDa was found in the first UV elution peak at 25 mL after injection, predominantly in fractions 1-3. Fraction 1 contained the r-amelogenin at the highest purity with fractions 2-3 containing increasing amounts of the lower molecular weight contaminants ranging 7 – 17 kDa. At 60-100 mL after elution, a number of lower UV peaks (below 50 mAU UV value) were visible. This likely corresponded to small UV absorbing molecules (amino acids, metabolites, salts or other non-protein components) that were retained through the column pores, due to their smaller radii.

3.1.2.2.2 Size exclusion chromatography using Bio-gel P-10 matrix, 35 cm bed height

The Bio-gel P-10 polyacrylamide matrix, whose nominal exclusion limit range is 1.5-20 kDa, was expected to improve the isolation of amelogenin away from the lower molecular weight contaminants compared with matrix P-30. In theory, the 27 kDa r-amelogenin should be excluded from the matrix beads and elute quickly in advance of smaller contaminants below 20 kDa which would be retained as they diffuse into the matrix beads during the chromatographic run. The result is shown in Figure 47 below. Analytical SDS PAGE with Coomassie Blue staining showed that the 27 kDa r-amelogenin was eluted in the first 4 fractions, corresponding to the first UV peak. Again, the very first fraction, representing the very first protein to elute from the column, contained r-amelogenin in the most pure state but even here, low molecular weight contaminants were visible. Most of the r-amelogenin eluted over the next 3 fractions and contained far more low molecular weight contaminants.
Isolation of 27 kDa r-amelogenin was, however, not improved compared with that observed with Bio-gel P-30 (Figure 46), as the bulk of r-amelogenin was collected with lower molecular weight contaminants.

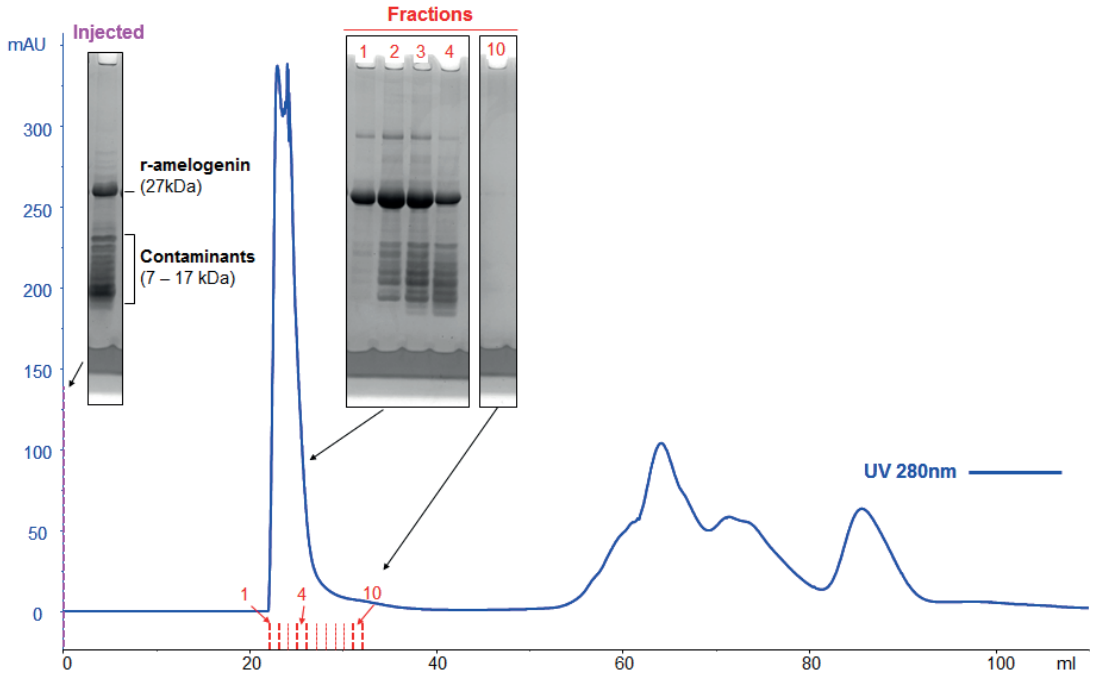


Figure 47 Separation of acetic acid extract components by SEC using Bio-gel P-10 matrix, 35 cm bed height. The chromatogram and accompanying SDS PAGE analysis showed that r-amelogenin at 27 kDa was found in the first UV elution peak, 20 mL after injection and predominantly in fractions 1 – 4. The first fraction contained r-amelogenin in the most pure state but also contained low molecular weight contaminants. The majority of the r-amelogenin eluted over the next three fractions and contained far more low molecular weight contaminants.

3.1.2.2.3 Size exclusion chromatography using Bio-gel P-10 matrix, 95 cm bed height

As shown above, columns of Bio-gel P10 and P30 matrices with 35-cm bed height failed to provide an adequate purification of r-amelogenin. In an attempt to improve the resolution, the bed height was increased to 95 cm. The content of each fraction collected was analysed by SDS PAGE with Coomassie Blue staining, shown below in Figure 48. Increasing the bed height to 95 cm appeared to provide a slightly better resolution of lower molecular weight contaminants (Figure 48) compared with the previous conditions working with 35 cm-bed heights (Figures 46 and 47, pp. 163 and 165), however, the 27 kDa r-amelogenin was still not isolated in a sufficiently pure state. In all of the size exclusion chromatographies tested, the bulk of r-amelogenin was never found at single band purity (Figures 46 - 48). The reasons for such a less-than-ideal performance are discussed later in Section 4.1.2.3, p. 225

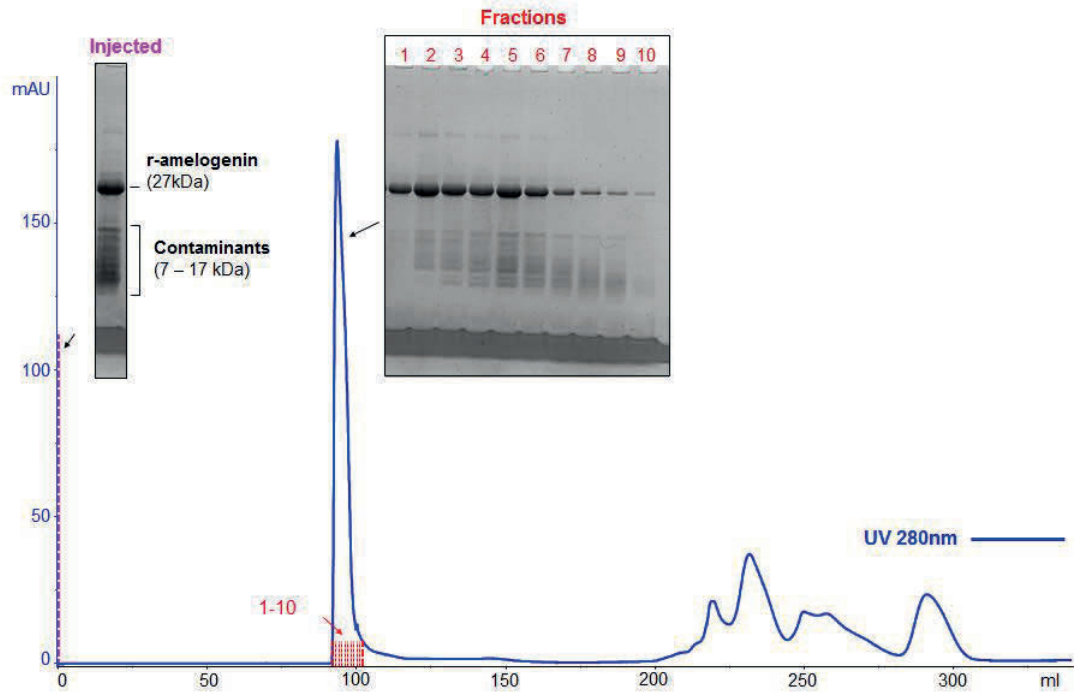


Figure 48 Separation of acetic acid extract components by SEC Bio-gel P-10, 95 cm bed height. The chromatogram and accompanying SDS PAGE analysis showed that r-amelogenin at 27 kDa was found in the first UV elution peak, 20 mL after injection and predominantly in fractions 1 – 9. The first fraction contained r-amelogenin in the most pure state but most of the r-amelogenin was eluted over the next 8 fractions, which contained far more low molecular weight contaminants.

3.1.2.3 Purification of r-amelogenins using preparative SDS PAGE

The purification techniques tested previously (nickel column chromatography, Section 3.1.2.1 and SEC, Section 3.1.2.2) failed to isolate r-amelogenin to single band purity by analytical SDS PAGE (as determined by Coomassie Blue and silver staining) from the acetic acid extract. This may be due interactions of r-amelogenin with the column matrices, or the aggregative nature of r-amelogenin, possibly interacting with the contaminants (discussed later in Sections 4.1.2.2.2, p. 221 and 4.1.2.3, p. 225). In either case, this would be refractory to achieving a good separation. Since analytical SDS PAGE could evidently resolve the 27 kDa r-amelogenin from contaminants (as evidenced by the analytical SDS PAGE gels shown in the previous pages), preparative SDS PAGE was trialled to isolate r-amelogenin away from the bacterial protein contaminants comprised in the crude acetic acid extract, according to their apparent molecular weight. Experiments were therefore carried out (1) to isolate His-tagged r-amelogenin from the acetic acid extract (Section 3.1.2.3.1, summarised in Figure 49), and (2) to isolate cleaved r-amelogenin (Section 3.1.2.3.2, summarised in Figure 53) using preparative SDS PAGE. Preparative SDS PAGE involves collecting protein fractions as the proteins separating on the gel are deliberately run off the bottom of the gel (an outcome to be avoided at all costs during analytical SDS PAGE).

3.1.2.3.1 Isolation of His-tagged r-amelogenin from acetic acid extracts by preparative SDS PAGE

The crude acetic acid extract obtained as described in Section 3.1.1.2.3 (p. 143) was desalted, resuspended into SDS PAGE sample buffer and loaded on the preparative SDS PAGE gel as described in the Section 2.1.2.3, p. 115 (details on the composition of the gel are provided in Table 4 p. 102). The resulting fractions were collected and analysed using analytical SDS PAGE with Coomassie Blue staining using this technique. The procedure is summarised in Figure 49. Following this method, WT^{+His} r-amelogenin ready for cleavage was found at single band purity (Figure 50, p. 170) in fractions 25-52.

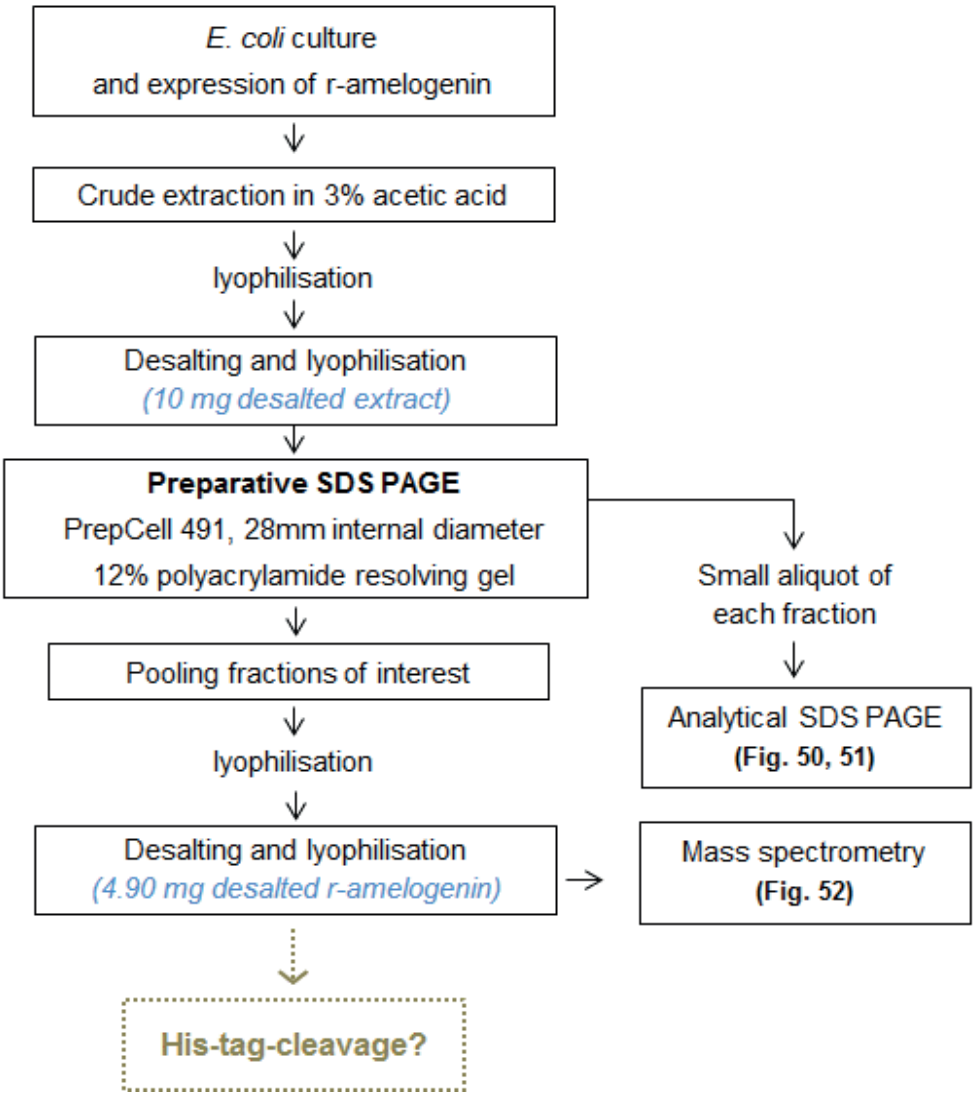


Figure 49 Purification strategy of WT^{+His} r-amelogenin from acetic acid extracts using preparative SDS PAGE (Figure 26 reproduced). The desalted acetic acid extract was resuspended into SDS PAGE sample buffer and heated, to solubilise the proteins. The mixture was loaded on to the preparative SDS PAGE gel to separate its components according to their apparent molecular weights. From the fractions collected, small aliquots were taken and their content was analysed by analytical SDS PAGE with Coomassie Blue staining (Figure 50), silver staining and western blot (Figure 51) to identify the fractions containing r-amelogenin. These fractions were pooled together, desalted and lyophilised. The lyophilisate was subjected to mass spectrometry (Figure 52) and was ready to be subjected eventually to His-tag cleavage and nickel column chromatography, to obtain a purified 'cleaved' r-amelogenin.

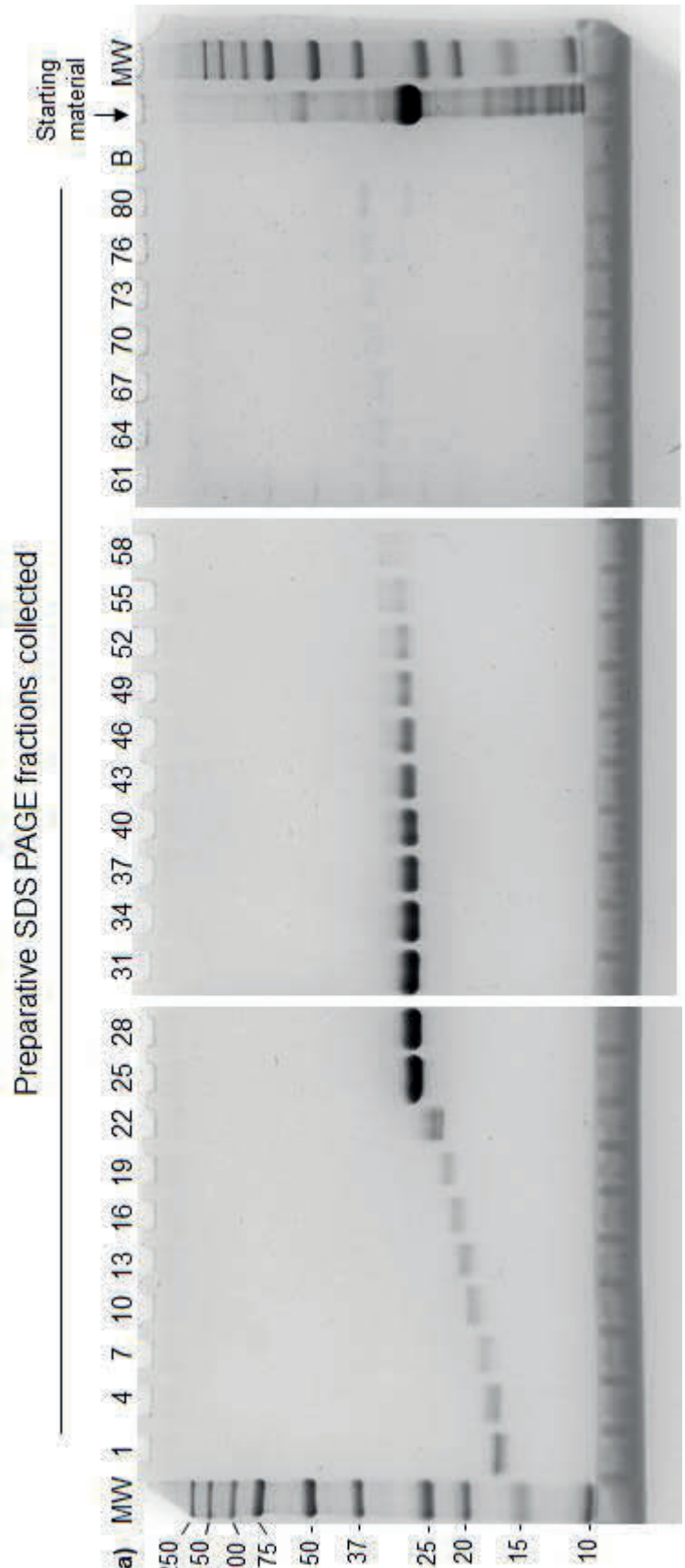


Figure 50 Analytical SDS PAGE with Coomassie Blue staining to identify the fractions of interest (those containing r-amelogenin) collected from preparative SDS PAGE. Prominent bands were in fractions 25 – 52.

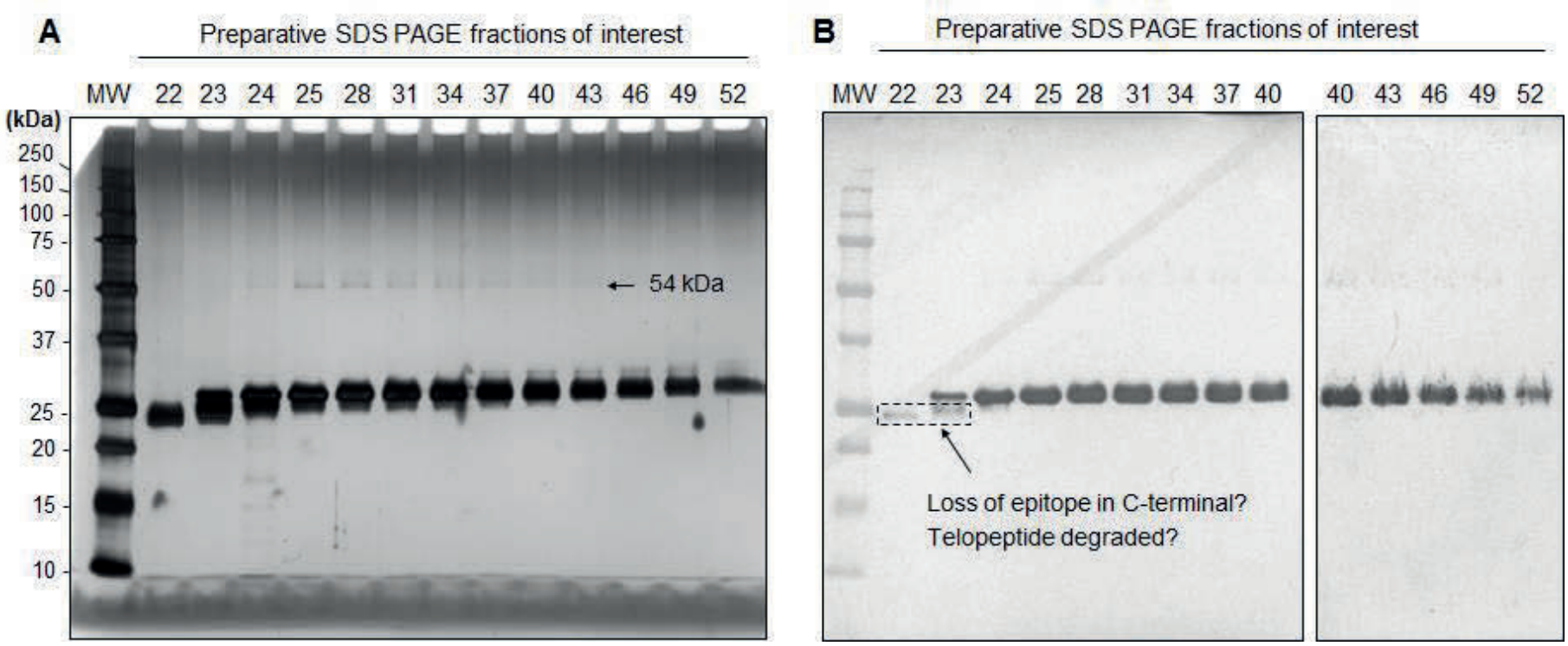


Figure 51 Analytical SDS PAGE of the fractions of interest obtained by preparative SDS PAGE with silver staining (A) and by western blotting using anti-amelogenin antibodies(B). (A) The purity of the samples in fractions 25 to 49 appeared to approach single band purity as judged by silver staining and western blotting. Fractions 22-23 contained silver stained bands (boxed) migrating just below 27 kDa and these bands were only poorly cross-reactive with the anti-amelogenin telopeptide antibody. Bands clearly detected by the antibody all appeared to migrate at 27 kDa. Minor bands at 54 kDa are likely due to dimerisation of the 27 kDa r-amelogenins. (B) The western blot with anti-amelogenin antibodies confirmed the identity of r-amelogenin in the fractions of interest.

After the fractions of interest (fractions 25 – 52) were analysed by analytical SDS PAGE with Coomassie Blue staining (Figure 50, p. 170), the purity and identity of r-amelogenin in the fractions was further assessed by analytical SDS PAGE with silver staining and western blotting (Figure 51 p. 171).

Analytical SDS PAGE with silver staining (Figure 51A) showed single band purity in fractions 25 to 49. The identity of r-amelogenin in these fractions was confirmed by an anti-amelogenin western blot (Figure 51B). Interestingly, strong immuno-cross-reactivity was seen from fraction 24 onwards whereas silver staining detected proteins that appeared to migrate just under 27 kDa in fractions 22 - 23 and were poorly cross-reactive with anti-amelogenin antibody (Figure 51B). Fractions 25 to 49 were then pooled together and desalted. The lyophilisate obtained after the round of preparative SDS PAGE yielded 4.90 mg of purified r-amelogenin from the 10 mg of desalted extract (a yield of ~50%).

The lyophilisate obtained was then subjected to mass spectrometry. The spectrum obtained is shown in Figure 52 p. 173. The spectrum showed peaks at 24749.60, 24763.15, 24780.55, 24812.4. 24827.85, 24844.45, 24860.20, 24876.25, 24892.4 Da. The peak at 24749.60 Da possibly correspond to WT^{+His} r-amelogenin which has lost a methionine (likely Met^{1})[1] and has one oxidation (+16 Da)[2]. The presence of the other peaks at 24763.15, 24780.55, 24812.4. 24827.85, 24844.45, 24860.20, 24876.25, 24892.4 Da indicate additions of 16 Da, which can be caused by oxidations at various degrees (from 2 to 10 additions of 16 Da). For clarity, the hypothesis is summarised in Table 9 (p. 174).

[1] As described in the Section 3.1.1.2.3, the loss of the N-terminal methionine is likely caused by E. coli endogenous methionine aminopeptidase (Ben-Bassat et al., 1987).

[2] The MW predicted on ProtParam (Gasteiger et al., 2005) for r-amelogenin (minus Met^{1}) is 24733.23 Da. One oxidation on a residue adds 16 Da, resulting in a total mass of 23749.23 Da.

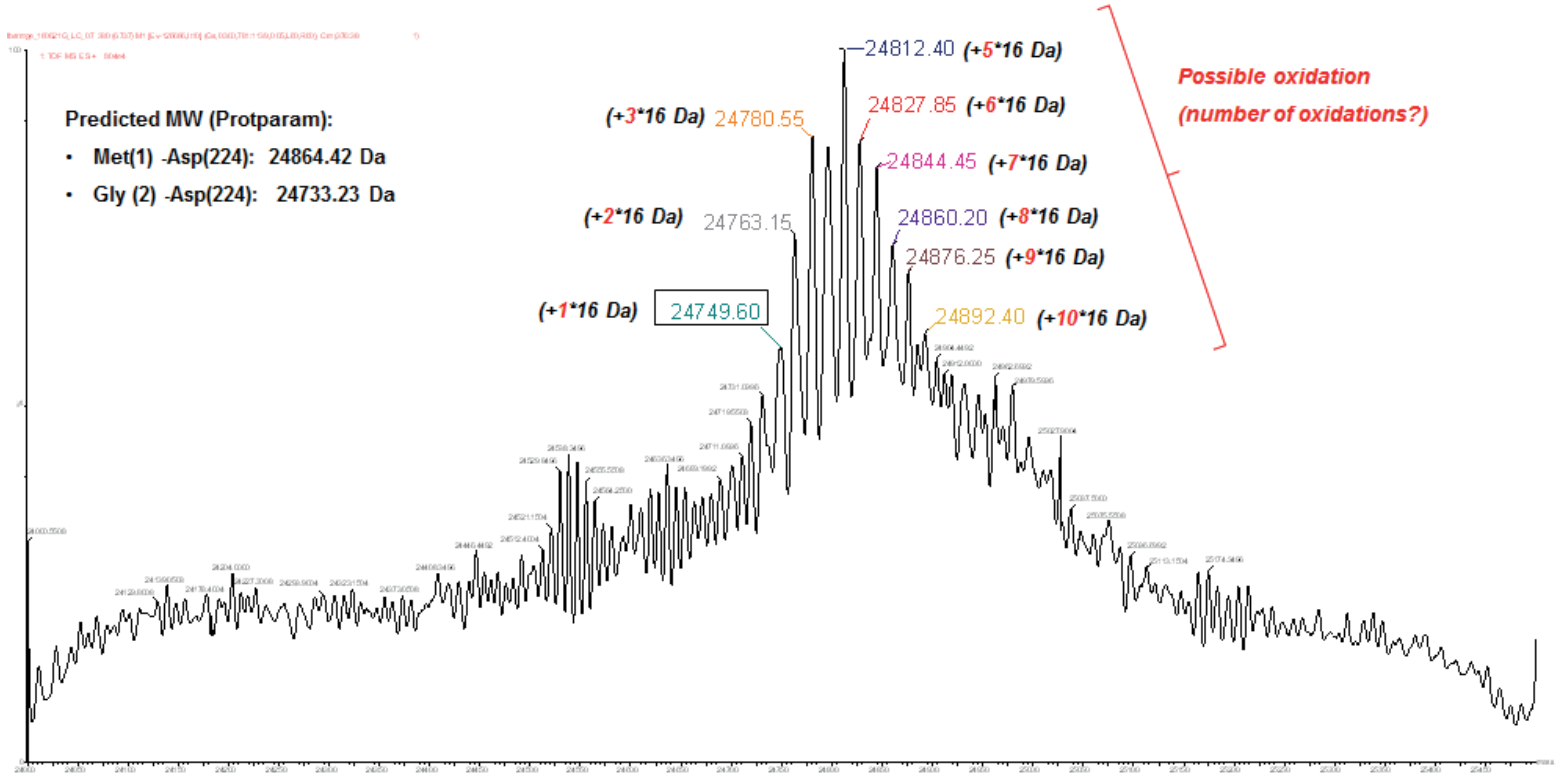


Figure 52 Mass spectrometry analysis of the fractions of interest (25-49) obtained from preparative SDS PAGE, pooled together and prepared as described in Figure 49 (p. 169) of r-amelogenin WT^{+His}. The spectrum showed 10 peaks (labelled). These peaks possibly correspond to the WT^{+His} r-amelogenin with a single methionine depletion with MW predicted: 24733.23 Da on ProtParam tool[1], which was oxidised to various degrees (from 1 to 10 additions of 16 Da).

[1] ProtParam (Gasteiger, 2005) is a protein analysis tool accessible on the ExPASy server (https://web.expasy.org/cgi-bin/protparam/protparam)

Table 9 Interpretation of mass spectrometry spectrum (peaks observed, Figure 52) obtained from the final lyophilisate). In the hypothesis that additions of 16 Da correspond to oxidations, the number of oxidations are referred to as "+ … Ox."

r-amelogenin state	Predicted MW	Total MW observed on spectrum
Full length	m_0 = 24864.42 Da	None
Full length –Met[1]	m_1 = 24733.23 Da	None
[Full length –Met[1]] +1x Ox.	= m_1 + 16 Da	24749.60 Da
[Full length –Met[1]] +2x Ox.	= m_1 + 32 Da	24763.15 Da
[Full length –Met[1]] +3x Ox.	= m_1 + 48 Da	24780.55 Da
[Full length –Met[1]] +5x Ox.	= m_1 + 80 Da	24812.4 Da
[Full length –Met[1]] +6x Ox.	= m_1 + 96 Da	24827.85 Da
[Full length –Met[1]] +7x Ox.	= m_1 + 112 Da	24844.45 Da
[Full length –Met[1]] +8x Ox.	= m_1 + 128 Da	24860.20 Da
[Full length –Met[1]] +9x Ox.	= m_1 + 144 Da	24876.25 Da
[Full length –Met[1]] +10x Ox.	= m_1 + 160 Da	24892.40 Da

3.1.2.3.2 Isolation of cleaved r-amelogenin by preparative SDS PAGE

The previous section described the results when using preparative SDS PAGE to isolate pure His-tagged-r-amelogenin from the crude acetic acid extract. The strategy was that the purified protein would then be subjected to enzymatic cleavage to remove the His-tag (see Figure 49, p. 169) and then the cleaved r-amelogenin could theoretically be purified from uncleaved r-amelogenin using nickel column chromatography. This method would provide a product free of any bacterial contaminants since initial preparative SDS PAGE of the crude acetic extract was highly efficient at delivering purified uncleaved 27 kDa r-amelogenin (see previous Section). However, as evidenced in Figure 44B (p. 156), nickel column chromatography was not 100% efficient in separating cleaved and uncleaved r-amelogenin. Given the exceptional purification power of preparative SDS PAGE (see previous Section), it was therefore decided to develop a modified approach that involved desalting and lyophilising the crude acetic acid prior to subjecting the entire spectrum of proteins present to His-tag cleavage. Preparative SDS PAGE would then be used to purify the cleaved r-amelogenin from the uncleaved r-amelogenin and bacterial contaminants. In effect, this approach would speed up the purification process as the nickel column chromatography steps would be no longer required. The process is summarised in Figure 53, below.

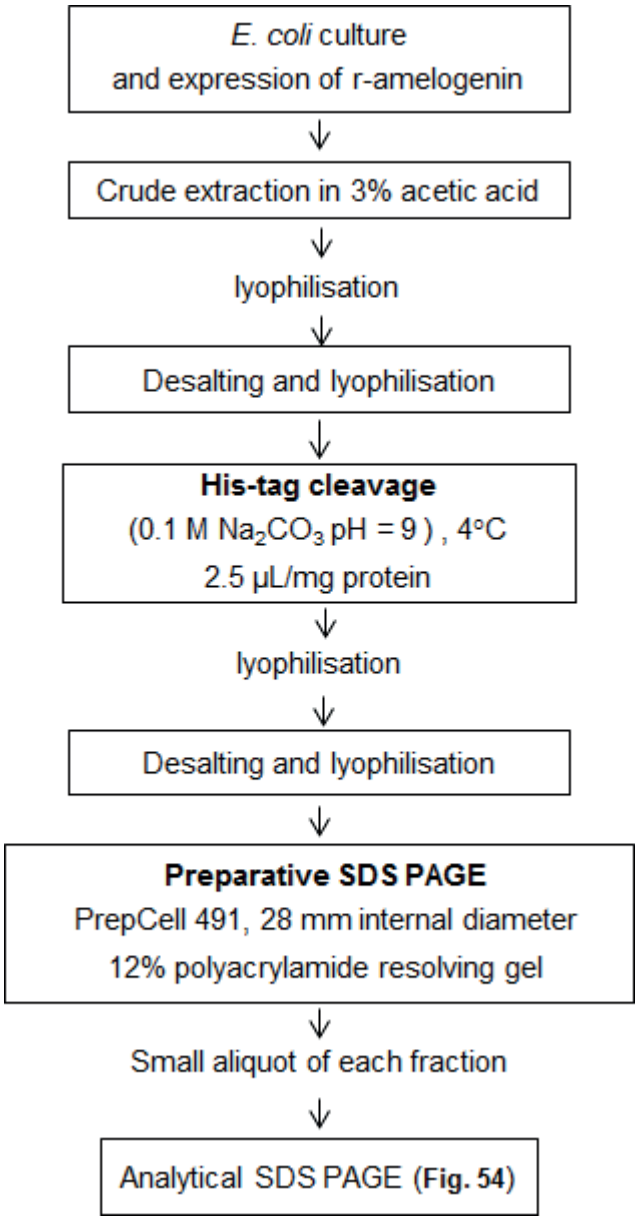


Figure 53 One-step purification of His-tag-free r-amelogenin from acetic acid extracts using preparative SDS PAGE. The methodology carried out is the same as that described in Section 3.1.2.3.1, except that a His-tag cleavage (with subsequent lyophilisation, desalting and lyophilisation) was added. The contents of the fractions obtained from preparative SDS PAGE purification are shown in Figure 54.

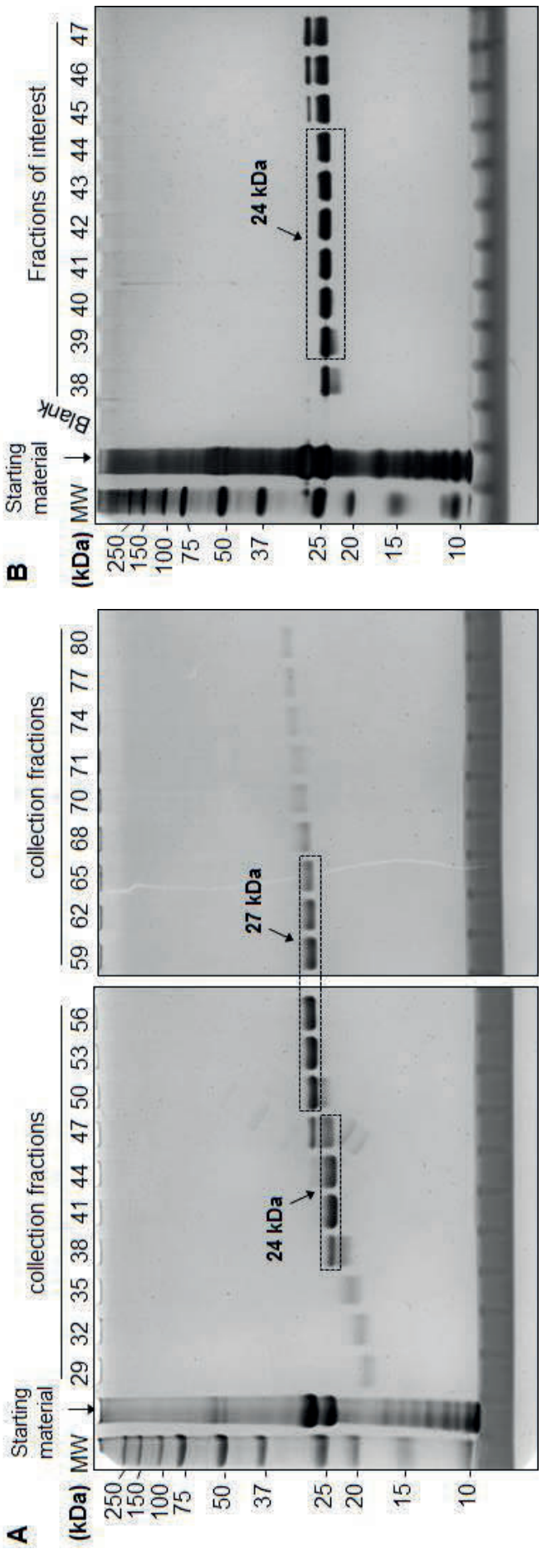


Figure 54 Analytical SDS PAGE showing the isolation of pure 'cleaved' r-amelogenin (24 kDa) from uncleaved r-amelogenin (27 kDa) using preparative SDS PAGE. (A)The cleaved r-amelogenin was found in the fractions 38 - 44 on SDS PAGE with Coomassie Blue staining. (B) SDS PAGE analysis of the fractions of interest with silver staining showed that r-amelogenin was present at single band purity in the fractions 40 - 44.

Analytical SDS PAGE with Coomassie Blue staining of every third fraction collected (Figure 54A) showed that preparative SDS PAGE of the crude acetic extract following cleavage was able to isolate cleaved r-amelogenin migrating at 24 kDa (fractions 38 - 47). Figure 54B, showing silver stained analytical SDS PAGE for all fractions between 38 and 47 shows that contaminating, uncleaved r-amelogenin at 27 kDa began to appear from fraction 45 onwards. Trace amounts of a slightly lower molecular weight contaminant were present in fraction 39, though in negligible amounts and typically fractions 39 to 44 could be pooled, desalted and lyophilised to prepare purified 24 kDa His-tag-free r-amelogenin.

.

3.2 Development of protein-binding assays

In the search for mechanisms underlying AI and possibilities for a therapy, AI was identified as a conformational disease for the first time, using a mouse model carrying the mutation p.Y64H, which phenocopies the human mutation p.P70T (Barron et al., 2010, Brookes et al., 2014). These reports hypothesised that the p.Y64H mutation causes amelogenin to aggregate during trafficking through the ER, which in turn leads to ER stress and ultimately to ameloblast apoptosis (Further description of these studies are provided in Section 1.3.4.1.3, p. 70). Having developed methods for producing purified r-amelogenin, the next aim was to use WT and mutant p.Y64H r-amelogenin to investigate the effect of the mutation on amelogenin-amelogenin interactions *in vitro,* in order to test the above hypothesis.
Several approaches can be used to study protein-protein interactions (e.g. QCM, SPR, ELISA). In the present study, the aim was to develop a cost effective and high throughput microplate-based assay to study amelogenin–amelogenin interactions. Such an assay would not only be useful for studying amelogenin-amelogenin interactions *per se* but would also provide a screening tool for testing potential therapeutics that might modulate any abnormal amelogenin-amelogenin interactions. To this end, attempts were made to develop two binding assays.

3.2.1 Method 1: Fluorescence-based binding assay

The basic strategy for method 1 involved immobilising amelogenin (bait protein) to microplate well surfaces to which was added FITC labelled amelogenin (the free ligand in solution). After incubation, microwells were washed and any amelogenin-amelogenin binding was determined by measuring the fluorescence remaining bound to the microwells (see Figure 29, p. 123). To conserve r-amelogenin during method development, the 20 kDa amelogenin abundant in enamel matrix derivative (EMD – the non-commercial form of Emdogain) was purified by preparative SDS PAGE to be used as a surrogate for initial method development.
The results presented below detail:

(i) FITC-labelling of EMD
(ii) purification of the FITC labelled 20 kDa amelogenin from EMD and
(iii) trialling of FITC-20 kDa amelogenin in the microplate based binding assays

3.2.1.1 FITC-labelling of whole EMD.

Whole EMD protein was dissolved in 0.1 M Na_2CO_3 at pH = 9, and labelled overnight with FITC (Section 2.2.1.1, p. 120). The labelled proteins were subjected to analytical SDS PAGE and visualised by UV excitation. The gels were also stained with Coomassie Blue to confirm the efficiency of the labelling (Figure 55).

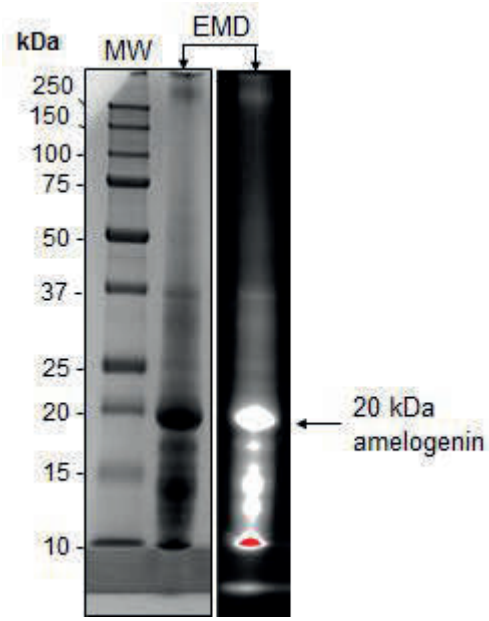


Figure 55 SDS PAGE analysis showing that EMD proteins were successfully labelled with FITC. Coomassie Blue staining and fluorescein-specific fluorescence revealed similar banding patterns.

SDS PAGE analyses as shown in Figure 55 revealed that all proteins in EMD were successfully labelled with FITC, since all bands visible by Coomassie Blue staining were also visible using UV excitation of the fluorescent label.

3.2.1.2 Purification of the 20 kDa amelogenin from FITC-labelled EMD using Preparative SDS PAGE

Following overnight labelling, the FITC reaction mixture was lyophilised, desalted against formic acid, lyophilised and subjected to preparative SDS PAGE. The process is summarised in Figure 56 below. The fractions obtained were analysed by analytical SDS PAGE. UV excitation was used to visualise the bands (Figure 57A), along with Coomassie Blue staining (Figure 57B) and silver staining (Figure 57C).

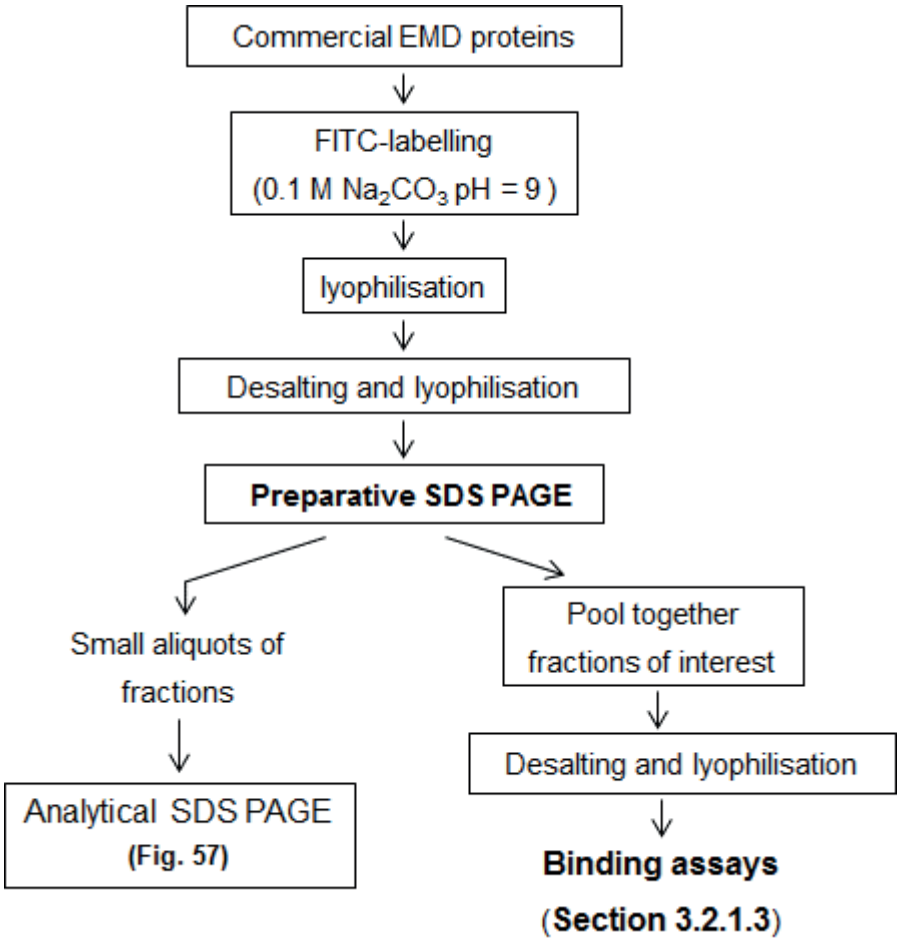


Figure 56 Diagram summarising the preparation and purification of FITC-labelled EMD proteins.

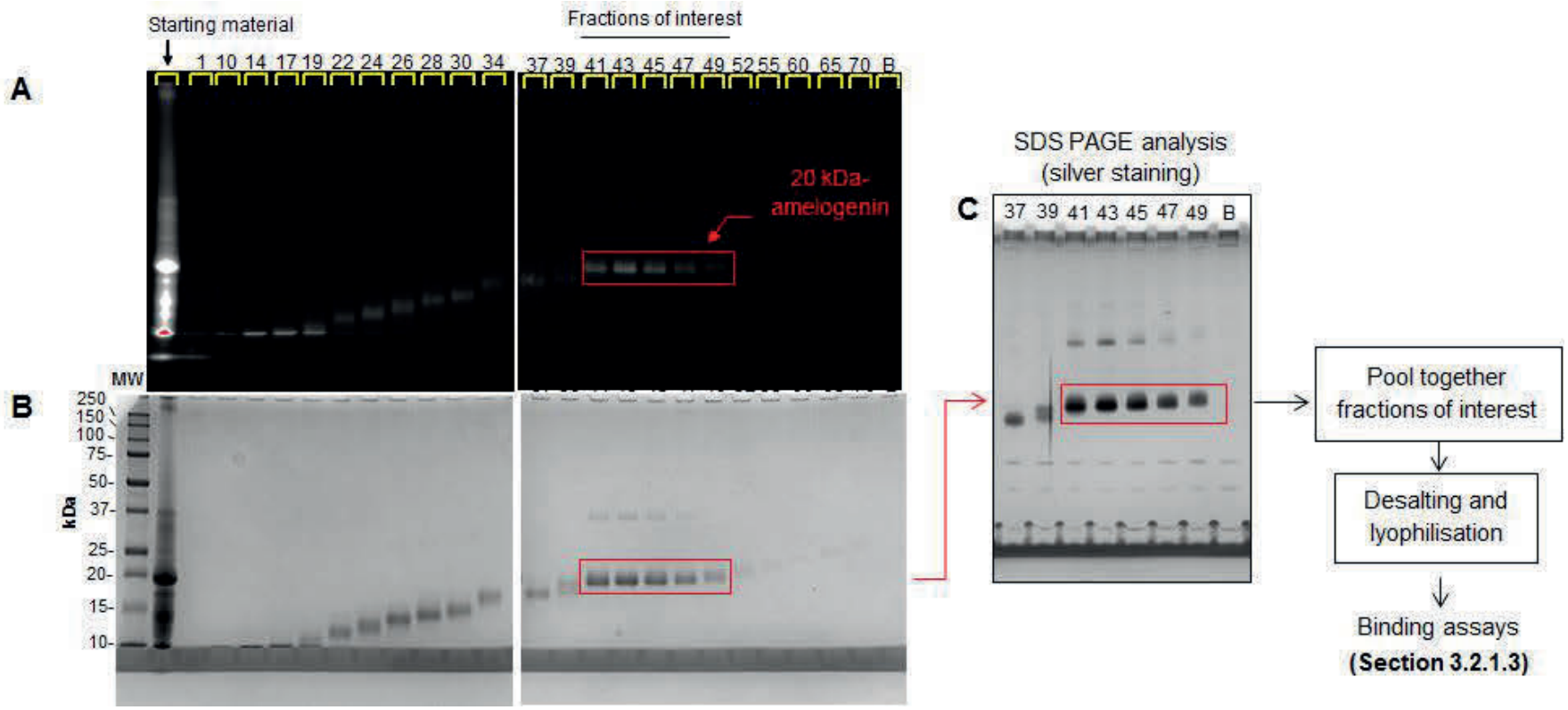


Figure 57 Analytical SDS PAGE showing the fractions obtained when FITC labelled EMD proteins were subjected to preparative SDS PAGE. The contents of fractions 1-70 were visualised by fluorescence (A) and Coomassie Blue staining (B), Fractions of interest (red box) were analysed further by SDS PAGE with silver staining to verify their purity (C). Note that the low molecular weight contaminants on this gel were also present in the blank lane (lane "B") and do not appear to be associated with the actual 20 kDa fractions *per se*. These fractions were pooled together, then desalted and lyophilised, providing the material for future binding assays.

The SDS PAGE analyses in Figure 57 (pictures A-B) showed that on UV excitation and Coomassie Blue staining, the 20 kDa bands appeared well resolved, indicating that the purification process was successful. The fact that the FITC-labelled proteins were visible on fluorescence (Figure 57A) indicated that the purification process did not affect the fluorescence.

The component of interest, the 20 kDa-amelogenin, appeared in fractions 41-49. The purity of these fractions, and that of the neighbouring fractions (37 and 39) was analysed by silver staining of the analytical gel (Figure 57C). In the fractions of interest, the 20 kDa band appeared as the strongest signal by far. Other components were visible at 36 kDa, possibly dimers. Other bands were visible at lower molecular weight in all of the fractions analysed, including the blank lane, which suggested that these bands originated from extraneous contamination that occurred during the preparation of the electrophoretogram (e.g. dirty gel comb) rather than being present in the fractions themselves. The fractions 41 to 49 were then pooled together and desalted for the binding assays.

3.2.1.3 Measuring amelogenin binding between free FITC-labelled amelogenin and immobilised amelogenin as binding partners

For the fluorescence-based binding assay (described in Section 2.2.1.3, pp. 122 - 123), it was necessary to ensure that the microplate wells were saturated with adsorbed bait protein and establish whether the adsorbed proteins could resist subsequent washing, blocking and handling steps. For convenience, FITC-labelled 20 kDa amelogenin (purified from EMD as described previously) was used as a surrogate protein to investigate the binding kinetics of amelogenin to the microwell surfaces and to determine the conditions required to achieve saturation of the microwells. The FITC-labelled 20 kDa amelogenin was dissolved in binding buffer at concentrations between 1 and 20 µg/mL, to determine the concentration required to saturate the microwell surface . To optimise binding conditions, three buffers were compared for binding: Na_2CO_3 at pH = 9, PBS at pH = 7.4 or TBS at pH = 7.2. For each of these buffers, the incubation was carried out overnight either at room temperature or at 4°C. After incubation, the microwells were washed 3 times and blocked with 1% bovine serum albumin (BSA) solution (see Section 2.2.1.3 , pp. 122 - 123) for one hour at room temperature. Afterwards the microwells were washed three times and pure water was added before the fluorescence associated with adsorbed FITC-labelled 20 kDa amelogenin was measured. The process is summarised in Figure 58.

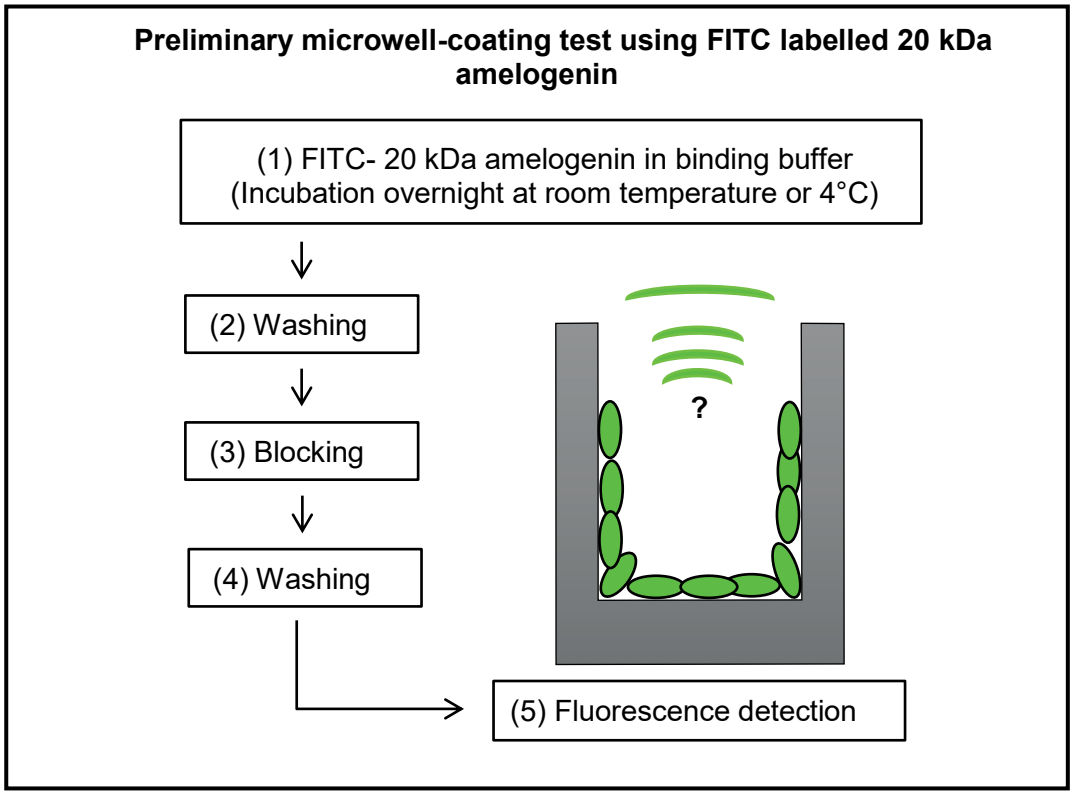


Figure 58 Experimental design of a preliminary test to determine the binding conditions necessary for FITC-labelled 20 kDa amelogenin to adsorb to, and saturate, microwell surfaces. Any fluorescent amelogenin that bound to the microwell surface was detected using a fluorescence plate reader (Figure reproduced from Figure 29).

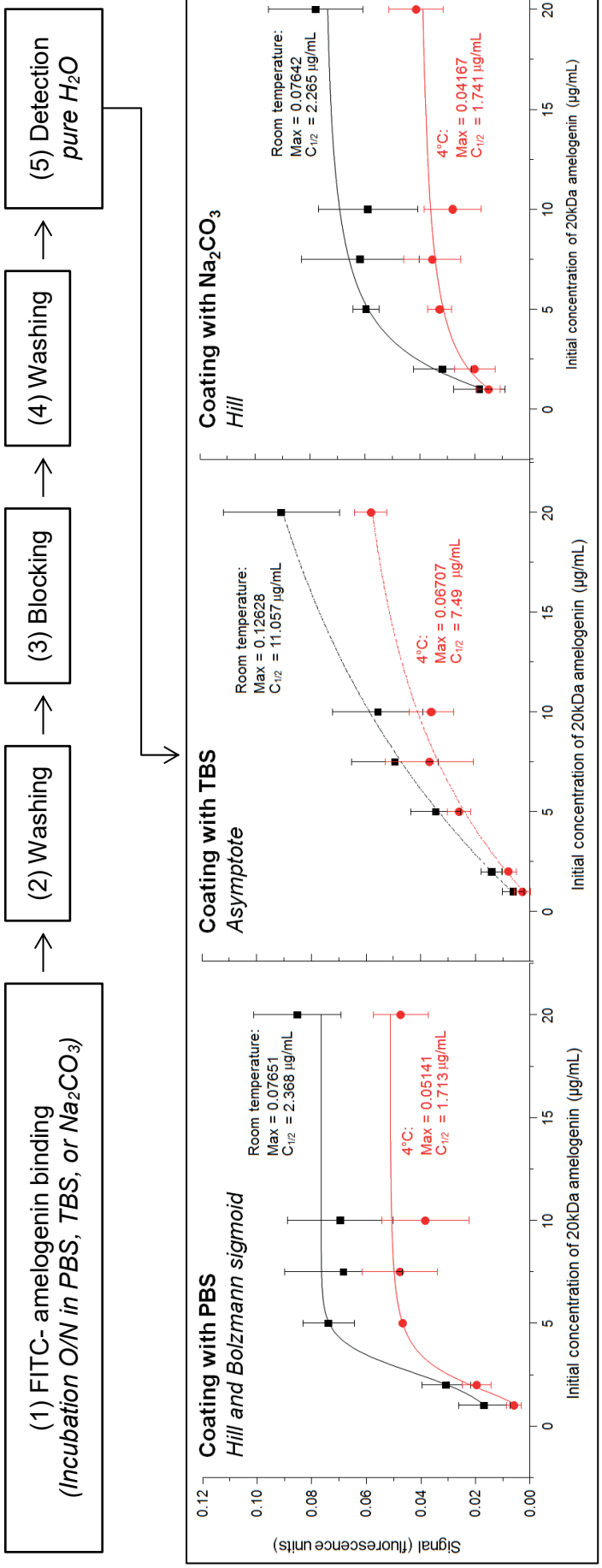


Figure 59 Fluorescence signals obtained at the end of the preliminary coating test were plotted against the initial concentrations of 20 kDa amelogenin in either: PBS, TBS and Na_2CO_3. For all plots, curve-fitting analyses (using OriginPro software) showed that the concentrations ($C_{1/2}$) required to obtain half saturation were lower when the plates were coated at room temperature than at 4°C. Data shows mean ± SD, n=3

After carrying out the incubation, blocking and washing steps (as summarised in Figure 58, p. 185) the fluorescence readings were recorded (Figure 59, p. 186). The data obtained were plotted as averages of 3 repeats. For all conditions, the fluorescence detected increased with increasing initial concentration of the FITC-labelled 20 kDa amelogenin added to the microwells. With PBS (pH 7.4) and Na_2CO_3 (pH 9) based binding buffers, fluorescence approached a plateau at around 5 µg/mL FITC-labelled 20 kDa amelogenin and the amount of protein adsorbed appeared to be greater at room temperature. In contrast, when TBS (pH 7.2) was used as the binding buffer, a clear plateau region, indicative that the microwell surface had been saturated, was not evident and the fluorescent signal continued to increase with increasing initial concentration of FITC labelled 20 kDa amelogenin added to the microwells.
The data were subjected to OriginPro curve-fitting in order to estimate parameters such as the concentration at which half-'saturation' ($C_{1/2}$) occurs and the maximum fluorescence signal achieved. The optimum models for incubations in PBS (pH 7.4) and Na_2CO_3 (pH 9) found were sigmoidal/growth curves (Hill and Bolzmann) while an asymptotic based model best fitted the data for TBS (pH 7.2). The parameters generated by the models are reported in Table 10.

Table 10 Comparison of buffers and incubation temperatures used to coat microwells with FITC labelled 20 kDa-amelogenin. This table compares the parameters generated by the fitting models (Figure 59). $C_{1/2}$ is the concentration at which, in theory, the signal would be half of the saturation value; and the maximum signal, 'Signal max' is the signal at saturation. Standard deviations are specified between brackets below values reported.

Buffer	**PBS** *Bolzmann or Hill*		**TBS** *Asymptote*		**Na_2CO_3** *Hill*	
Temperature	**Room T**	**4°C**	**Room T**	**4°C**	**RoomT**	**4°C**
$C_{1/2}$ (µg/mL)	2.368 (±0.281)	2.211 (±0.138)	10.072 (±5.160)	7.470 (±1.350)	2.265 (±1.099)	1.742 (±1.112)
Signal max (units)	0.077 (±0.019)	0.051 (±0.016)	0.119 (±0.033)	0.067 (±0.005)	0.076 (±0.036)	0.042 (±0.020)

The fluorescence signals were read at neutral pH in pure water which would theoretically maximise the signal obtained as fluorescein fluorescence is inhibited at low pH (Chen et al., 2008). Even so, the signals in this case were very weak (below 0.1 on average), even though labelled protein was supposedly saturating the microwell surfaces. This raised doubts as to whether this method would be sensitive enough to accurately determine the binding of labelled protein free in solution to unlabelled protein immobilised on the microwell surfaces. This, together with the fact that the labelling chemistry protein modifies lysine residues (Maeda et al., 1969), which may

modify protein behaviour, switched focus towards the second method of using the microplate format to measure amelogenin interactions. This second method was developed with the aim of monitoring the interaction of unlabelled protein free in solution to unlabelled protein pre-adsorbed to the microwell surfaces by the reduction in UV absorbance of the microwell contents as the free protein in solution interacted with the immobilised bait protein.

3.2.2 Method 2: Monitoring the kinetics of amelogenin-amelogenin interactions as a function of a reduction in UV absorbance

The second method involved simply incubating unlabelled amelogenin in UV transparent microplate wells and monitoring the decrease in UV absorbance (illustrated in Figure 60 below) as amelogenin initially adsorbed to microwell surfaces (protein-polystyrene binding) followed by binding of amelogenin remaining free in solution to the immobilised amelogenin covering the microwell surfaces (protein-protein binding). Protein concentration in solution is usually monitored spectrophotometrically at 280 nm, which is the absorbance maximum of tryptophan and tyrosine. However, the EMD proteins (mostly amelogenins) used as a model protein are not particularly rich in these residues. Therefore absorbance was measured at 220 nm, which is within the range of wavelengths at which peptide bonds absorb maximally (Goldfarb et al., 1951)

The results presented in this Section detail development of the microplate assay with regards to:

(i) the effect of the starting EMD concentration on the EMD-polystyrene binding equilibrium (Section 3.2.2.1),

(ii) the optimisation of the methodology to investigate the kinetics of EMD protein-protein interactions (Section 3.2.2.2),

(iii) the effect of EMD proteins adsorbed to the bottom of the microwell and attenuating the UV before it passes through the free proteins still in solution potentially confounding the absorbance readings (Section 3.2.2.3)

(iv) attempts to block EMD adsorption to the bottom of the microwells to circumvent the issue of UV attenuation by EMD adsorbing to the bottom of the microwells (Section 3.2.2.4).

(v) an attempt to use the assay to study mutant p.Y64H and WT r-amelogenins interactions (Section 3.2.2.5).

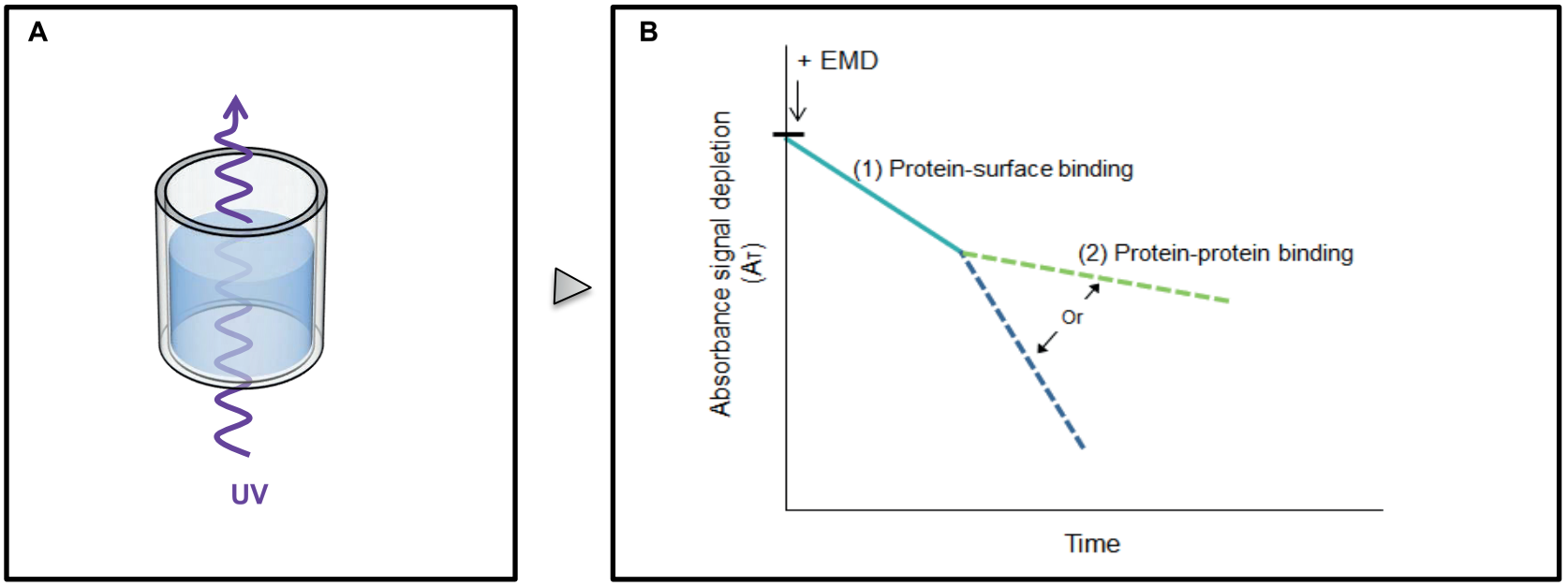


Figure 60 Principles of protein binding kinetics based on protein depletion from solution as detected using UV absorbance (Reproduced from Figure 30). (A) Photometric measurements are based on UV light transmission through a sample contained in a microwell. A UV light source located below the microplate illuminates the protein sample. The light beam passes up through the sample solution and the absorbance is measured by a detector. (B) By Beer-Lambert's law, absorbance will fall as solubilised proteins adsorb to the microwell surface and are taken out of solution (step (1), solid turquoise line). The hypothesis is that a point will be reached when the microwell surfaces become saturated. At this time, proteins would continue to be removed from solution by interaction with the protein already immobilised to the microwell surfaces. This second phase of protein removal from solution may occur at a different rate, resulting in an inflection point (step (2) dashed lines). If protein-protein binding is faster than protein-surface binding, the signal will resemble the dark blue dashed line. If protein-protein binding is slower than protein-surface binding, the signal obtained will resemble the green dashed line.

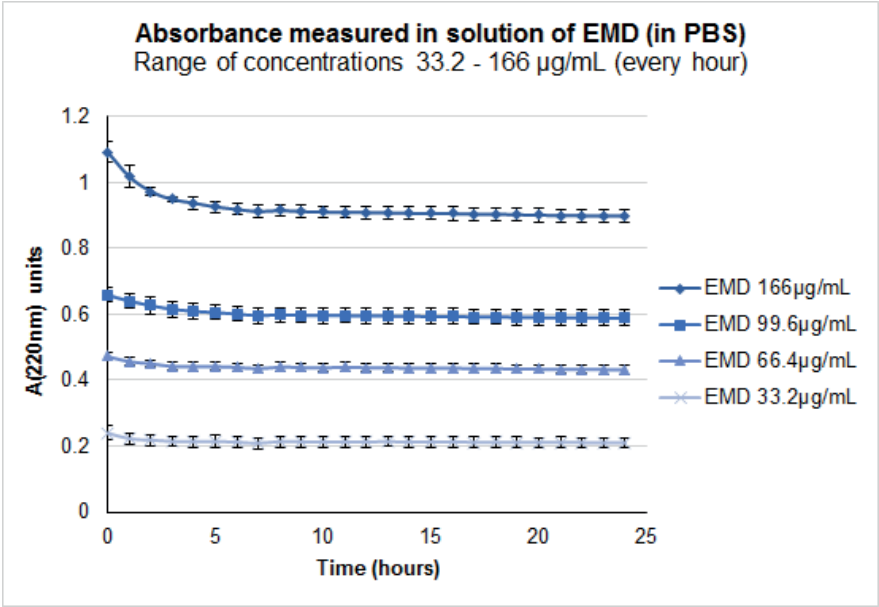
3.2.2.1 Effect of initial EMD concentration on binding equilibrium: determination of a concentration required to saturate the microwell surfaces.

As illustrated in Figure 60 above, the hypothesis behind the microplate assay was to monitor protein interactions between proteins free in solution and proteins already immobilised to the polystyrene microwell surfaces at saturating levels, by tracking the depletion of protein from solution.

To optimise the starting concentration of EMD proteins required to saturate the polystyrene microwell surfaces, freshly prepared EMD solutions in PBS:1% acetic acid (290:1) ranging in concentration between 33.2 and 166 and µg/mL were added into empty microwells and left to incubate for 24 hours with the absorbance monitored at 1 hour time-points. The absorbance measurements (n=6 corrected by blank subtraction) are plotted in Figure 61A[1]. The graph obtained showed that for all concentrations tested, the absorbance decreased over 24 hours and the decrease of absorbance was characterised by a fast drop (in one or two phases) within the first 10 hours of incubation, after which it stabilised, indicating that the binding reactions, whatever their nature (EMD-polystyrene and/or EMD-EMD), had reached equilibrium. SDS PAGE analysis with Coomassie Blue staining of the solutions left in the microwells after 24 hours (Figure 61B) supported the spectrophotometric data in that protein concentration was decreased in these solutions after 24 hours incubation.

[1] Note: The absorbance measurements started 10 minutes after the EMD solution was prepared. As 10 minutes only represent 0.7% of the total 24-hour incubation, the starting time of absorbance measurement was considered to be the starting time of experiment.

A



B

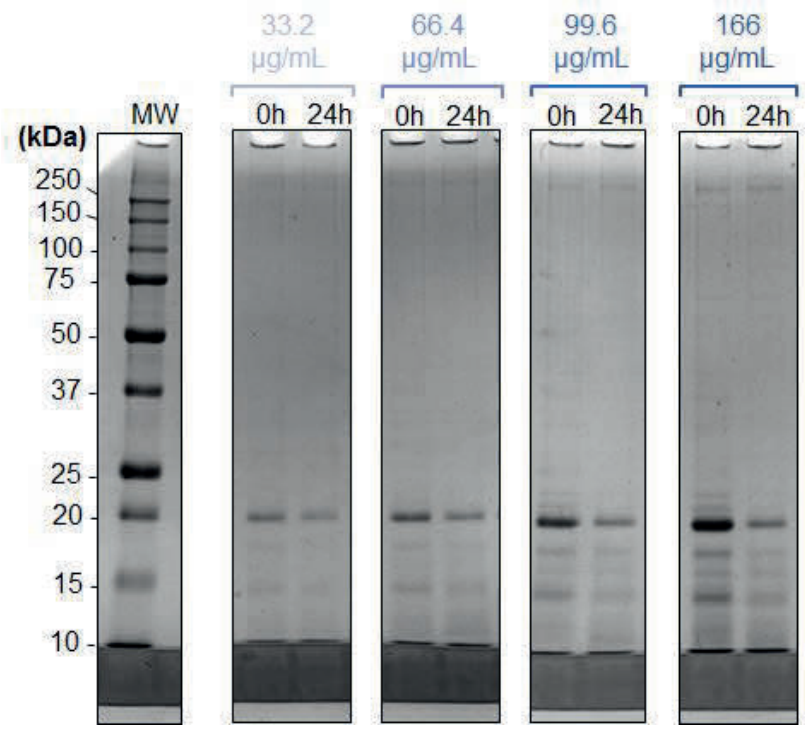


Figure 61 (A) A range of initial EMD concentrations (33.2 – 166.0 µg/mL) was used to investigate EMD adsorption onto the microwell surfaces by measuring the depletion of EMD from solution spectrophotometrically over time. (B) To confirm that proteins were being removed from solution by adsorption to the microwell surfaces, the contents of solutions before incubation ("0h") and at the end of incubation ("24h") were analysed by SDS PAGE. All the incubations were carried out at room temperature, in PBS:1% acetic acid (290:1). Absorbance data shows mean ± SD, n=6.

For the highest starting EMD concentrations (99.6 to 166 µg/mL), the contents of the solutions at the end of incubation showed similar signals on SDS PAGE analysis with Coomassie Blue staining (Figure 61B). This implies that the maximum equilibrium concentration was reached. At lower initial EMD concentrations (33.2 and 66.4 µg/mL), in contrast, the proteins remaining in solution appeared less concentrated on analytical SDS PAGE (Figure 61B) suggesting that equilibriums were reached, but not to saturation. So, within for initial EMD concentrations of 99.6 to 166.0 µg/mL, the polystyrene surface was apparently entirely covered by EMD proteins; this is illustrated below in Figure 62. At 166.0 µg/mL the drop of EMD concentration to the saturation level resulted from the binding equilibrium of EMD proteins still free in solution to those immobilised on the polystyrene surface (EMD-EMD-polystyrene interactions).

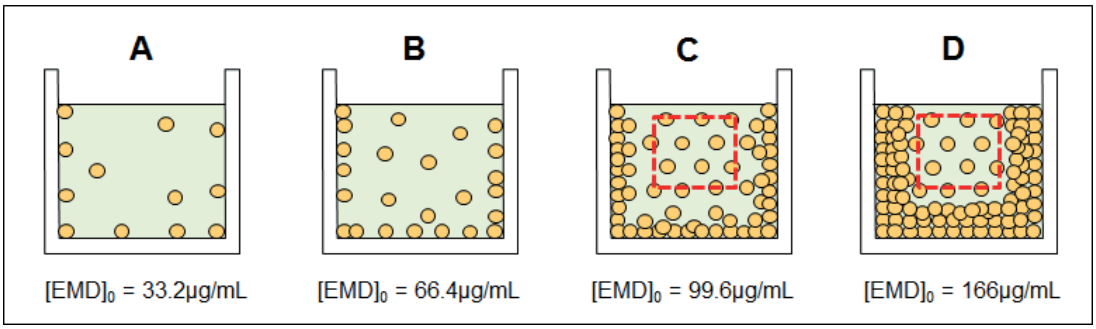


Figure 62 Diagram illustrating the hypothesis in respect of equilibrium reached in the microwells, at initial EMD concentrations of (A) 33.2 µg/mL, (B) 66.4 µg/mL, (C) 99.6 µg/mL and (D) 166.0 µg/mL). (A-B) At lower initial EMD concentrations (33.2 and 66.4 µg/mL) the solutions at equilibrium appeared less concentrated as proteins bound to polystyrene. The equilibrium concentrations in solution were lower than at the initial concentrations of 99.6 and 166.0 µg/mL. (C-D) At the higher initial concentrations (99.6 and 166 µg/mL), the equilibrium concentration appeared to reach its maximum value, being identical in both cases (*highlighted by a dashed red frame*). The polystyrene surfaces were therefore assumed to be saturated, and within that range of starting concentration, the binding equilibrium reflected that of protein-protein binding, that is: between EMD proteins still free in solution and those covering the polystyrene.

As described in Figure 60 (p. 190), the initial hypothesis speculated that an initial binding phase, involving EMD-polystyrene interactions, would occur until the polystyrene microwell surfaces were saturated with protein and thereafter, a second binding phase comprising EMD-EMD interactions would become evident. However, there was no obvious inflection point identifying a time point when a distinct EMD-EMD binding phase became prominent over the initial EMD-polystyrene binding phase during which the microwell surfaces were brought to saturation. It may well be that these two phases may occur simultaneously, explaining why there was no obvious inflection points on the graphs (Figure 61A). It was concluded that it would not be possible to monitor protein loss from solution that was attributable to EMD-EMD interactions in isolation from the protein lost from solution due to EMD-polystyrene interactions occurring during saturation of the polystyrene microwell surfaces.

In order to overcome this problem it was decided to temporally separate the saturating of the surface of the microwells from the subsequent phase involving EMD interactions with immobilised EMD saturating the microwells. Quite simply, at the end of the incubation to saturate the microwell surfaces, the depleted protein solution was discarded and replaced with fresh EMD solution. The reduction in UV absorbance of this fresh solution over time would then more likely reflect the kinetics of EMD binding to the EMD already immobilised and saturating the polystyrene microwell surfaces.

3.2.2.2 Optimisation of the methodology to characterise the kinetics of protein-protein (EMD-EMD) interactions

As described above, it appeared to be impossible to track the initial phase of EMD immobilisation on to the polystyrene microwells separately from any subsequent EMD interactions between EMD still free in solution and the immobilised EMD. To monitor the EMD- immobilised EMD binding reaction in isolation, the assay was adapted by first saturating the surface of the microwells during an initial incubation with EMD solution and then once saturation equilibrium was achieved the solution was discarded and replaced with fresh EMD solution for a second phase of incubation during which protein-protein interactions could be tracked. As illustrated in Figure 63, showing an hypothetical result, any decrease of absorbance /protein concentration in solution during the second incubation, could be assumed to reflect the rate of protein-protein binding rather than the combination of [(EMD-polystyrene)+ (EMD-immobilised EMD)] binding[1].

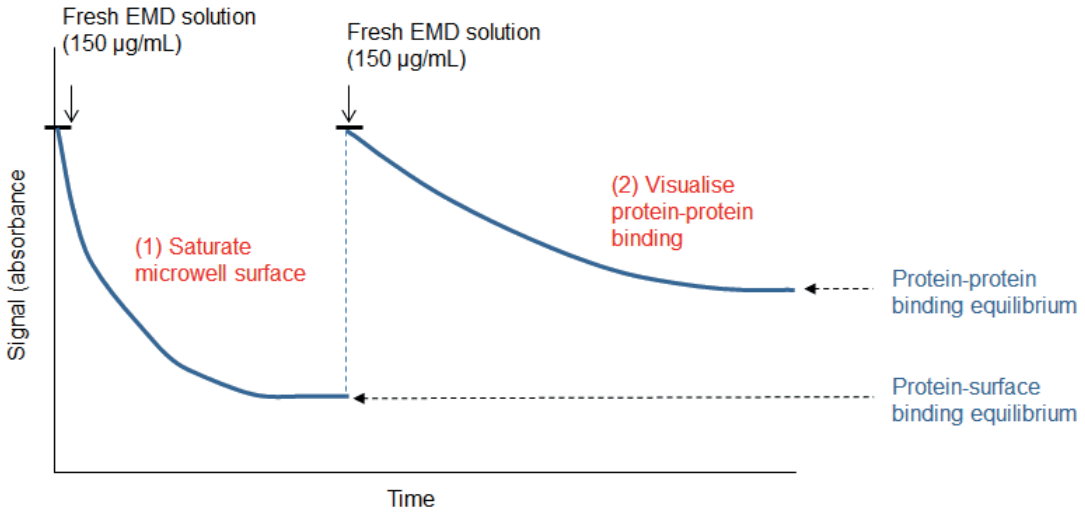


Figure 63 Hypothetical signal obtained while monitoring initial EMD-polystyrene interactions that saturate the microwell surfaces and subsequent EMD-EMD interactions occurring after the addition of fresh protein solution. The absorbance decrease in this second incubation period is assumed to be due to depletion of the solution due to the freshly added EMD interacting with the EMD already immobilised to the polystyrene surfaces during the initial saturation step.

[1] For clarification (and to relate back to the research question) the final aim of this thesis is not to compare protein-protein interactions with protein-polystyrene interactions but to compare WT r-amelogenin – WT r-amelogenin interactions with mutant p.Y64H r-amelogenin – mutant p.Y64H r-amelogenin interactions. Here, while the protein-polystyrene interactions were of no interest with regard to answering the research question; they are an incidental and necessary component of the method.

EMD proteins were dissolved at 150 µg/mL in PBS:1% acetic acid (284:1) (which is within the range known to saturate the microwell surfaces, Section 3.2.2.1 p. 192) was incubated at 37°C for 24 hours in the microwells (Phase 1) of polystyrene plates. At the end of the incubation period, aliquots were taken from the microwells for SDS PAGE analyses of the protein content of the solution, and the microwells were then emptied. Fresh solutions of EMD (at 150 µg/mL in PBS:1% acetic acid (284:1)) were added to the saturated microwells in order to track ensuing EMD-EMD binding (Condition "EMD-EMD-polystyrene"). At this point EMD was also added to fresh microwells as a control ("Ctr", or Condition "EMD-Polystyrene"). The decrease in UV absorbance was then monitored over the next 24 hours as protein-protein interactions depleted the free protein free in solution (Phase 2). At the end of this second incubation period, aliquots were taken for SDS PAGE analysis.

The decrease in absorbance recorded over the whole 48 hours (n=6), was corrected by blank subtraction and plotted as reported in Figure 64 along with the SDS PAGE analysis of protein remaining free in solution at the end of each of the two incubation periods.

Analytical SDS PAGE confirmed that proteins were depleted from solution during the first 24-hour incubation period during which the microwells became saturated. After 24 hours, this solution was removed and replaced with fresh EMD ready for the 2nd incubation period (24 – 48 hours) during which it was assumed that EMD-EMD interactions would occur that could be followed as a separate set of events from the initial protein-polystyrene interactions that occurred during the previous 24 hours. As would be expected, the SDS PAGE profile of the fresh EMD added after the initial 24 hours was identical to the profile of the EMD added at time 0[1]. However, after 48 hours the profile showed evidence of significant depletion that presumably occurred due to binding of free protein in solution to the previously saturated microwell surfaces. The protein profile associated with the control microwell, (comprising fresh EMD added to empty microwells at the 24 hour point) showed that the protein content was reduced at 48 hours (this is essentially a repeat of the initial incubation designed to saturate the microwells – it simply facilitates comparing the reduction in absorbance arising from protein–protein binding in the 24-48 hour period to the protein surface binding). Remarkably, SDS PAGE analysis suggests that not all proteins adsorbed to the

[1] Note: The absorbance measurements started 4.5 minutes (Phase 1) and 8 minutes (Phase 2) after the EMD solutions were prepared. As 4.5 and 8 minutes represent respectively 0.55 and 0.3% of the 24-hour incubation periods, the starting times of absorbance measurements were considered to be the starting times of experiment.

polystyrene surfaces, as low molecular weight (< 10 kDa) components remained at the same concentration after 24 hours of incubation on analytical SDS PAGE

It is noteworthy that the plateau region of the plot representing EMD-EMD binding in phase 2 was at a higher absorbance level than the plateau region obtained for the control sample and the plateau region associated with phase 1 during which the microwell surfaces were being saturated with EMD. This difference, marked ΔAbs on Figure 64, is thought to have been due to the absorbance of EMD on to the bottom of the microwell during phase 1 attenuating the UV light passing through fresh EMD solution added for phase 2. The cause of this difference will be reviewed again later, below.

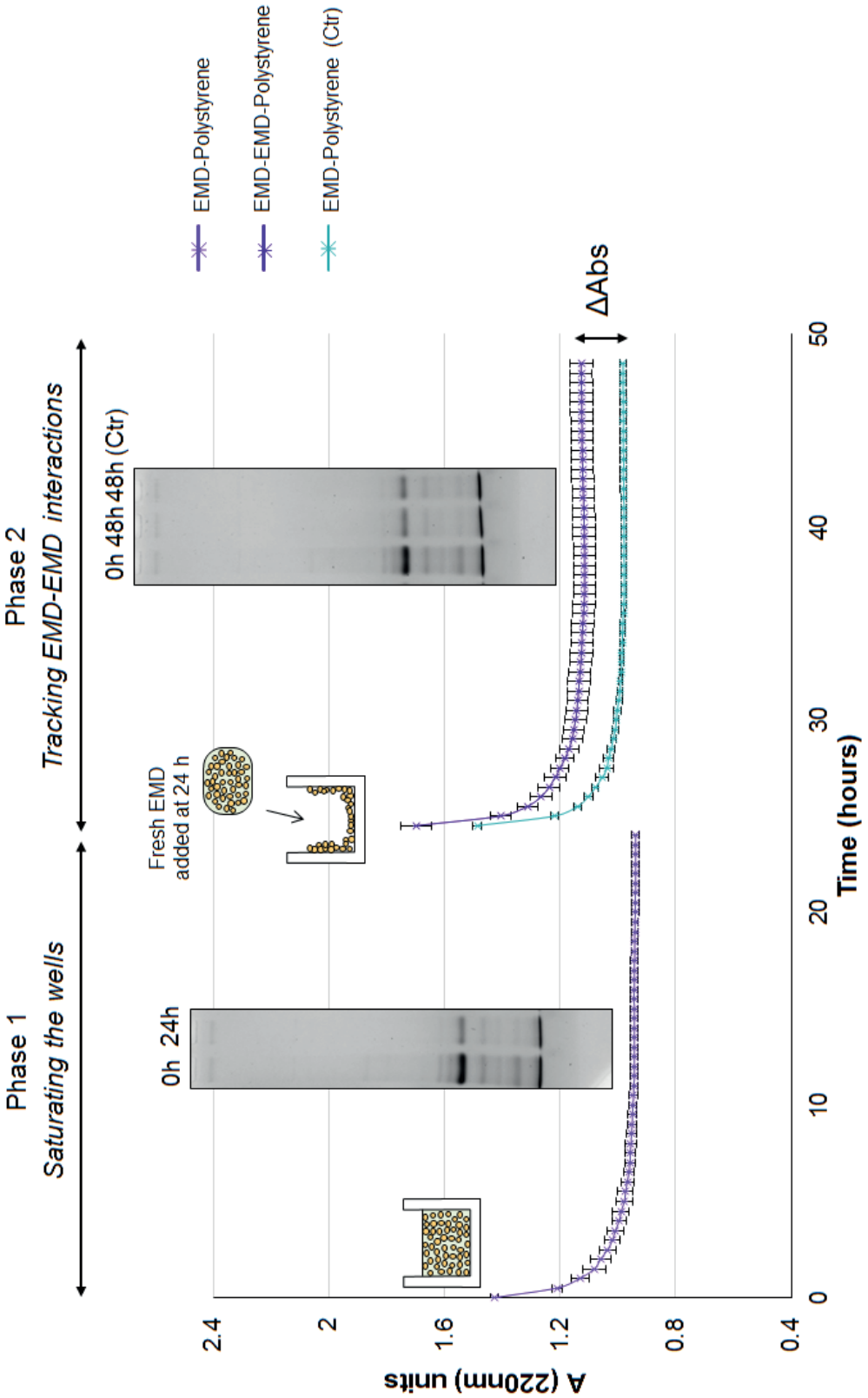

Phase 1
Saturating the wells
Phase 2
Tracking EMD-EMD interactions
0h 24h
Fresh EMD
added at 24 h
0h 48h 48h (Ctr)
EMD-Polystyrene
EMD-EMD-Polystyrene
EMD-Polystyrene (Ctr)
ΔAbs
A (220nm) units
Time (hours)
2.4
2
1.6
1.2
0.8
0.4
0
10
20
30
40
50

Figure 64 Analysis of EMD-EMD binding kinetics. In the first phase (0-24 hours) freshly added EMD proteins bind to the polystyrene surfaces of the microwells until they are saturated. After discarding the initial solution, fresh EMD solution is added to the microwell and left to incubate for a further 24 hours. The curves above show the decrease in UV absorbance as proteins are removed from solution as they bind to the microwell surfaces in phase 1 or, presumably, to proteins immobilised on the microwell surfaces in phase 2. The dotted line represents the control corresponding to the binding of EMD solution freshly added to a new microwell (essentially a repeat of phase 1 but run concomitantly with phase 2 for direct comparison). All the incubations were carried out at 37°C, in PBS:1% acetic acid (284:1). ΔAbs indicates the possible effect of EMD adsorbed to the bottom of the microwells on measurements recorded during phase 2. SDS PAGE analyses of the solutions added to microwells and the solutions removed from the microwells at the end of phase 1 and phase 2 are also shown. Absorbance data shows mean ± SD, n=6.

The data points shown in Figure 64 (p. 198), obtained during the second incubation period (and so describing EMD-immobilised EMD binding and EMD-polystyrene binding (control)), were replotted to represent the gain of protein by the polystyrene surface rather than the loss of protein from solution. The points were replotted as A_0-A_T (where A_0 = intial absorbance at the start of phase 2 and A_T = absorbance measured at time T between the start and the end of phase 2). The curves were then subjected to curve-fitting using OriginPro software and sigmoid Hill fit was identified as the best fitting model (Figure 65 below). The parameters extracted from this model are compared in Table 11. The Hill equation suggested that the times required to reach half equilibrium ($T_{1/2}$, identified in Hill equation as 'K') were similar both when EMD was interacting with immobilised EMD (EMD-EMD-Polystyrene) or when EMD was interacting directly with polystyrene (EMD-EMD-Polystyrene). This similarity could be because there is still a possibility that "Phase 1" might be a combination of EMD-polystyrene and EMD-EMD binding. However, the depletion of free EMD from solution was greater when it was interacting with immobilised EMD compared to when it was interacting with polystyrene (with a 'Maximum gain' value predicted 14% higher).

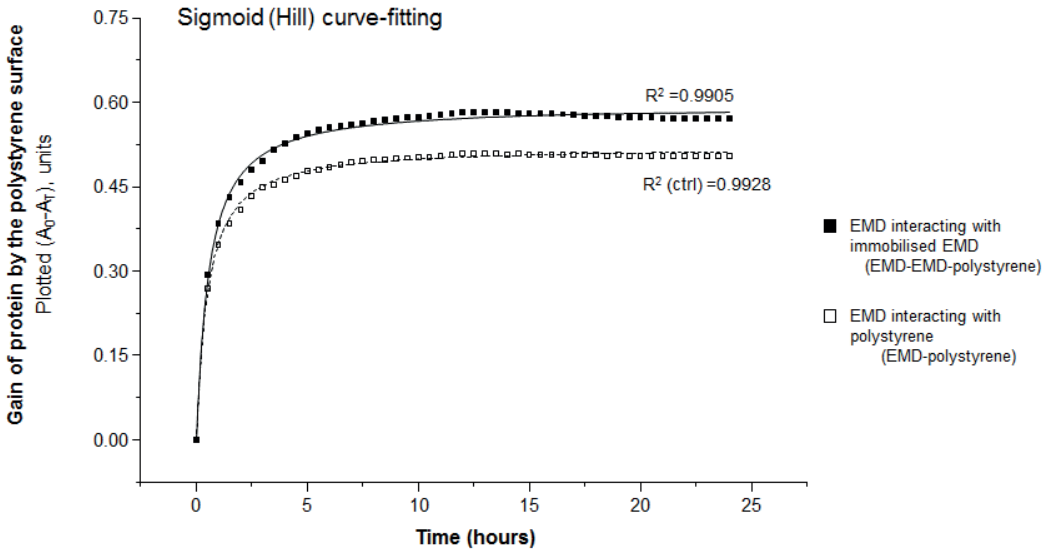


Figure 65 Depletion of EMD from solution (calculated as A_T subtracted from A_0 and plotted against time) using Hill curve-fitting. The conditions compared show measurements of absorbance depletion after adding 150 µg/mL EMD solution to empty microwells (control - white marker), or to microwells previously saturated with EMD (solid black markers). The Hill model shows that both conditions had similar half-time rates, but apparent maximum gain was greater (14% higher) in the case of microwells previously saturated with EMD during phase 1. This difference could be due to adsorption by EMD already immobilised to the bottom of the microwell elevating the absorbance readings obtained during phase 2 when freshly added EMD was interacting with the immobilised EMD.

Table 11 Kinetics parameters predicted by OriginPro curve-fitting with Hill Sigmoid equation (displayed above in Figure 65), describing the gain of EMD proteins by polystyrene for EMD-EMD-Polystyrene and EMD-Polystyrene (ctrl).

Hill sigmoid Parameters	Maximum gain	$T_{1/2}$
EMD-Polystyrene (Ctr)	0.5199	0.5148
EMD-EMD-Polystyrene	0.5939	0.5583

The Hill equation describes the effect of ligand concentration on the proportion of macromolecules saturated by that ligand and was originally derived to describe the sigmoidal curve associated with haemoglobin binding (Hill, 1910). The Hill equation best fitted the plots and can be used to predict $T_{1/2}$ and the maximum ordinate value reached (maximum gain). Any conclusions drawn need to be considered carefully as the Hill equation usually relevant to describe other types of reaction than those reported throughout the Section 3.2.2.

To summarise, the data suggest that EMD-EMD interactions can be monitored quantitatively by spectrophotometric measurement of the depletion of free EMD remaining in solution. Obviously, the ultimate aim would be to compare WT and mutant p.Y64H r-amelogenins in the system by tracking the depletion of solubilised WT r-amelogenin as it interacts with immobilised WT r-amelogenin saturating the microwell surfaces. The kinetics of this process would then be compared to the kinetics associated with solubilised mutant p.Y64H r-amelogenin as it interacts with immobilised mutant p.Y64H r-amelogenin saturating the microwell surfaces. However, as described above, these experiments also raised the issue that EMD adsorbed to the bottom of the microwell during the saturation phase attenuates the UV light passing through the EMD solution placed in the microwell during phase 2 when the interaction of freshly added EMD with the immobilised EMD was being monitored (manifesting as ΔAbs shown in Figure 64, p. 198). Attempts to eliminate the effect of this confounding factor are the subject of the next Section.

3.2.2.3 Effect of proteins adsorbed on the bottom of the microwells on the binding assay performance

As described in the previous Section, EMD immobilised to the bottom of the microwells of the polystyrene plates may potentially confound absorbance readings taken during phase 2 of the binding assay, when freshly added EMD was added to the microwells already saturated with immobilised EMD. This was investigated by measuring how the absorbance was affected by EMD binding to the bottom of the microwells. A solution of 150 µg/mL EMD proteins in PBS:1% acetic acid (284:1) was left to incubate in the microwells at 37°C for 24 hours[1]. At the end of incubation, the solution was discarded from the microwells, immediately after which the absorbance was read again. The absorbance measurements (n=6 corrected by blank subtraction) are plotted in Figure 66.

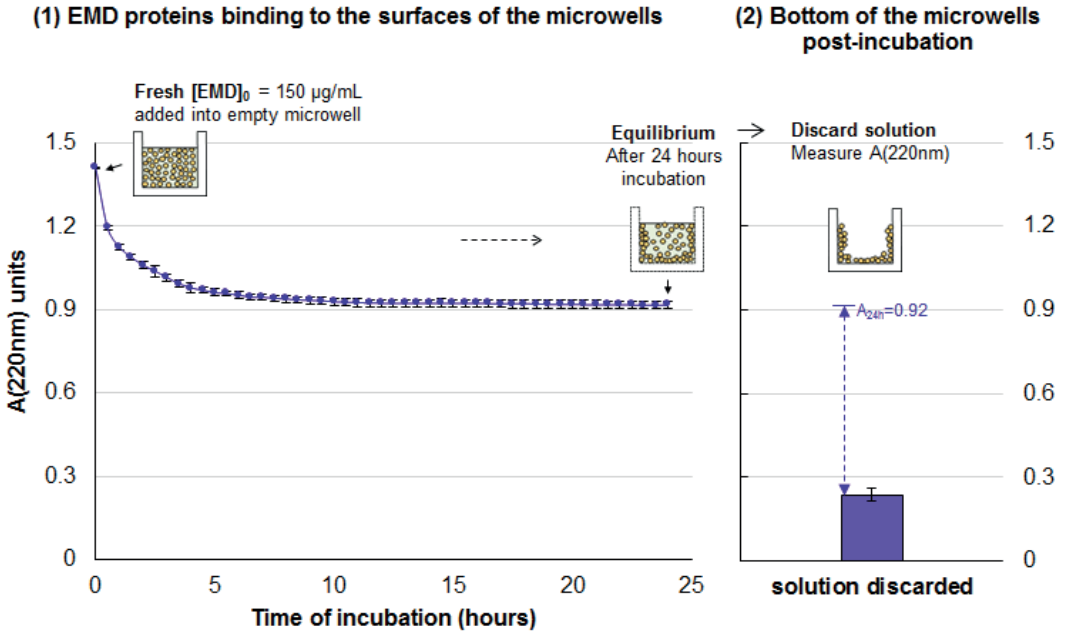


Figure 66 The effect of EMD proteins bound to the bottom of the microwells on absorbance values. (1) Over an incubation period of 24 hours, the absorbance decreased by 0.49 units after addition of fresh EMD solution (150 µg/mL), presumably due to EMD proteins binding to the surfaces of the microwells. The incubation was carried out at 37°C, in PBS:1% acetic acid (284:1). (2) The absorbance read after discarding the solution (after incubation) was 0.24 units, which is due to the layers of EMD that adsorbed on the bottom of the microwells. Data shows mean ± SD, n=6.

[1] Note: The absorbance measurements started 4.5 minutes after the EMD solution was prepared. As 4.5 minutes represent only 0.3% of the 24-hour incubation period, the starting time of absorbance measurement was considered to be the starting time of experiment.

In the first 24 hours of incubation the absorbance dropped in a similar fashion to that observed previously (Figure 64), stabilising at a value of around 0.92. After discarding the EMD solution, the absorbance dropped to 0.24. This residual absorbance likely reflects EMD proteins that bound and accumulated onto the bottom surface of the microwells during the first 24 hours of incubation. This experiment confirms therefore that proteins bound to the bottom surface of the microwells had a significant effect on absorbance reading during the binding assays reported previously.

An absorbance value of 0.24 is not negligible and in fact is close to the value 0.3, which is considered the ideal value to use for spectrophotometric measurements, as it is the point where 50% of the photons emitted by the light source are absorbed, which allows for readings of the highest precision. This actually opens up the possibility to compare the extent of binding or aggregation of WT and p.Y64H r-amelogenins by measuring the direct accumulation of the proteins on the bottom of the microwells rather than by measuring their depletion from solution. However, these would be end-point measurements.

For continuous kinetics analyses based on depletion of protein from solution, it would be preferable to eliminate the signal bias caused by EMD proteins bound to the bottom surfaces of the microwells. One method to achieve this would be to block the bottom surface of the microwells with a blocking protein; in a similar way to that employed when using microwell plates in classic ELISA where a blocking protein is applied after the target antigen has adsorbed to the microwells to prevent non-specific adsorption of detection antibodies to exposed microwell surfaces. As detailed in the next Section (3.2.2.4), a preliminary assay was therefore designed to test the efficacy of gold standard blocking proteins used in ELISAs for this purpose.

3.2.2.4 Attempts to block the microwells

To prevent any EMD proteins from binding to the bottom of the microwells and affecting the absorbance values, the use of a blocker (covering the bottom surface of the microwells only) was trialled. This required the availability of an effective blocker that would prevent EMD proteins from binding to the surface of the microwells. Gold standard blockers BSA and NFDM were tested for this purpose.

Blocking was carried out by adding 200 µL of 1% BSA or NFDM solutions into the microwells, incubating and washing before the addition of 200 µL EMD solution (made of EMD dissolved in PBS:1% acetic acid (290:1)). The volume of the blocking solution was the same as that of EMD solution in this preliminary test (to ensure that the surface covered by EMD solution could be entirely blocked, which would show whether blocking was effective). If that was indeed the case, then one should expect that EMD would not bind to the blocked microwell surfaces and instead remain in solution, providing a constant absorbance signal overtime (unlike the control with no blocker where absorbance would steadily decrease as the EMD adsorbed to the microwell surfaces). The theory for this is illustrated in Figure 67.

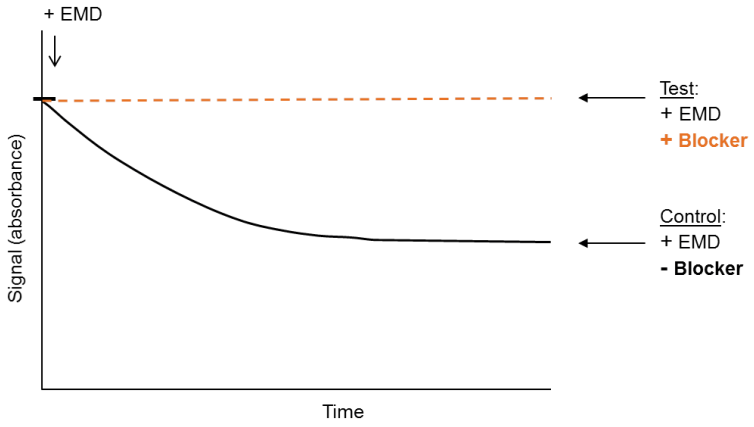


Figure 67 Hypothetical signal expected over incubation time (orange dashed line) in microwells effectively blocked compared to control with no blocking.

The absorbance measurements (averages, n=6, corrected by blank subtraction) over time[1] are plotted in Figure 68A. Analytical SDS PAGE (Figure 68B) shows the content of the aliquots taken before incubation and at the end of incubation.

[1] Note: The absorbance measurements started 10 minutes after the EMD solution was prepared. As 10 minutes represent only 0.7% of the 24-hour incubation period, the starting time of absorbance measurement was considered to be the starting time of experiment.

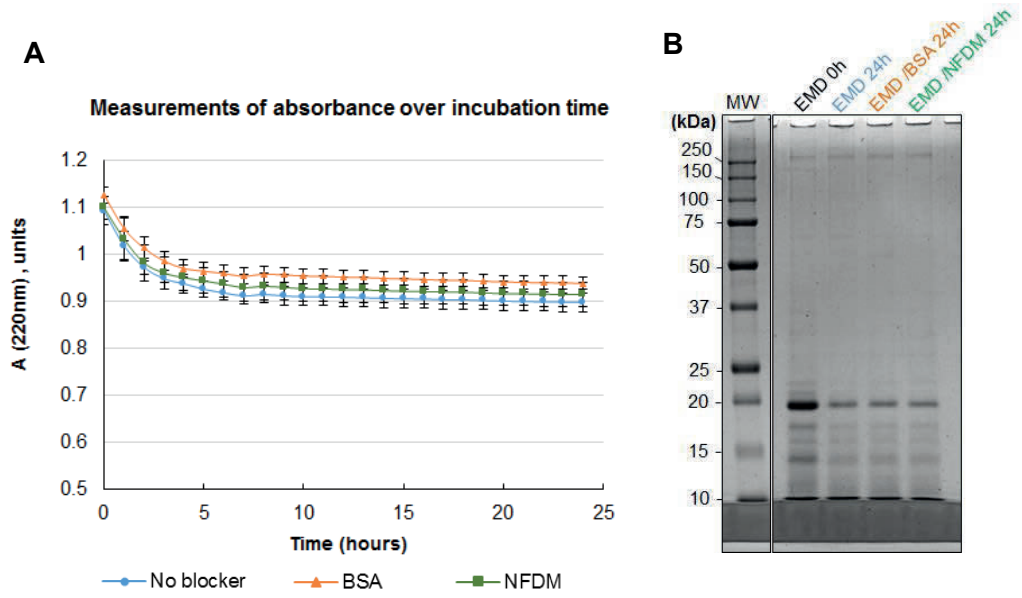


Figure 68 Evaluation of the ability of the gold standard blocking proteins bovine serum albumin (BSA) and non-fat dry milk proteins (NFDM) to block EMD binding to microwell surfaces. (A) After adding fresh EMD solution, absorbance measurements in empty microwells (control) or in microwells previously blocked followed the same pattern, indicating that the blocking was not effective. (B) The solutions in the microwells at the end of the incubation were analysed using analytical SDS PAGE which confirmed that the amount of EMD remaining in solution after 24h incubation, and showed that blocking the microwells with either blocking protein had little effect on the depletion of EMD from solution (i.e. did not block EMD adsorbing to the microwell surfaces). All the incubations were carried out at room temperature, in PBS:1% acetic acid (290:1). Absorbance data shows mean ± SD, n=6.

The absorbance measurements (Figure 68A) show that blocking with neither BSA nor NFDM significantly altered depletion kinetics. Consistently, the results obtained on SDS PAGE analysis (Figure 68B) showed that under all conditions tested, the concentration of EMD decreased by similar levels leaving equal amounts of EMD left in solution at the end of incubation. These data together indicated that blocking the microwells was not successful. The possible causes for this are hypothesised and discussed in Section 4.2.2.3. (pp. 247 – 252).

As described above, previous experiments (Section 3.2.2.3) showed that it was feasible to spectrophotometrically determine the amount of EMD adsorbed on to the bottom of the microwell after incubation with end point absorbance values (reflecting the amount of proteins that bound to the bottom of the microwell after 24-hour incubation of EMD solution), of 0.24. This value is close to the value 0.3, which as described above is the most accurate absorbance value for quantitative spectrophotometric measurements. At the beginning of the project, it was not

envisaged that enough protein could be adsorbed onto the microwell surface to be directly measured spectrophotometrically in this way. However, the fact that the accumulation of EMD on the microwell bottom could be measured directly, raised the possibility of comparing interactions between WT and mutant p.Y64H r-amelogenins simply by direct spectrophotometric measurement of the proteins accumulating on the microwell bottoms at the end of the incubation period.

3.2.2.5 Characterisation of the behaviour of WT and mutant p.Y64H r-amelogenins in the microplate-based assay

Throughout this thesis, the focus has been on the production, purification and use of His-tagged r-amelogenins (as the initial strategy was to use nickel affinity chromatography, the standard purification method for recombinant proteins, to obtain purified WT^{+His} and Mut^{+His} r-amelogenins) in protein-protein binding studies. However, during the course of the work it became clear that nickel chromatography was not completely suitable for use with these proteins and preparative SDS PAGE was able to produce purer fractions. This led to the realisation that nickel chromatography and the associated His-tag were not actually required as the initial acetic acid extraction and subsequent preparative SDS PAGE were sufficient to achieve a high degree of purification. A further advantage of using His-tag-free r-amelogenins is that the time-consuming and costly His-tag cleavage step can be eliminated. To this end, a commercial company (Novoprotein Scientific, NJ, USA) was tasked with generating plasmids encoding WT and mutant p.Y64H r-amelogenins without His-tag, respectively "WT^{-His}" and "Mut^{-His}" r-amelogenins. However, this proved problematic and two separate batches of plasmids delivered during the final phase of the experimental work failed to result in the expression of protein for reasons unknown. A third batch of plasmids delivered during the write up phase of the thesis did express both WT^{-His} and Mut^{-His} r-amelogenins. Details regarding their sequences and the plasmid used are provided in Appendix A. Their identities were confirmed by mass spectrometry (see Appendix B) the spectra obtained showed high degree of purification with a clear peaks corresponding to r-amelogenins at 20160.10 Da for WT^{-His} and 20134.2 to Mut^{-His} after acetic acid extraction (kindly performed by Dr Sarah Myers and Mr. Matthew Percival). WT^{-His} and Mut^{-His} r-amelogenins were tested along with WT^{+His} and Mut^{+His} r-amelogenins in the microplate assays based on simply measuring the absorbance of protein accumulating on the bottom of the microwells directly (see Section 3.2.2.3 pp. 202 - 203).

WT^{+His} and Mut^{+His} r-amelogenins were used in the uncleaved state due to time constraints at this late stage in the research. The use of His-tagged proteins in functional studies is widely practiced with the assumption that the tag has no effect on protein conformation or function (Chant et al., 2005) and there are several examples where His-tagged enamel proteins have been used in functionality studies (e.g. His-tagged amelogenin, ameloblastin and enamelin were used to investigate interaction of these enamel proteins with fibronectin (Beyeler et al., 2010) and His-tagged ameloblastin was used to investigate growth factor-like activity of ameloblastin and its

effect on the effect on cell attachment and proliferation of periodontal ligament cells (Zeichner ‐ David et al., 2006).

Freshly prepared r-amelogenin solutions (comprising 7.45 μM solutions of WT^{-His}, Mut^{-His}, WT^{+His}, Mut^{+His} r-amelogenins in PBS: 1% acetic acid at 216.4:1 ratio) were added to empty polystyrene microwells and left to incubate for 24 hours at 37°C. The absorbance of the microwells (n=6 corrected by blank subtraction) was tracked over 24 hours and the absorbance measurements are plotted in Figure 69 (p. 209). The microwell contents were then discarded and the absorbance attributable to protein bound to the bottom surface of the microwells (see Section 3.2.2.3 p. 202 for further explanation) was measured as reported in Figure 70 (p. 210); The contents of the solutions before (“0h”) and after (“24h”) incubation were analysed using Analytical SDS PAGE with Coomassie Blue staining (Figure 69B) with the staining intensity of the protein bands compared by gel densitometry.

For WT^{-His}, both absorbance measurements and analytical SDS PAGE showed a minor decrease (~8-10%) of r-amelogenin concentration in solution after 24 hours incubation and in keeping with this, the absorbance due to bound protein on the bottom of the microwells was close to zero (Figure 70). This data indicated that WT^{-His} r-amelogenin did not associate with the polystyrene surface of the microwells but instead remained in solution. In contrast, Mut^{-His}, WT^{+His} and Mut^{+His} r-amelogenin all disappeared from solution over 24 hours incubation, presumably as they associated with the polystyrene microwell surfaces. For Mut^{-His}, the absorbance appeared to decrease by ~37% over the 24 hours incubation period and SDS PAGE showed that the staining intensity of the 24 kDa band decreased by ~45% over that time (Figure 69). As for WT^{+His} and Mut^{+His} r-amelogenins, the absorbance decrease over 24 hours was even greater, in that the absorbance appeared to decrease by ~62% and the staining intensity of the 27 kDa (His-tagged) bands was reduced by ~80% after 24h incubation on analytical SDS PAGE (Figure 69). The absorbance measurements of what was left on the bottom of the microwells demonstrated average readings ranging between 0.16 and 0.29 units (Figure 70), which confirmed that the Mut^{-His}, WT^{+His} and Mut^{+His} r-amelogenins associated with the polystyrene surface of the microwells in contrast to WT^{-His}, which did not. The result is summarised in Table 12.

Table 12 Comparison of binding behaviours of WT^{-His}, Mut^{-His}, WT^{+His}, Mut^{+His} r-amelogenins in the microplate based assay. Summary of the data obtained in the Section 3.2.2.5.

r-Amelogenin	Associates with microwell surface?
WT^{-His}	No
Mut^{-His}	Yes
WT^{+His}	Yes
Mut^{+His}	Yes

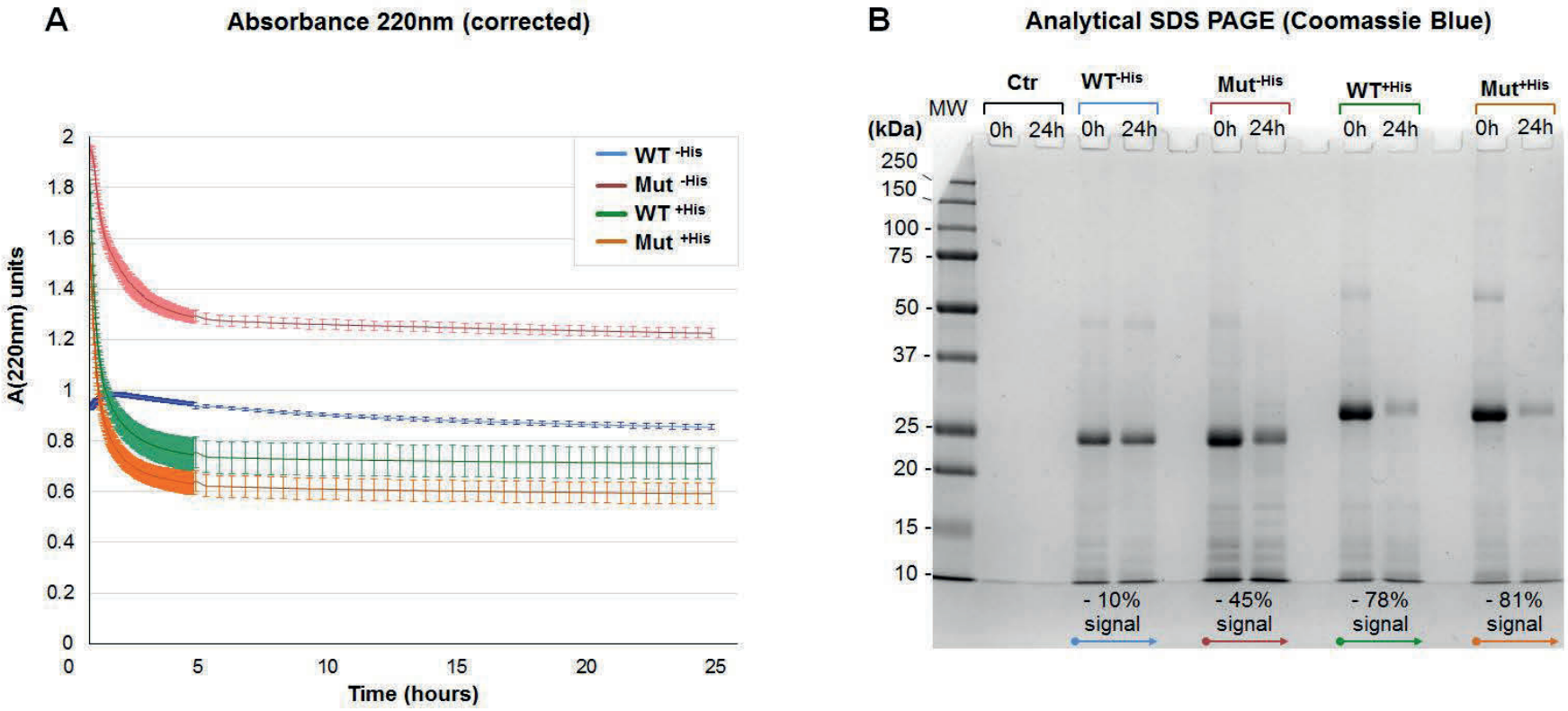


Figure 69 Comparison of r-amelogenin protein depletion from solution after 24 h incubations in micro-wells for WT^{-His}, Mut^{-His} (His-tag-free), WT^{+His}, Mut^{+His} (His-tagged) r-amelogenins. (A) Absorbance measurements over 24 hours incubation showed a minor decrease of signal for WT^{-His} r-amelogenin but a major decrease of signals for Mut^{-His}, Mut^{+His} and WT^{+His} r-amelogenins. (B) Analytical SDS PAGE with Coomassie Blue staining showing the solution content before incubation ("0h") and at the end of incubation ("24h"). All the incubations were carried out at 37°C, in PBS:1% acetic acid (216.4:1). Absorbance data shows mean ± SD, n=6.

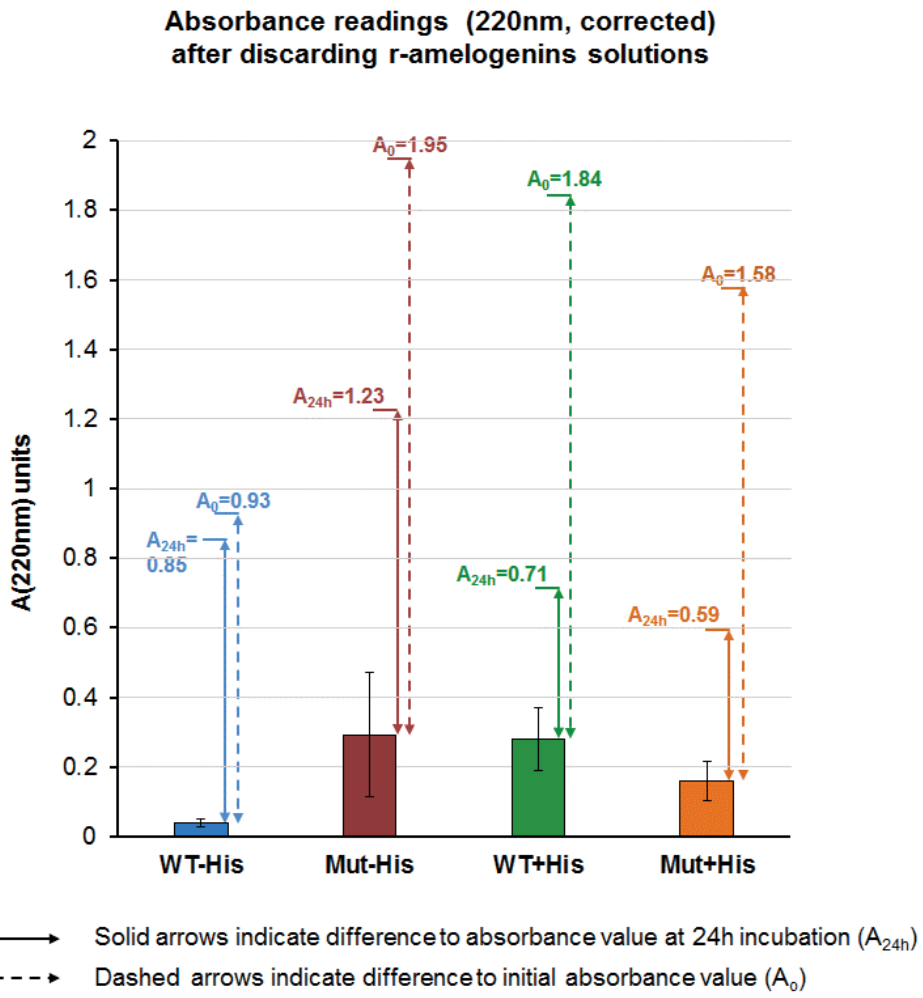


Figure 70 Comparison of absorbance readings in the microwells after discarding the r-amelogenin solution (after 24-hour incubation, see Figure 69A) for WT and mutant p.Y64H r-amelogenins ± His-tag. Solid-line arrows reflect the difference of absorbance with the absorbance at the end (24 hours) of incubation (Figure 69A). Dashed-line arrows reflect the difference of absorbance with the absorbance at the beginning of incubation (Figure 69A). The initial and final absorbance values of incubation (Figure 69A) are indicated, as A_0 and A_{24h}. Absorbance data shows mean ± SD, n=6.

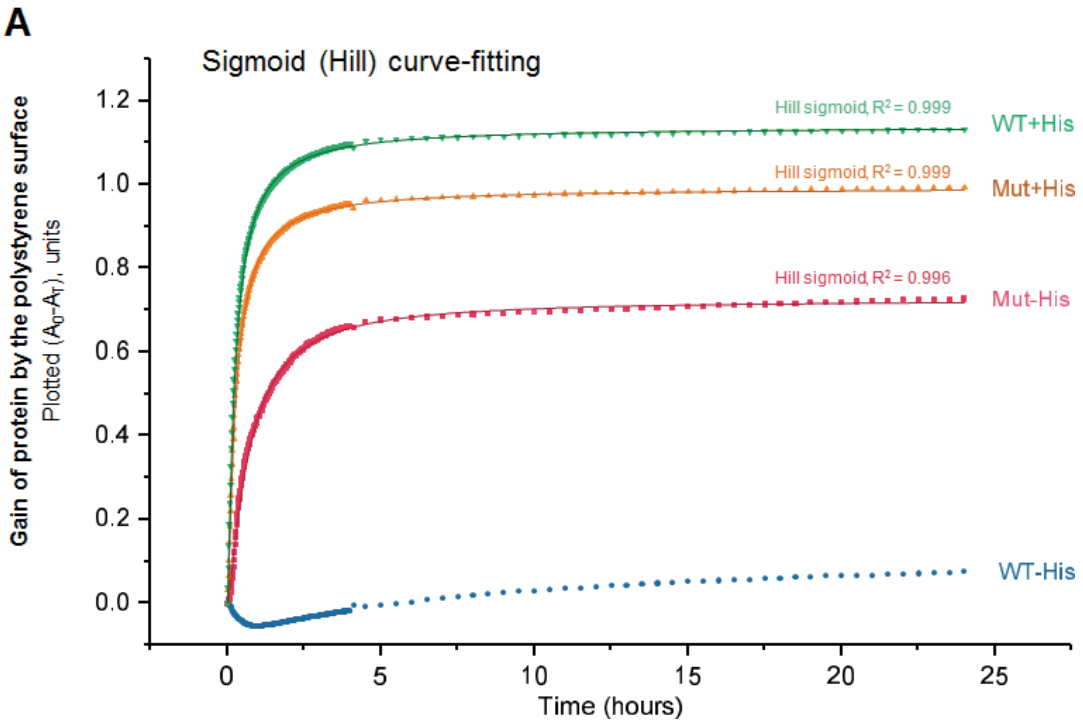
Figure 69A suggested that both WT^{+His} and Mut^{+His} disappeared from the solution more rapidly than WT^{-His} and Mut^{-His} r-amelogenins. The data in Section 3.2.2.3 (p. 202) previously opened up the possibility of comparing the extent of binding or aggregation of WT and mutant p.Y64H r-amelogenins by measuring the direct accumulation of the proteins on the bottom of the microwells rather than by measuring their depletion from solution. It is important to bear in mind that this does not account for an exact comparison of protein binding kinetics but rather a preliminary descriptive estimation as the proteins binding to the bottom of the microwells mask the actual kinetics of protein depletion from solution.

As it was done previously in Section 3.2.2.2 (Figure 65 p. 200) the data obtained in this Section (Figure 69A) were replotted[1] to represent the gain of protein by the polystyrene surface (A_0-A_T (where A_0 = initial absorbance, time 0 and A_T = absorbance measured at any time T between the start and the end of incubation), which allowed all curves to start from the same point, at abscissa and ordinate = 0. The graphs obtained are shown in Figure 71 below. The curves representing Mut^{-His}, WT^{+His} and Mut^{+His}, subjected to curve-fitting using OriginPro software, successfully fitted the Hill sigmoid model (previously used in Figure 65 to analyse EMD depletion from solution). The data points representing WT^{-His} were not subjected to curve-fitting because, as explained previously, WT^{-His} did not significantly disappeared from solution.

The models generated for Mut^{-His}, Mut^{+His} and WT^{+His} recombinants (Figure 71) predicted that the maximum gain of r-amelogenin to polystyrene surface was ~40-60% higher with the His-tagged r-amelogenins Mut^{+His}, WT^{+His}, and also, the depletion pace was faster, with the time taken to reach half saturation predicted to be 15 minutes with the His-tagged r-compared with 41 minutes for the non-tagged Mut^{-His} recombinant. This model highlights, despite the 'masking' effect of proteins binding to the bottom of the microwells, that:

- In His-tag-free r-amelogenins, the single amino acid change of a tyrosine into a Histidine caused a significant gain of protein by the polystyrene surface, which can fit to a Hill sigmoid curve with R^2=0.996.
- However, the addition of a His-tag in the primary sequence in both WT and mutant r-amelogenins increased the amplitude and pace of gain of proteins on polystyrene surfaces to such an extent that both curves representing WT^{+His} and Mut^{+His} had similar shapes. Moreover, WT^{+His} r-amelogenin appeared to adhere to the polystyrene surface more than Mut^{+His} r-amelogenin.

[1] Note: The absorbance measurements started 7 minutes after the EMD solution was prepared. As 7 minutes represent less than 0.5% of the 24-hour incubation period, the starting time of absorbance measurement was considered to be the starting time of experiment.



B

Hill sigmoid parameters		
Recombinant	Maximum gain	Half-equilibrium time
WT +His	1.14 units	15 minutes
Mut +His	0.99 units	16 minutes
Mut –His	0.73 units	41 minutes

Figure 71 Comparison of gains of protein by the polystyrene microwell surfaces for r- amelogenins: Mut^{-His} (His-tag-free), WT^{+His}, Mut^{+His} (His-tagged). (A) Absorbance values obtained previously in Figure 69A were replotted as A_T subtracted from A_0 , against time. The curves were subjected to curve-fitting using the Hill sigmoid model. (B) The Hill equation parameters extracted in panel B suggest that His-tagged recombinants adhered to polystyrene surfaces at least twice faster, and to a higher extent than, the His-tag-free Mut^{-His} r-amelogenin. WT^{+His} and Mut^{+His} 'Gain' curves reach half their maximum value (predicted) at 15-16 minutes of incubation. The adhesion ('gain by the polystyrene surface') is higher in WT^{+His} than Mut^{-His}, indicating that the presence of a His-tag altered the binding behaviour of r-amelogenins.

Chapter 4 Discussion

4.1 Purification strategy used to obtain purified r-amelogenin: preparative SDS PAGE identified as the most effective technique

4.1.1 Purification of r-amelogenin

As described in the literature review (Section 1.3.5 p. 84), efforts have been made previously to purify r-amelogenin. One of the most frequently published purification methods for producing recombinant mouse M180 amelogenin expressed by *E. coli* starts with cell lysis under denaturing conditions (6 M guanidine hydrochloride) followed by centrifugation, precipitation with ammonium sulphate and preparative C4 reverse phase column chromatography in 0.1% trifluoric acid/acetonitrile (Simmer et al., 1994). This method has been widely used to provide r-amelogenin for use in *in vitro* studies e.g. characterising amelogenin assembly into nanospheres (Moradian-Oldak et al., 1994). The method had the advantage of not being dependent on a His-tag but did not provide r-amelogenin at single band purity on SDS PAGE. Moradian-Oldak et al (1994) noted that the methodology generated mouse recombinant amelogenins that co-eluted with contaminating *E. coli* proteins. This purification method originally yielded 4–11 mg purified proteins (in acetonitrile) per litre cell culture (Simmer et al., 1994). The method was subsequently modified by adding a His-tag (MRGSHHHHHHGS) to the N-terminal with the introduction of a nickel column affinity chromatography step. With this method the yield was increased to 70 mg/L cell culture (Buchko et al., 2013).

Nickel column affinity chromatography was used by others to purify r-amelogenins for use in functional or structural studies (Moradian-Oldak et al., 2000, Taylor et al., 2006). The purity of the r-amelogenin obtained directly by nickel column chromatography was not commented on by Moradian-Oldak et al (2000) while the fractions obtained by Buchko et al. (2013), comprising His-tagged r-amelogenin were subjected to further purification using reverse phase C18 columns to increase purity (Buchko et al., 2013). Later reports provided more detailed analyses and assessments of the ability of nickel column affinity chromatography to purify r-amelogenin: these include a publication (Taylor et al., 2006) and this thesis. The relevance of nickel column to purify r-amelogenin is discussed in greater detail in Section 4.1.2.2 (pp. 218 – 224).

Other purification methods include salting out (e.g. using ammonium sulphate) followed by cation exchange chromatography and direct use of reverse phase chromatography. These were used to obtain rP172 and rM179 to characterise their *in vitro* cleavage by MMP-20 (Ryu et al., 1999).

A recent method based on the preferential solubility of amelogenin in acetic acid allowed the extraction and significant enrichment of r-amelogenins from bacterial pellets (Svensson Bonde and Bulow, 2012). This method is discussed in greater detail in Section 4.1.2.1 (pp. 216 - 217). The solubility of amelogenin in acetic acid was also used to extract native amelogenins from porcine teeth (Wang et al., 2018) but the method would also have extracted the other enamel matrix proteins such as ameloblastin and enamelin which are equally soluble in acetic acid. The use of acetic acid as a selective solubilising agent for r-amelogenin, published by Svensson Bonde and Bulow (2012) negated the need for His-tag based purification and provided a simple route to His-tag-free r-amelogenin preparations for use in functional studies and the methodology has since been used by others. For example, Buchko and Shaw, (2015) used it to prepare isotopically-labelled amelogenin for NMR analyses (Buchko and Shaw, 2015). They extracted r-amelogenin from *E. coli* in acetic acid and dialysed the extracts against acetic acid to remove salts and reduce the ionic strength to limit amelogenin aggregation. The r-amelogenin was then further purified using reverse-phase HPLC. Dialysis and HPLC improved the purity of the fractions but contaminants were still visible on analytical SDS PAGE after Coomassie Blue staining. Buchko et al. (2015) adjusted their protocol by using 6M guanidine hydrochloride for *E. coli* cell lysis, instead of 2% acetic acid and reported that the amelogenin yield was improved but more contaminants were co-extracted.

To date then, the extraction and purification of r-amelogenin is complex with most reported methodologies using denaturants such as guanidine hydrochloride (Simmer et al., 1994, Taylor et al., 2006, Buchko and Shaw, 2015) or urea (Taylor et al., 2006) to achieve an initial extraction of r-amelogenin followed by further downstream purification steps.

This thesis tested a panel of purification methods to extract and purify r-amelogenin from *E. coli.* The first priority was to define a method that could provide r-amelogenin at single-band purity on analytical SDS PAGE after silver staining in sufficient amounts for use in downstream functionality studies**.** Single-band purity on analytical SDS PAGE with silver staining generally reflects a satisfactory purity due to the high sensitivity of silver staining: it has a detection limit “100 times lower than that of Coomassie Blue staining” as reviewed by Shevchenko et al (1996) (Shevchenko et al., 1996) and currently the silver staining kit (cat no. 24612, ThermoFisher Scientific,

Leicestershire, UK) employed in this thesis allows for detection down to 0.25 ng proteins per band.

4.1.2 Purification methods tested to determine the optimum purification of r-amelogenin

This section critically discusses: (i) the optimisation of acetic acid extraction of His-tagged mouse r-amelogenin (Section 4.1.2.1), (ii) His-tag-affinity nickel column purification of His-tagged mouse r-amelogenin (Section 4.1.2.2), (iii) SEC purification of His-tagged mouse r-amelogenin (Section 4.1.2.3) and (iv) the use of preparative SDS PAGE to purify His-tagged mouse r-amelogenin (Section 4.1.2.4). The purification strategies tested are summarised below in Figure 72.

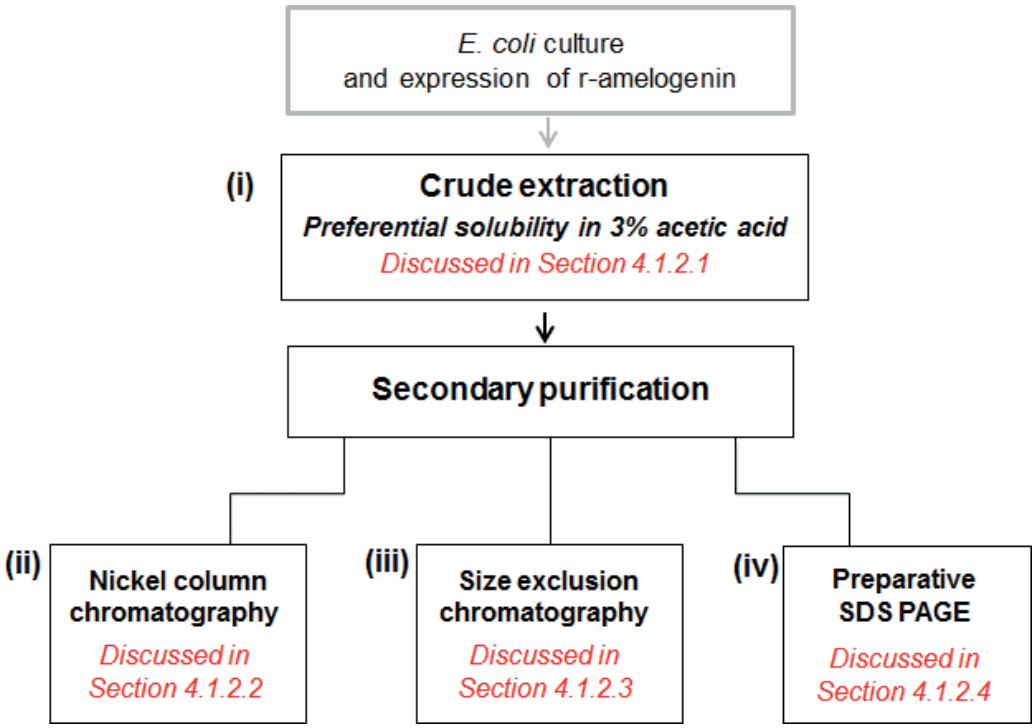


Figure 72 Strategies tested to extract and purify mouse r-amelogenin from *E. coli*. (i) The extraction of r-amelogenin in acetic acid, based on its preferential solubility in acidic, was optimised. Strategies for optimisation are discussed in Section 4.1.2.1. After optimisation of step (i) was achieved, the acetic acid extract (lyophilised, desalted and lyophilised) was subjected to secondary purification, to isolate r-amelogenin. The purification methods tested were (ii) the gold-standard nickel column chromatography method, discussed in Section 4.1.2.2, (iii) SEC, discussed in Section 4.1.2.3 and (iv) preparative SDS PAGE, discussed in Section 4.1.2.4.

4.1.2.1 Acetic acid extraction effectively reduced bacterial contamination

The acetic acid extraction technique reportedly producing r-amelogenin from *E. coli* at > 95% purity in a single purification step was previously published (Svensson Bonde and Bulow, 2012). This technique is based on the simple premise that amelogenin is soluble in 3% acetic acid at 80°C whereas *E. coli* proteins are insoluble under these conditions. First, the volume of acetic acid for extraction was optimised (Section 3.1.1.2.1, p. 138). This optimisation is also referred to as "optimisation of 'weight of E. coli to volume of acetic acid") as Svensson Bonde and Bulow (2012) reported that the total yield increased with increasing volumes of acetic acid (using 0.035 - 0.35 g/mL acetic acid) without, however, providing any specific details. The extractions using 0.33 and 0.17 g/mL acetic acid provided a good yield (Section 3.1.1.2.1, pp. 138 - 139). and conveniently small working volumes. Increasing the ratio of solvent to solute allows more solute to enter the solution until equilibrium is re-established. Therefore extractions using 0.033 and 0.0033 g/mL acetic acid were predicted to yield more dilute extracts of larger volume that contained a greater absolute amount of amelogenin. Indeed, taking the larger dilution factor into account (Figure 36B p. 139), the highest yield of amelogenin was obtained with 0.0033 g/mL acetic acid. This is in agreement with the observation by Svensson Bonde and Bulow, that the total extraction yield increased with increasing volume of acetic acid (Svensson Bonde and Bulow, 2012). However, although extraction at 0.0033 g/mL provided an optimum yield, the excessive dilution led to handling problems and made downstream tasks such as lyophilisation more difficult. Therefore, extraction at 0.033 g/mL acetic acid was chosen as the preferred option since the yield was acceptable and extraction volume was manageable.

While Svensson Bonde and Bulow (2012) reported that the final fraction purity was higher when only heat treatment was used for cell lysis/purification, compared to when sonication was used, the results obtained (Section 3.1.1.2.2, pp. 140 - 141) showed that ultrasonic mixing did not make any significant difference to sample purity. In contrast to Svensson, Bonde and Bulow (2012)'s statement that "sonication is not very feasible in large scale", it was found that ultrasonication facilitated mixing and extraction of r-amelogenin Mut^{+His} r-amelogenin in particular (Section 3.1.1.2.2, p. 140). However, Svensson Bonde and Bulow (2012) did not specify how sonication was carried out and the equipment used in this thesis (see Section 2.1.1.2.2, p. 95) may have been more powerful given that it was a sonic probe specifically designed to disrupt tissue. The results obtained in Section 3.1.1.2.2 supported Svensson Bonde and Bulow (2012)'s finding that heating the sample improved the extraction of r-

amelogenin (Figure 37B, p. 141). Heat treatment is a convenient cell lysis method to extract recombinant proteins from *E. coli* cells (Middelberg, 1995). By disrupting the membranes it facilitates the release of proteins. It may, however, denature some proteins (Kim et al., 2000) but to extract r-amelogenin, it was assumed to be safe to use as enamel matrix derivatives were reported to resist heat treatments “without precipitating (Gestrelius et al., 2000) or losing their bioactivities” (Nagano et al., 2004), as reviewed by Svensson, Bonde and Bulow (2012).

Following the optimisation of volume, mixing regimen and temperature, the protocol established was carried out as shown in Section 3.1.1.2.3 (p. 143). While the volume of acetic acid was optimised for the extraction of r-amelogenin WT^{+His} (Section 2.1.1.2.1, p. 95), the mixing regimen and temperature were optimised for r-amelogenin Mut^{+His} (Section 2.1.1.2.2, p. 95). For consistency, the optimisations should have been ideally carried out for both r-amelogenin WT^{+His} and Mut^{+His}, but at this stage this was not important since the final extraction method adopted (Section 3.1.1.2.3, p. 143) was highly reproducible for both r-amelogenins, WT^{+His} and Mut^{+His}.

The mass spectrometry analysis of the extracted WT^{+His} (lyophilised, desalted and lyophilised) showed 5 major peaks ranging from 24732.99 Da to 24812.20 Da (Figure 39, p. 145). According to the data generated using the Protparam tool (Gasteiger et al., 2005) the first peak at 24732.99 Da corresponded to the predicted molecular weight of r-amelogenin WT^{+His} but minus 131 Da, indicating that the N-terminal methionine had been removed by *E. coli* methionine aminopeptidase (Ben-Bassat et al., 1987), an issue that was previously reported in respect of the expression of r-amelogenins (Simmer et al., 1994, Svensson Bonde and Bulow, 2012). The other peaks, which are found 16 Da apart, correspond most likely to oxidised species (see Table 8, p. 144). The origin of this oxidation is unclear but as described in the Material and Methods (Section 2.1.1.2.3, p. 97) the acetic acid extract was lyophilised, desalted and lyophilised again before being subjected to mass spectrometry. These additional steps may cause oxidation of some amino acids (Challener, 2017).

The spectra obtained (Appendix B) for WT^{-His} and Mut^{-His} extracted in acetic acid (extraction kindly performed by Dr Sarah Myers) and subjected directly to mass spectroscopy without prior lyophilisation showed a high degree of purification with a clear peak corresponding to r-amelogenin at 20160.10 Da for WT^{-His} and 20134.2 to Mut^{-His}, which also appeared to have lost the N-terminal methionine (Appendix B). Less pronounced peaks at 20175.30 and 20149.60 Da were visible on the spectra for WT^{-His} and Mut^{-His} respectively again suggesting possible oxidation, though at much lower

levels than those observed for r-amelogenin WT^{+His} (Section 3.1.1.2.3, pp. 144 – 145). The His-tagged r-amelogenins may be more sensitive to oxidation simply because they were subjected to desalting and lyophilisation and of course the His-tag contains 3 methionines and 7 histidines, which are amino acids sensitive to oxidation (Uchida, 2003, Ji et al., 2009):
"MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSLEVLFQ"

Figures 36 (p. 139) and 37 (p.141) indicated that the r-amelogenin made up > 90% of the initial acetic acid extract. This initial extraction was effort- and cost-effective, and provided an extremely efficient clean-up, generating hundreds of milligrams of extract out of 10 g of *E. coli* (wet weight cells harvested). However, as contaminants were clearly detectable by analytical SDS PAGE with Coomassie Blue and silver staining techniques, secondary purification steps were needed to reach optimal purity for use in binding studies: (i) nickel column chromatography (Section 4.1.2.2), SEC (Section 4.1.2.3) and (ii) preparative SDS PAGE (Section 4.1.2.4).

4.1.2.2 His-tag purification was not an optimum method to purify r-amelogenin

Following expression in *E. coli* and extraction in acetic acid, which removed major contaminants, the r-amelogenin extract was subjected to standard nickel column purification. The procedures tested are summarised in Figure 73 (p. 220) (omitting all lyophilisation and desalting steps for clarity).

4.1.2.2.1 Principles and general considerations of recombinant proteins His-tag purification

Nickel affinity chromatography is a popular methodology to purify recombinant proteins by including a fusion tag at the N- or C-terminal. The His-tag is widely used (Hochuli et al., 1988) as the imidazole rings of histidine are electron donors that allow for the formation of co-ordination bonds with the immobilised nickel (Terpe, 2003) which preferentially retains the recombinant His-tagged protein on the column. The method is simple and in its basic form consists of expressing His-tagged proteins in a given organism (e.g. bacteria) followed by a first round of nickel column chromatography to remove bacterial contaminants. The His-tagged recombinant protein is isolated and then the His-tag can be removed enzymatically if a suitable cleavage site is included in the His-tag sequence. A second round of nickel column chromatography removes any remaining uncleaved recombinant protein. It may also remove the protease if this also carries a His-tag, yielding a purified product. The inclusion of His-tag in r-amelogenins for purification purposes has been previously reported (Moradian-Oldak et al., 2000,

Tarasevich et al., 2007, Tao et al., 2015). Researchers reported that the addition of a His-tag into the sequence of mouse full-length r-amelogenin rM179 increased its expression levels in *E. coli*, which they proposed to be due to the fact that the His-tag increased the hydrophilicity of the r-amelogenin (Svensson et al., 2006).

The choice of buffers for nickel column chromatography used is critical since the protein of interest needs to be solubilised in order for it to be applied to, and bind to, the nickel column and then to be successfully eluted and isolated from bacterial contaminants. His-tag cleavage buffers were also optimised to obtain a maximum enzyme activity. The enzyme-to-protein ratio was optimised (see Appendix C) so as to provide a maximum yield with minimum amounts of protease.

However, r-amelogenin was never successfully purified by the His-tag-based purification protocol, as illustrated in Figure 73 and discussed, below.

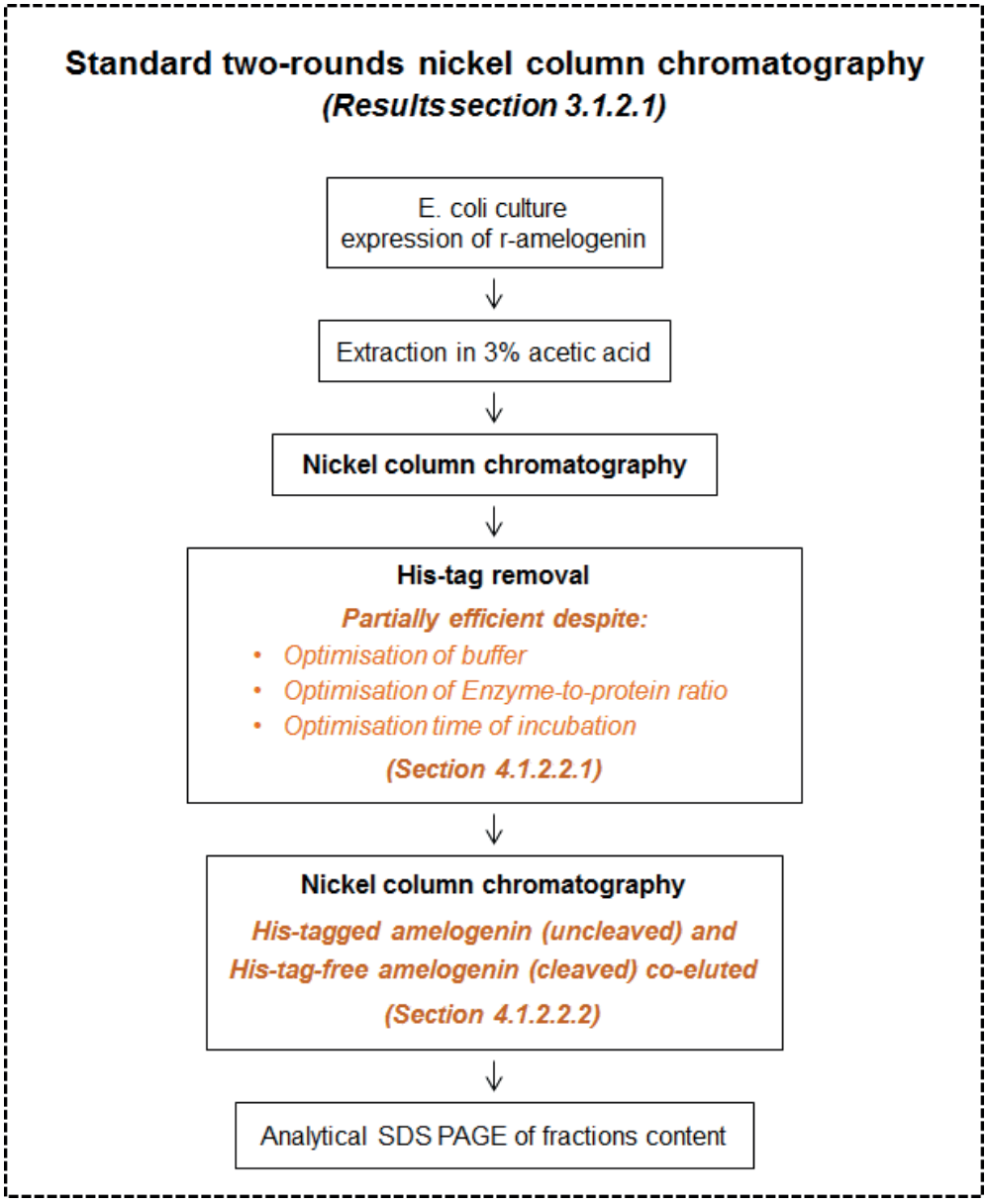


Figure 73 Summary of His-tag-based purification procedures tested and optimisations carried out. Standard nickel column chromatography included an initial round of nickel column chromatography before His-tag removal. His-tag removal was not 100% efficient and remaining uncleaved amelogenins and enzyme had to be removed by a second round of nickel column chromatography. However, a second round of nickel column chromatography after His-tag removal did not completely isolate cleaved r-amelogenin from uncleaved r-amelogenin.

4.1.2.2.2 Partial efficiency of His-tag removal from r-amelogenin

From Figure 43 (p. 153) it is clear that the efficiency of the His-tag cleavage from r-amelogenins was considerably less than 100%. Analytical SDS PAGE of the cleavage reaction mixture showed the very clear presence of cleaved r-amelogenin at 24 kDa and uncleaved r-amelogenin at 27 kDa. Increasing the concentration of cleavage enzyme or increasing the incubation time had little effect on improving the efficiency of the cleavage reaction, as described in Appendix C.

Previous reports also indicated that cleavage of the His-tag from r-amelogenin is inefficient. For example, Taylor et al (2006) used rTEV protease to cleave His-tags from r-amelogenin (Taylor et al., 2006) and noted the poor yield obtained. In general, cleavage efficiency of other recombinant proteins was as low as 50% and this was suggested to be caused by steric hindrance blocking the cleavage site and preventing enzyme access (Waugh, 2011). Amelogenin aggregation at physiological pH and temperatures has been well studied (Moradian-Oldak et al., 1994, Moradian-Oldak et al., 1995, Moradian-Oldak et al., 1998, Simmer et al., 1994, Tan et al., 1998, Wiedemann-Bidlack et al., 2007, Aichmayer et al., 2010) and amelogenin-amelogenin interactions or intra-amelogenin interactions may cause the His-tag cleavage site to be inaccessible to the HRV3C cleavage enzyme under certain conditions. An attempt was made in this thesis to increase the solubility of the r-amelogenin by carrying out the cleavage reaction at pH 8 and 9; conditions known to reduce the tendency of amelogenin aggregation (Tan et al., 1998). However, as described in Appendix D increasing the pH to 8 or 9 did not help matters.

4.1.2.2.3 Nickel column purification did not successfully isolate r-amelogenin.

Ideally, cleaved r- amelogenin, lacking a His-tag, should not be retained on the nickel column and is expected to elute in the flow-through fraction. In contrast, the uncleaved r-amelogenin, cleaved His-tag and His-tagged HRV3C protease should bind to the column. However, Figure 44 (pp. 156 - 157) showed that the second round of nickel chromatography was unable to completely separate cleaved from uncleaved r-amelogenin because the cleaved r-amelogenin appeared to show significant affinity for the column. It could be eluted by increasing the imidazole concentration to 60 mM but this resulted in the co-elution of uncleaved r-amelogenin and other contaminants (reproduced in Figure 74A below). Another secondary nickel column purification was performed independently by Dr Sarah Myers using a more refined stepped imidazole gradient (at 50, 60, 70, 90, and 200 mM imidazole) in the second round of nickel column chromatography. SDS PAGE analysis of these fractions, shown in Figure 74B, showed that cleaved and uncleaved r-amelogenins were visible in all eluted fractions

with the higher concentrations of imidazole eluting increased amounts of uncleaved r-amelogenin .

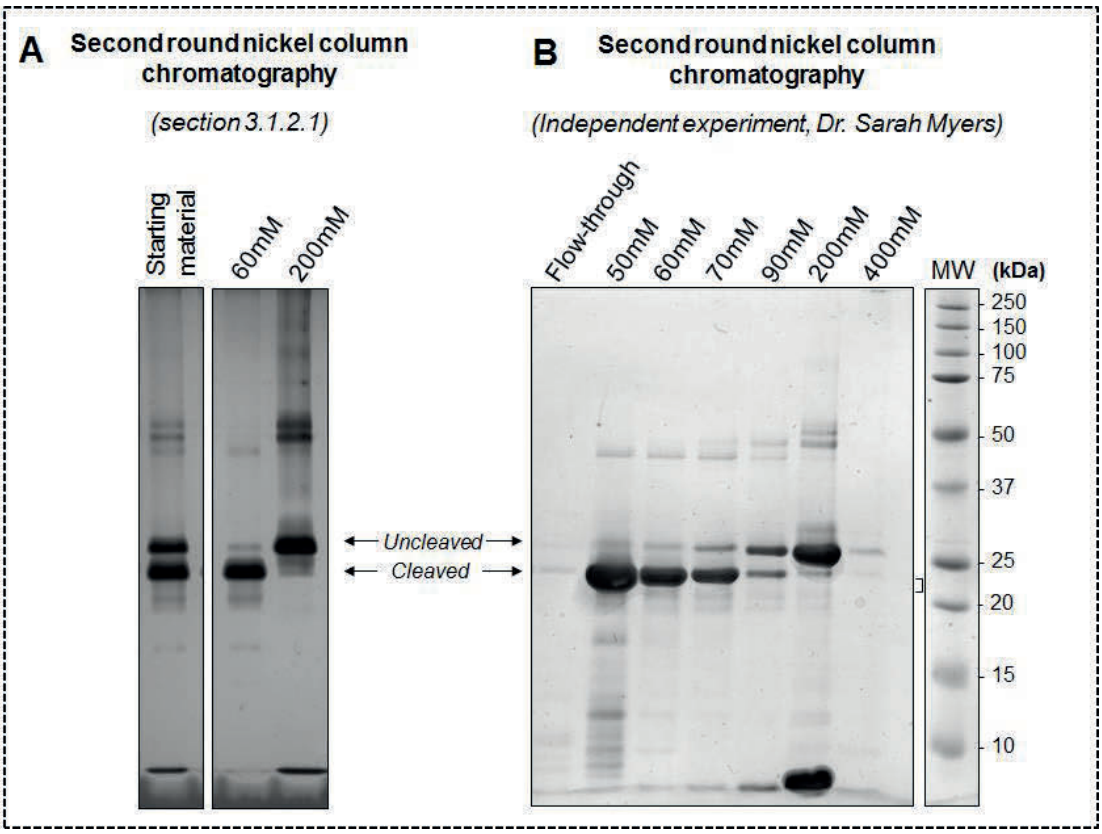


Figure 74 Comparison of isolation efficiency of cleaved r-amelogenin by (A) two-round nickel affinity chromatography or (B) second round of Nickel column chromatography with 'refined' stepped elution at 50, 60, 70, 90, 200 mM imidazole. The cleaved r-amelogenin, which does not carry a His-tag, was not collected in the flow-through as expected but was eluted by increasing the concentration of imidazole to 50 mM and higher concentrations. At 50 mM imidazole, the bulk of cleaved r-amelogenin was collected with a minor trace of uncleaved r-amelogenin. At 60 and 70 mM imidazole, significant amounts of cleaved r-amelogenin were clearly co-eluted with the uncleaved r-amelogenin, with a higher proportion of uncleaved r-amelogenin at higher concentrations of imidazole. The bulk of uncleaved r-amelogenin was eluted at 200 mM imidazole . A trace amount of HRV3C protease (His-tagged) is visible in elution fractions at 50 - 200 mM imidazole. Figure 74 is taken from Gabe et al, 2017[1]; the data for Figure 74B was acquired by Dr Sarah Myers.

[1] Figure 74 is reproduced from Gabe et al. (2017) under the terms of the Creative Commons license CC BY (https://creativecommons.org/licenses/by/4.0/).

- **Retention of His tag free cleaved r-amelogenin to the nickel column**

The reason why the cleaved r-amelogenin still has affinity for the nickel column may be due to the high histidine content of amelogenin (14 residues in M180). Moreover, as illustrated below, amelogenin includes di- and tri-histidine repeats which may act as a pseudo His-tag and cause non-specific binding to the nickel column:
MPLPP**H**PGSPGYINLSYEVLTPLKWYQSMIRQPYPSYGYEPMGGWL**HH**QIIPVLSQQ
HPPS**H**TLQP**HHH**LPVVPAQQPVAPQQPMMPVPG**HH**SMTPTQ**HH**QPNIPPSAQQPF
QQPFQPQAIPPQS**H**QPMQPQSPL**H**PMQPLAPQPPLPPLFSMQPLSPILPELPLEAWP
ATDKTKREEVD
If amelogenin is an unstructured protein as described in Section 1.2.1.3.2 (p.19) the di- and tri-histidine repeats could be in close proximity in space and form co-ordinate complexes with the nickel in a similar fashion to the hexahistidine motif making up the His–tag even though they may be distant from each other in terms of their linear position in the amelogenin primary sequence.
As described in the Section 1.2.2.1 (pp. 20 - 21) amelogenin has a strong tendency to aggregate and one other explanation for the retention of cleaved His-tag free r-amelogenin on the nickel column is that the cleaved r-amelogenin is simply interacting with the uncleaved r-amelogenin that is strongly bound to the column by virtue of its His-tag. However, this is unlikely since the imidazole elution buffers were 4 M with respect to urea; a chaotropic agent well known for its ability to solubilise amelogenin (Eggert et al., 1973, Brookes et al., 2002). In addition, from Figure 73B it is clear that cleaved r-amelogenin is eluted with 50 mM imidazole whereas the uncleaved r-amelogenin remains bound to the column. Imidazole is not a chaotropic agent and this data indicates that the cleaved r-amelogenin was being displaced from the nickel column by competing imidazole molecules as opposed to being solubilised from an aggregated state involving column-bound uncleaved r-amelogenin.

To recap, His-tag-based purification methodology was not suitable to obtain r-amelogenins for future uses because:

1) His-tag cleavage with HRV3C protease was not efficient.

2) Cleaved r-amelogenin could not be isolated from uncleaved r-amelogenin using nickel column chromatography due to the apparent interaction of cleaved r-amelogenin with the nickel column.

Although nickel chromatography is the standard methodology for purifying His-tagged proteins, it is reported elsewhere that the methodology can provide satisfactory yields of recombinant proteins but the purity of the proteins obtained is not always ideal (Lichty et al., 2005).

Practically, the necessary inclusion of urea in the buffers to solubilise the amelogenin is not compatible with the enzymatic removal of the His-tag. The urea denatures the HRV3C enzyme, and according to the manufacturer, will completely inhibit its activity. This incompatibly required the use of an additional time-consuming desalting step prior His-tag removal (see Figure 41 p. 148 in Section 3.1.2.1). Since His-tag-based purification was not an optimum method to obtain r-amelogenin for functional studies, other purification techniques were tested and optimised. Their respective efficiencies are addressed and compared in the Sections 4.1.2.3 and 4.1.2.4.

4.1.2.3 Size exclusion chromatography on Bio-gel P columns did not isolate r-amelogenin from lower molecular weight contaminants

SEC separates proteins within a mixture according to their apparent molecular weights or more accurately according to their hydrodynamic radii (Grubisic et al., 1967). Larger molecules are excluded from the column matrix and quickly pass through the column carried along in the mobile phase and are eluted first while lower molecular weight components with a smaller hydrodynamic radius can diffuse from the mobile phase into the beads. Their progress through the column is retarded and depends on their rate of diffusion in and out of the beads.

Figures 46 - 48 (Section 3.1.2.2, pp.163, 165, 166) showed that none of the SEC methods employed to isolate r-amelogenin from bacterial contaminants present in the crude acetic acid extracts was successful; this is despite adjusting the column pore size and the column length. The very first fraction to elute from the columns was reasonably pure r-amelogenin but the bulk of the r-amelogenin was eluted over several fractions and these fractions were contaminated with bacterial low molecular weight contaminants (previously extracted, with r-amelogenin, in acetic acid).

SEC was carried out using 125 mM formic acid which solubilised the r-amelogenin. Formic acid in conjunction with Bio-gel P-30 (as used here) was successfully used to isolate the full-length amelogenin from the lower molecular weight enamel matrix proteins present in extracts of developing enamel matrix (Fincham et al., 1981). However, the fact that most of the r-amelogenin co-eluted anomalously along with the low molecular weight contaminants suggested that formic acid was not effective in eliminating interactions between the r-amelogenin and the low molecular weight contaminants or perhaps more likely, did not prevent interactions between the r-amelogenin and the column matrix (which would retard the migration of r-amelogenin down the column such that it eluted with the low molecular weight contaminants. Another possibility is that amelogenin, described as an IDP (Delak et al., 2009) may adopt a continuum of different conformations, each with different hydrodynamic volumes or shapes. These different conformers of r-amelogenin may diffuse at different rates and co-elute with contaminants of their respective hydrodynamic radii.

4.1.2.4 Preparative SDS PAGE successfully purified r-amelogenin from acetic acid extracts

While r-amelogenin was co-eluted with lower molecular weight contaminants on SEC, an alternative 'molecular sieving' method, preparative SDS PAGE, avoided this issue. Analytical SDS PAGE was used in this thesis to check the effectiveness of nickel column chromatography and SEC at resolving r-amelogenin from contaminants. It was apparent that the technique itself can successfully resolve r-amelogenin from contaminants. It is possible to cut bands out of the gels and extract the protein but the yield is in the microgram range at best. However, large scale preparative SDS PAGE promised the resolving power of analytical SDS PAGE with yields associated with chromatographic methods and was successfully trialled as an alternative to the chromatographic methods tested.

sodium dodecyl sulphate (SDS) is an amphoteric detergent that binds to proteins via its dodecyl aliphatic tail through hydrophobic interactions (Laemmli, 1970, Reynolds and Tanford, 1970a). The sulphate group the effectively covers the protein with negative charges which repel each other and force linearization of the protein so that it is denatured and takes on a rod like conformation (Reynolds and Tanford, 1970b) as indicated in Figure 75 below. The proteins thus attain a more or less constant charge to mass ratio and when driven through a sieving medium (cross linked acrylamide) by an electric field they migrate as a function of their molecular size (Reynolds and Tanford, 1970b).

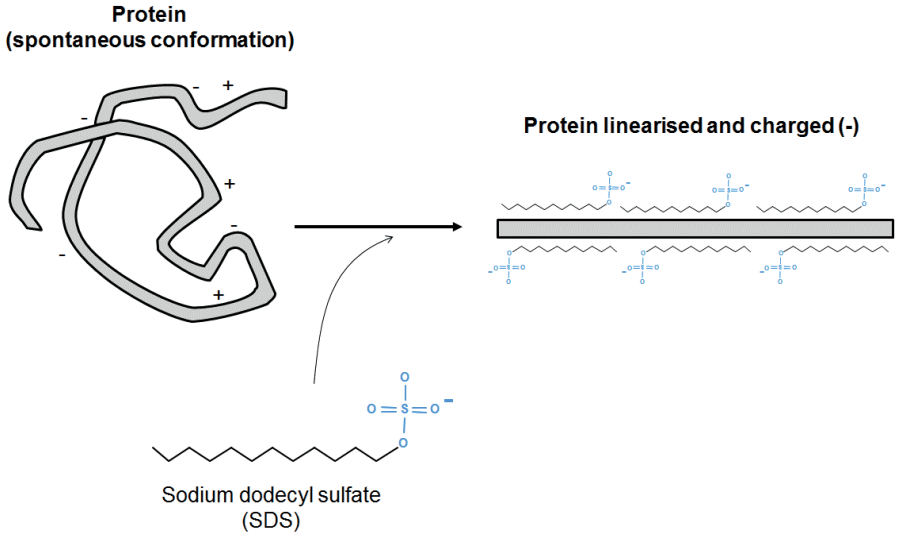


Figure 75 Principle of protein linearisation by SDS. SDS is a detergent that binds to proteins by hydrophobic interactions (Reynolds and Tanford, 1970a). It covers the proteins with negative charges and linearises them, which allow to separate them according to their molecular weight (which will be directly proportional to their molecular sizes in this case).

When SDS linearises amelogenin, every molecule will attain the same rod-like conformation reducing the continuum of hydrodynamic radii which means that all molecules will migrate through the molecular sieve at the same rate as a tight, well resolved band.

The results obtained (Section 3.1.2.3, pp. 167 - 178) showed that preparative SDS PAGE permitted the isolation of r-amelogenin to single-band purity. Furthermore, the resolving power was such that it not only separated His-tagged r-amelogenins from bacterial contaminants in the acetic acid extracts (Figures 50-51, pp. 170 - 171) but it also separated cleaved from uncleaved His-tagged r-amelogenins (Figure 54 p. 177). The purification procedures involving preparative SDS PAGE are summarised below in Figure 76. The results obtained indicated that the His-tag was redundant as it appeared likely that r-amelogenin, engineered without a His-tag, could be purified to single band purity from the crude acetic acid extracts simply be a single round of preparative SDS PAGE.

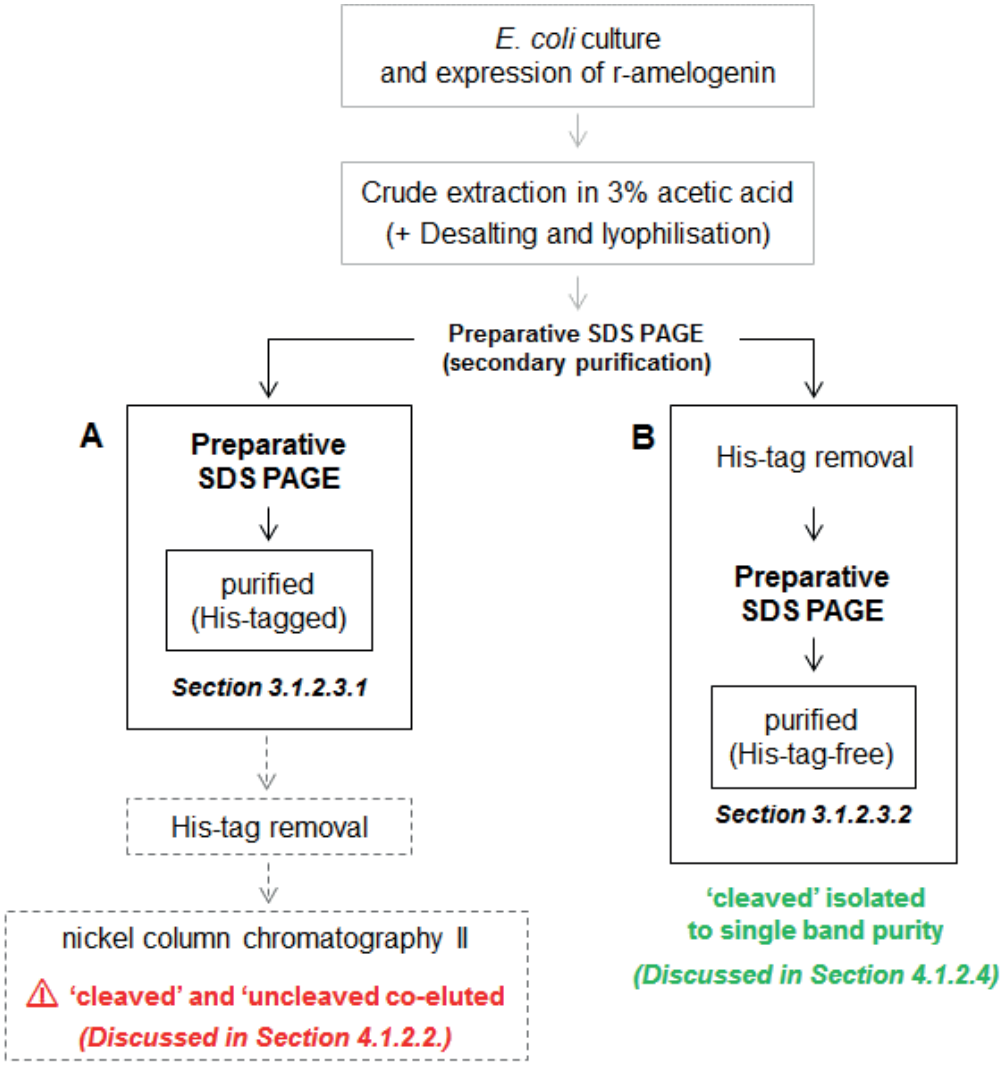


Figure 76 Secondary purification of r-amelogenin using preparative SDS PAGE. (A) Initially the strategy was to use preparative SDS PAGE to purify uncleaved r-amelogenin from the crude acetic acid extract (Section 3.1.2.3.1), cleave off the His-tag and then separate the cleaved from the uncleaved r-amelogenin by nickel column chromatography. However cleaved and uncleaved proteins were never completely separated by nickel column chromatography (discussed in Section 4.1.2.2, p. 222). (B) An alternative method (which omitted the nickel chromatography) was simply to subject the proteins in the crude acetic acid extract to His-tag cleavage and then use preparative SDS PAGE to isolate the cleaved r-amelogenin from uncleaved r-amelogenin and contaminating bacterial proteins. This method gave 'cleaved' r-amelogenin at single band purity on analytical SDS PAGE with silver staining (Section 3.1.2.3.2) thus rendering the His tag redundant. (Discussed in Section 4.1.2.4, pp. 230 - 231).

4.1.2.4.1 Preparative SDS PAGE provided a high resolution purification of His-tagged r-amelogenin from crude acetic acid extracts

Coupled with acetic acid extraction, which initially enriched the amelogenin fraction (discussed in Section 4.1.2.1, p. 216), preparative SDS PAGE allowed for the isolation of cleaved and uncleaved r-amelogenins and 20 kDa amelogenin from EMD to single band purity as determined using analytical silver stained SDS PAGE (Figures 51, 54 and 57 in pp. 171 - 182). This Section focuses on the isolation of WT^{+His} r-amelogenin using preparative SDS PAGE.

Figures 50, 51 (pp. 170 - 171), showed that fractions 25 to 49 contained WT^{+His} r-amelogenin migrating at an apparent MW of 27 KDa. The earlier fractions, 22- 23, running at a slightly lower molecular weight of 24.5 kDa also stained well with silver staining. However, western blotting of this gel, using a polyclonal anti-amelogenin to the 12 amino acid C-terminal telopeptide, showed that the immunoreactivity was reduced for the bands in these earlier slightly smaller fractions. This suggests that the antibody had a weaker affinity for the epitope (i.e. the telopeptide). This together with the slightly lowered molecular weight suggests that the telopeptide may have been attacked by exopeptidases which shortened the telopeptide sequence. In effect, this would have shortened the size of the epitope and reduced the proportion of polyclonal antibodies still able to recognise the truncated sequence leading to a reduced immunological staining on the blot. It appears therefore, that the r-amelogenin may be susceptible to C-terminal degradation by exopeptidases, possibly of bacterial origin. Fortunately, the amount of r-amelogenin affected (i.e. fractions 22-23) is relatively small compared to the bulk of protein present in fractions 25-49 that corresponds to the full length r-amelogenin. It should be noted that this apparent C-terminal degradation was only discovered thanks to the very high resolving power of preparative SDS PAGE. The degradation of the C-terminal telopeptide would have no effect on the His-tag and therefore r-amelogenin purified by nickel column chromatography would have generated a mixed fraction containing full length r-amelogenin and r-amelogenin suffering from C-terminal truncation.

4.1.2.4.2 Quantitative yield of r-amelogenin generated by acetic acid extraction coupled with preparative SDS PAGE

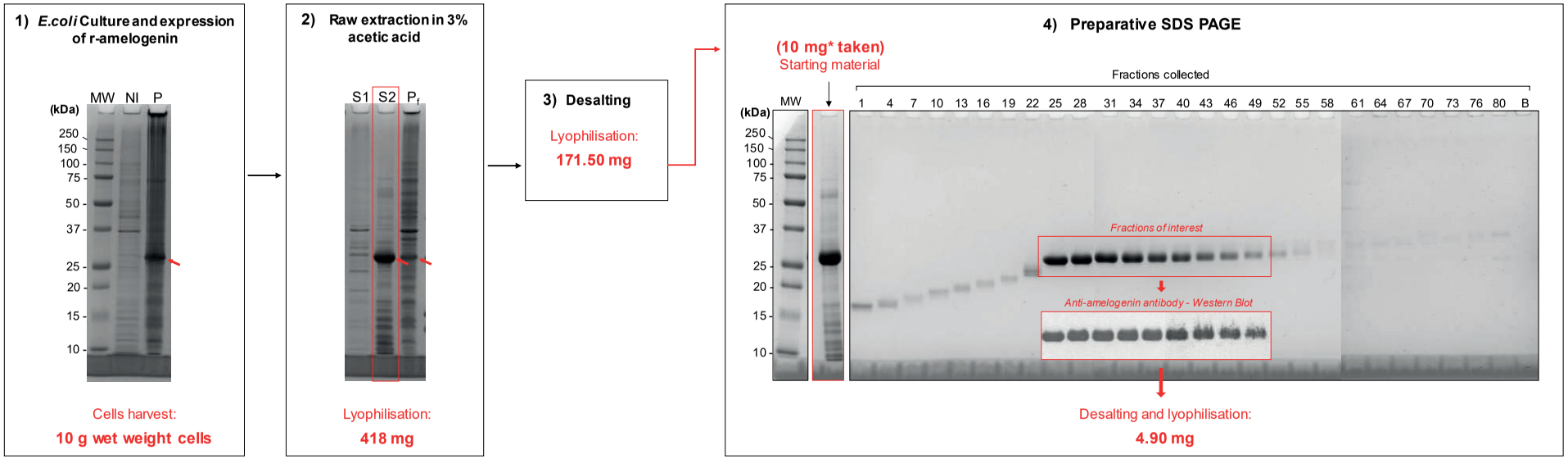
As shown in Figure 77 below (p.232), 10 g of *E. coli* from 1.6 L of culture generated 418 mg of crude lyophilised acetic acid extract. Following desalting and lyophilisation of this extract, 171.5 mg of WT^{+His} r-amelogenin protein were obtained. Ten milligrams of this were subjected to preparative SDS PAGE which yielded a total of 4.9 mg of protein following desalting and lyophilisation. Working backwards, this equates to a

yield of 52.5 mg/L bacterial culture, which compares well with the typical yield of recombinant protein technology which ranges from 4 to 200 mg/L (Lehmann et al., 2003, Qing et al., 2004, Sivashanmugam et al., 2009). It is all the more satisfactory since the r-amelogenin is obtained at single band purity on silver stained analytical SDS PAGE. A first round of His-tag nickel column chromatography (Section 3.1.2.1 pp. 149 - 151) yielded 14.5 mg r-amelogenin out of 37.4 mg of desalted extract; the yield, calculated as 41.5 mg per litre cell culture (Figure 78 below, p. 233), was therefore slightly lower, and crucially, was not at single band purity compared to preparative SDS PAGE. The yield and purity of the final r-amelogenin fractions obtained after preparative SDS PAGE and nickel column chromatography are compared below in Table 13.

Table 13 Strategies employed to purify r-amelogenin: Comparison of the yield and purity obtained in the final fractions after preparative SDS PAGE or nickel column chromatography. This table highlights the observations reported in Figures 77 and 78 (overleaf).

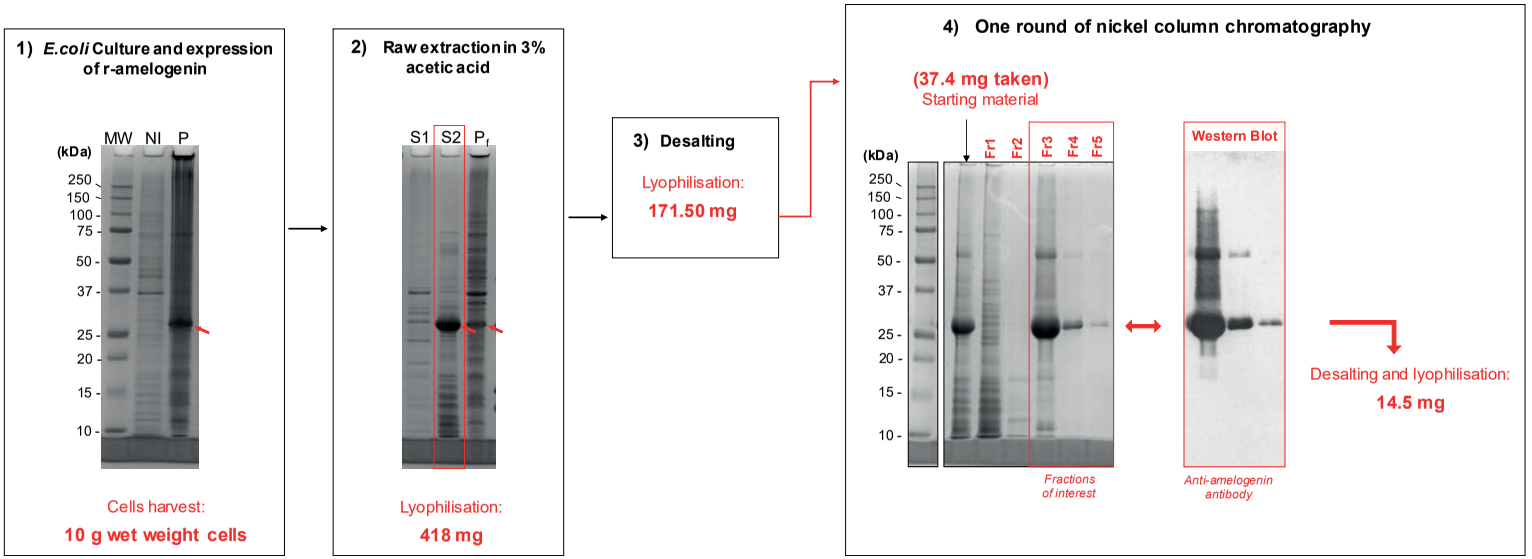
Method	Yield	Purity
Preparative SDS PAGE-based purification (Figure 77)	52.5 mg proteins per litre cell culture	Single band on analytical SDS PAGE with Coomassie Blue staining and with silver staining
Nickel column chromatography-based purification (Figure 78)	41.5 mg proteins per litre cell culture	Major band on analytical SDS PAGE with Coomassie Blue staining with visible contamination

Figure 77 Yield of His-tagged r-amelogenin obtained by coupling raw acetic acid extraction with preparative SDS PAGE



⇨ 1.6L *E. coli* culture (10g wet weight cells) can generate up to 84 mg r-amelogenin

Figure 78 Yield of r-amelogenin obtained by method coupling raw acetic acid extraction with one round of nickel column chromatography.



⇨ 1.6L *E. coli* culture (10g wet weight cells) can generate up to 66.5 mg r-amelogenin

- 233 -

4.1.3 Impact of preparative SDS PAGE on r-amelogenin production

4.1.3.1 Preparative SDS PAGE by-passes the need for a His-tag

As shown above in Figures 77 and 78 (Section 4.1.2.4.2), preparative SDS PAGE appeared to be superior to nickel column chromatography for purification of His-tagged r-amelogenin away from the milieu of bacterial contaminants present in the crude acetic acid extracts, giving a yield of ~50 mg His-tagged r-amelogenin per litre of bacterial culture. The initial plan was that the purified His-tagged r-amelogenin would then be subjected to His-tag cleavage and the cleaved r-amelogenin purified by nickel chromatography (Figure 76A, p. 229). However, the degree of resolution provided by preparative SDS PAGE was such that a simplified method was developed that involved subjecting all of the proteins present in the acetic acid extract to His-tag cleavage and then purifying the cleaved r-amelogenin from the uncleaved r-amelogenin and bacterial contaminants using preparative SDS PAGE (Figure 76A, p. 229). When this second approach was used by subjecting 10 mg of the cleaved lyophilised acetic acid extract to preparative SDS PAGE the final yield of purified cleaved r-amelogenin was in the order of 1 mg, which equates to around 10 mg of cleaved r-amelogenin per litre of bacterial culture. This apparent reduction in yield is a reflection of the fact that in this case, the 10 mg of starting material contains a mixture of cleaved and uncleaved r-amelogenin so there is less target protein in the 10 mg of starting material to begin with. In addition, the preparative SDS PAGE fractions containing the cleaved and uncleaved r-amelogenins overlapped slightly (see fractions 45 -47 in Figure 54, p. 177), so not all fractions containing the cleaved r-amelogenin were pooled to obtain the final yield.

In summary then, the results obtained not only support that a His-tag is now redundant to purify r-amelogenin but also show that the presence of a His-tag is in itself problematic as its removal is not 100% efficient; this complicates the purification process, particularly affecting the yield. The excellent resolving power of preparative SDS PAGE, coupled with the acetic acid extraction technique provides for a far better approach to obtaining purified r-amelogenins and that is to engineer a His-tag-free r-amelogenin that can be extracted in acetic acid and purified to single band purity from the contaminating bacterial proteins by preparative SDS PAGE. Abandoning the His-tag would dramatically reduce costs as the HRV3C protease is expensive and it would provide a rapid means of delivering mg quantities of His-tag-free r-amelogenin using one round of purification only. Crucially, the yield would be expected to return to ~50 mg of recombinant per litre of bacterial culture since the purification process would be exactly the same as that shown in Figure 77 above (p. 232). The only difference would be that the *E. coli* would be expressing His-tag-free r-amelogenin as opposed to His-

tagged r-amelogenin. For that, the choice of expression system is important, considering Svensson et al (2006)'s observation that the addition of a fusion His-tag increased bacterial growth and r-amelogenin expression (Svensson et al., 2006) as discussed earlier in Section 4.1.2.2 (p. 219). Svensson et al. (2006) used BL21(DE3) *E. coli* cells transfected with pET11a vector, which may be sensitive to the properties of the r-amelogenin they express.

An attempt was made to engineer an expression vector to express His-tag free r-amelogenin but the commercial partner undertaking the work experienced undisclosed technical difficulties in producing the vectors. Functional vectors resulting in the expression of both WT and mutant Y64H r-amelogenins at comparable levels to the vectors expressing the His-tagged r-amelogenins were finally delivered during the write-up period of this thesis but as discussed elsewhere, time limited their use experimentally.

4.1.3.2 Preparative SDS PAGE may provide a route to r-amelogenin produced by eukaryotic expression systems

One drawback of expressing amelogenin using bacterial expression systems is that Ser^{16} will not be phosphorylated as it is in the native protein. However, eukaryotic systems will produce correctly phosphorylated amelogenin.

Taylor et al. (2006) expressed human r-amelogenin in an eukaryotic baculovirus system which yielded 10 mg r-amelogenin per litre culture using nickel chromatography. However, the r-amelogenin was sensitive to degradation within the Sfx insect cells. These degradation products, presumably many of them carrying the His-tag, were clearly present in the nickel column elution fractions on analytical SDS PAGE following Coomassie Blue staining (Taylor et al., 2006) and full-length r-amelogenin could not be isolated to single band purity. The potential of preparative SDS PAGE to provide a ready source of purified phosphorylated r-amelogenin is an exciting prospect for the field.

4.1.3.3 Preparative SDS PAGE purification may increase protein oxidation.

The previous sections highlight the benefits of preparative SDS PAGE for purifying r-amelogenin. However, in the interests of providing a balanced and critical discussion it is important to mention one possible drawback; increased protein oxidation linked to SDS PAGE.

The mass spectra obtained from Figure 39 (p. 145) and Figure 52 (p. 173) are compared in this Section (Figure 79, below). Figure 79 shows numerous species corresponding to r-amelogenin carrying multiples of 16 Da (i.e. oxygen) in both the crude acetic acid extract and the r-amelogenin purified by preparative SDS PAGE. The number of species exhibiting evidence of oxidation was greater in the sample subjected to preparative SDS PAGE.

- **Preparative SDS PAGE caused oxidation**

The polymerisation of the acrylamide/bis acrylamide monomers to form the polyacrylamide gel used in SDS PAGE gels involves the oxidising agent ammonium persulphate (see Section 2.1.1.4, pp. 100 - 101). Excess persulphate in SDS gels is known to oxidise proteins as they migrate through the gel and histidine, cysteine and methionine side chains are particularly sensitive to oxidation. However, this problem has been addressed by Sun and Anderson, 2004 who advocated thorough degassing of the gel solutions to remove molecular oxygen and replacing the persulphate triggered polymerisation system with a photopolymerisation system using flavin as an alternative initiator. This, together with adding the antioxidant thioglycolate to the running buffer, reportedly eliminated protein oxidation during SDS PAGE (Sun and Anderson, 2004). The potential effect of protein oxidation on protein structure and function clearly needs to be considered and it would be preferable in future studies to take the preventative steps described above to protect the r-amelogenin from any oxidation incurred during preparative SDS PAGE.

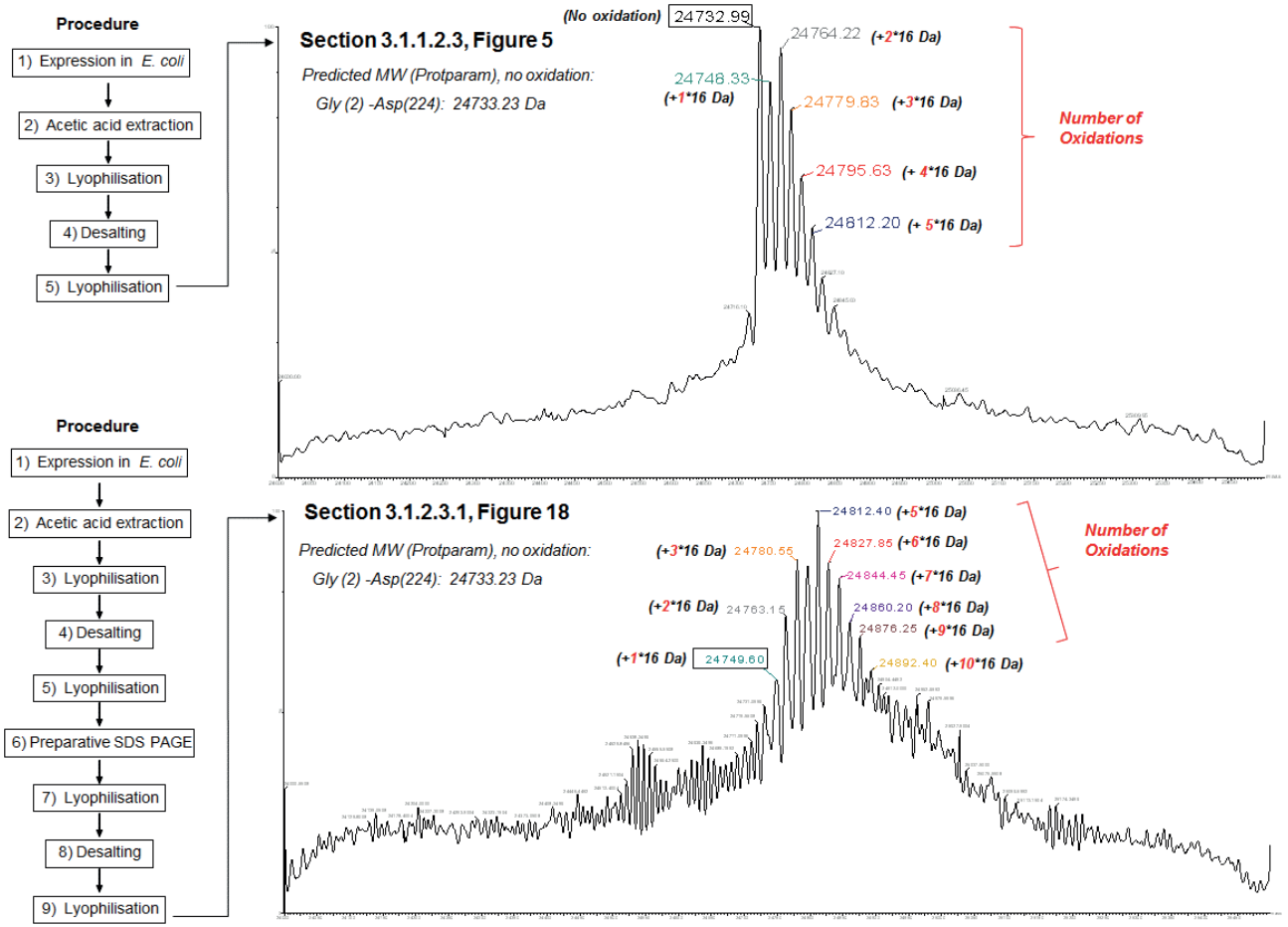


Figure 79 Comparison of spectra obtained for His-tagged WT r-amelogenin (A) extracted in 3% acetic acid and (B) isolated by preparative SDS PAGE. The main peaks differ by multiples of 16 Da which was presumed to be due to methionine or histidine oxidation.

4.2 Development of microplate binding studies to dissect the molecular mechanisms in AI.

As described in Section 1.3.4.1.3 (p. 70), the first evidence of a mechanism underlying AI was published in 2014 using mice carrying a p.Y64H mutation. This mechanism was proposed to be driven by ER stress that subsequently triggered a pro-apoptotic UPR. ER stress generally arises following pathological interactions (aggregation) of affected proteins transiting the ER and p.Y64H amelogenin in affected mice was abnormally retained in within the ameloblasts (Brookes et al., 2014). A compound, 4-phenyl butyrate, reported to relieve ER stress (Iannitti and Palmieri, 2011), rescued the phenotype in heterozygous female mice (Brookes et al., 2014) providing the first possibility that AI could be relieved therapeutically. However, further work is needed to further dissect the underlying mechanism; especially in terms of how the p.Y64H mutation affects amelogenin-amelogenin interactions and in turn how potential therapeutic agents might modulate such interactions or modulate the UPR. Key questions are: How did the single-point mutation trigger ER stress and a subsequent pro-apoptotic UPR? How did 4-PB rescue the phenotype? Further studies can help answer these questions.

This section (4.2) describes initial trials of microplate based binding tests for use in studying amelogenin-amelogenin interactions.

4.2.1 A fluorescence labelling-based microplate binding assay was not successful.

The fluorescence labelling-based microplate binding assay consisted of immobilising amelogenin (bait protein) to microwell surfaces and adding FITC-labelled amelogenin (the free ligand in solution) to allow it to bind to the unlabelled bait on incubation. After incubation, microwells would be washed and any amelogenin-amelogenin binding would be determined by measuring the remaining fluorescence. This would provide a simple method to analyse the effect of the p.Y64H mutation on amelogenin binding behaviour. First, to try and establish the conditions required to saturate the microwells with the bait amelogenin, a preliminary experiment was conducted using FITC-labelled amelogenin as a bait (see Section 2.2.1.3, pp. 122 – 123). Different concentrations and incubation conditions were tested and the protocol carried out (Figure 29 p. 124) was the same as that planned for the actual binding assay except that no FITC-labelled amelogenin was added as “free ligand” and the fluorescence measured at the end was to reflect the amount of bait adsorbed.

In order to determine the amount of bait r-amelogenin required to saturate all binding sites within the wells, FITC-labelled EMD was incubated in the microwells overnight in PBS, TBS or bicarbonate (Section 3.2.1.3, p. 184). The fluorescence signals obtained following coating of the microwells reached clear plateau values of 0.05141 (PBS, 4°C) and 0.04167 (bicarbonate, 4°C) fluorescence units (Figure 59, p. 186). These were reached at low starting concentrations of EMD (less than 10 µg/mL). This suggests that the protein may have only formed a monolayer and that no protein-protein interaction occurred. With TBS, the signal kept going up, suggesting it coated the microwell and then interactions between free and immobilised 20kDa amelogenins made the signal keep going up. The signals obtained by incubating the amelogenin solutions at room temperature (20°C) were higher than at 4°C, suggesting that higher temperatures of incubations for coating favoured amelogenin binding to microplates surfaces. This is understandable since amelogenin propensity to aggregate increases with the temperature (Moradian-Oldak et al., 1998).

Remarkably, the signals in all cases were very weak (below 0.1 on average), even though labelled protein may have saturated the microwell surfaces. This raised doubts as to whether this method would be sensitive enough to accurately determine the binding of labelled protein free in solution to unlabelled protein immobilised on the microwells surfaces. At the very least, it was clear that FITC as a label was not sensitive enough for the purpose.

FITC covalently modifies positive lysine R-groups and the N-terminal portion of proteins, and so, could modify their binding behaviour. This is what may have

happened with the surrogate 20 kDa amelogenin from EMD, interfering with its immobilisation on the polystyrene surface of the microwell. A similar issue was reported in other works using BSA, as FITC-labelling altered its blocking abilities: FITC-labelled BSA did not withstand PBS or PBS-0.1% Tween washing (Ahirwar et al., 2015) unlike unlabelled BSA, which is a gold standard blocker in ELISA protocols (Gibbs, 2001).
In the case of the actual proposed binding assay, unlabelled amelogenin would have been bound to the microwells as bait. In this case, lysine modification might be an advantage as FITC-labelled amelogenin ligand would not bind the microwells. This would avoid nonspecific binding to the polystyrene. However at this stage, the robustness of fluorescence-based microplate binding assays could not be demonstrated as the preliminary study was not successful.

A second method to study amelogenin-amelogenin binding was developed based on UV-absorption spectrometry to monitor the depletion of solubilised amelogenin from solution in real time as it bound to amelogenin immobilised on the microwell surfaces. As discussed in the next Section, this avoided the need for labelling or covalent modifications and no washing steps were involved. The method, utilising UV-transparent microplates could, in theory, be used to determine the kinetics of protein-protein interactions as proteins were depleted from solution over time.

4.2.2 Protein depletion from solution in UV-transparent microplates provided with an end-point measurement and limited kinetic information

Unlabelled amelogenin was incubated in UV transparent microwells and the decrease in UV absorbance was monitored over time, reflecting the depletion of amelogenin from solution, as amelogenin initially adsorbed to the microwell surfaces (protein-polystyrene binding) and then amelogenin remaining free in solution binding to the previously immobilised amelogenin (protein-protein binding). The hope was that this would then permit determination of the kinetics of amelogenin-amelogenin binding (illustrated in Figure 80).

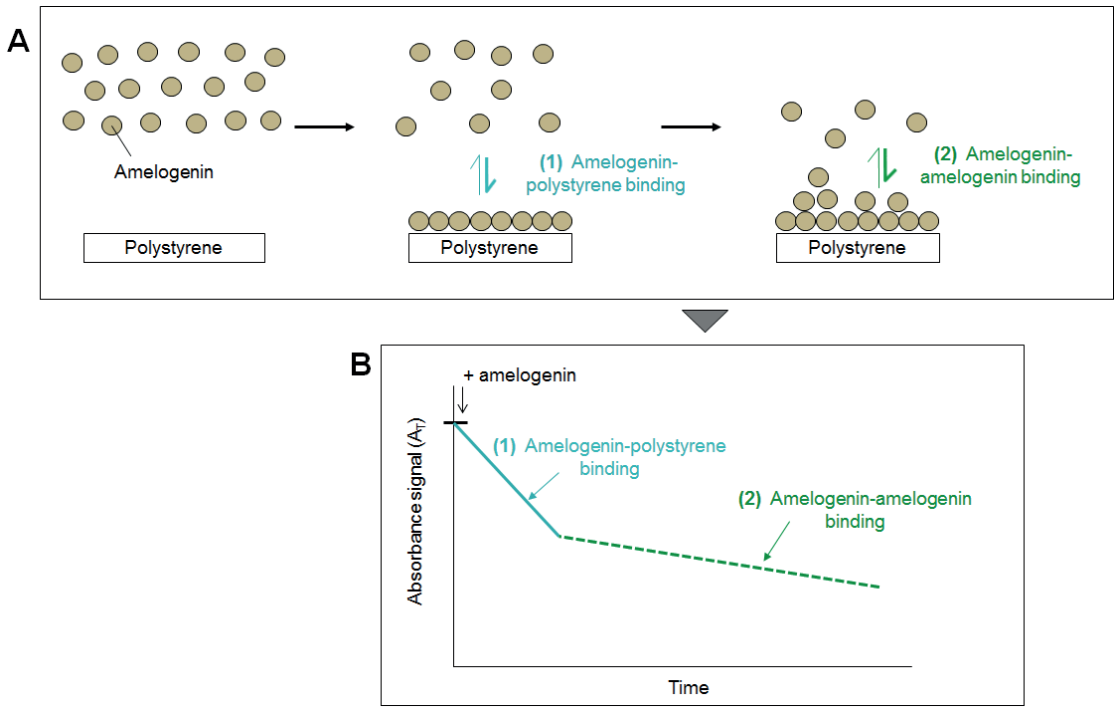


Figure 80 Principle of protein (amelogenin) depletion assay using UV-transparent microwells. Panel (A) illustrates the initial hypothesis that (1) amelogenin would bind to the polystyrene and then (2) free amelogenin would bind to immobilised amelogenin. Panel (B) displays the decrease of absorbance expected, reflecting the interactions occurring in (A). After phase (1) of amelogenin-polystyrene binding, there is an inflexion point and the second slope (2) would reflect previously free amelogenin binding to the immobilised amelogenin. The assumption is that amelogenin-polystyrene binding occurs more readily than amelogenin-amelogenin binding so that the microwells are saturated before amelogenin-amelogenin binding begins. The aim is to characterise phase (2), that is, amelogenin-amelogenin binding.

4.2.2.1 The absorbance decrease reflected the major EMD proteins binding to the polystyrene surfaces

4.2.2.1.1 Amelogenin bound to polystyrene surfaces

In all tests carried out, absorbance decreased throughout the incubation period (Section 3.2.2, Figures 61 p. 192, 64 p. 198, 66 p. 202, 68 p. 205), which indicated that EMD, used as a surrogate for r-amelogenin, bound to the polystyrene surface. Consistently, analytical SDS PAGE with Coomassie Blue staining showed that the concentration of amelogenins left in solution (looking at the 20 kDa band) was, as expected, reduced after incubation (Figures 61 p. 192, 64 p. 194, 68 p. 205).
The binding of hydrophobic amelogenin to polystyrene was expected as polystyrene (Figure 81 below) is hydrophobic. Amelogenin has a high content of hydrophobic amino acids such as proline and leucine. The binding of leucine to polystyrene was demonstrated by studies on LK_{14} peptide (Mermut et al., 2006). The adsorption of proteins to polystyrene is mediated by their hydrophobic regions (Norde et al., 1995). This is likely mediated by hydrophobic interactions between the benzene rings of polystyrene and hydrophobic amino acid residues (Qiang et al., 2017), as illustrated in Figure 81 below.

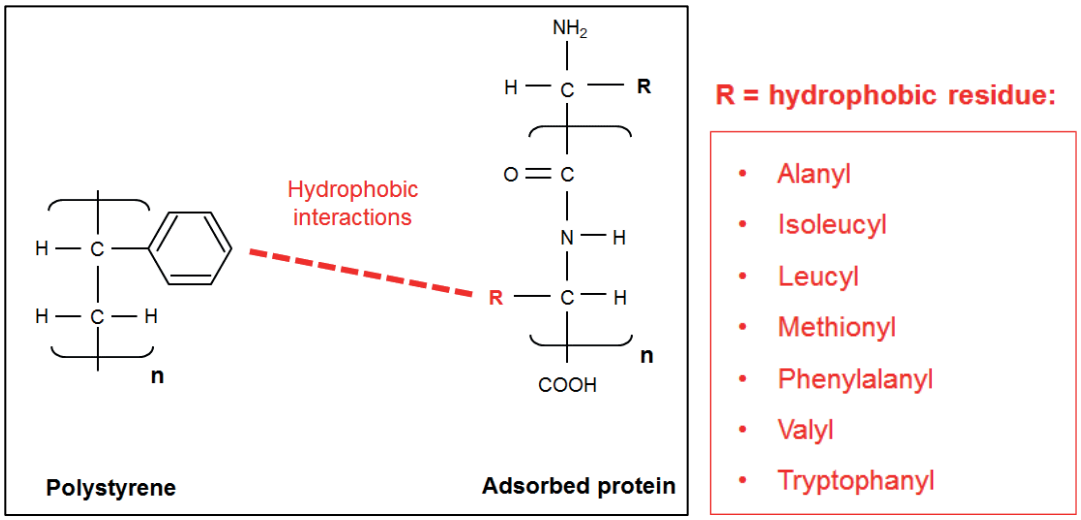


Figure 81 Principle of adsorption of EMD proteins on polystyrene surface. The adsorption of protein on polystyrene is mediated by hydrophobic interactions and the residues involved are most likely those of hydrophobic amino acids (red box, on the right).

However, analytical SDS PAGE of proteins remaining unadsorbed from solution indicated that low molecular weight (< 10 kDa) components were not adsorbed to any degree during the incubation period (shown in Figures 61, p. 192 and 64, p. 198). These could be soluble components of EMD such as amelogenin degradation

products (Brookes et al., 1995) which may not be able to engage in hydrophobically driven interactions and therefore did not bind to hydrophobic polystyrene.
The first step in developing the method was to establish the starting concentration of EMD sufficient to saturate the microwell surfaces. The results in Section 3.2.2.1 (pp. 191 – 192) showed that the equilibrium was reached after 10 hours of incubation as the absorbances plateaued from that point onwards (Figure 61A p. 192). The polystyrene microwell surfaces could be saturated using EMD at an initial concentration ranging 99.6 to 166 µg /mL. For both 99.6 and 166 µg/mL as starting concentrations, the contents of solution at equilibrium displayed equal signals on analytical SDS PAGE after Coomassie Blue staining in Figure 61B (p. 192). This was illustrated in Figure 62 (p. 193), reproduced below in Figure 82.

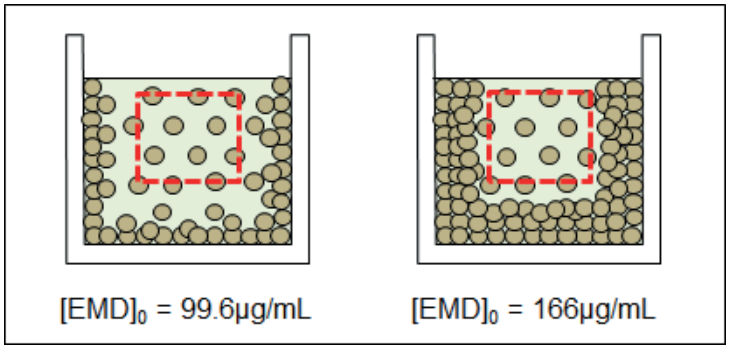


Figure 82 Saturation of the polystyrene surface was obtained using EMD at initial concentrations of 99.6 - 166 µg EMD / mL. As the EMD concentrations at equilibrium appeared to be the same on analytical SDS PAGE (illustrated by the dashed red boxes), the polystyrene surfaces were assumed to be saturated, with the equilibrium existing between free EMD and the immobilised EMD.

4.2.2.1.2 Characterising the kinetics of EMD binding: Mathematical modelling of the UV absorption curves

It is worth noting that in contrast to what was hypothesised initially (Figure 80, p. 241) the pattern of UV absorbance decrease (Figure 61A p. 192) was not a linear graph with inflexion point. This could be because:

- With time the concentration of free EMD in solution would fall, so binding did not occur under constant conditions and slowed down towards the equilibrium;
- Free EMD- immobilised EMD interactions may be occurring at the same time as free EMD- polystyrene interactions, confounding the protein-protein and protein-polystyrene binding kinetics, as illustrated in Figure 83 below. This precludes seeing a distinct inflexion point.

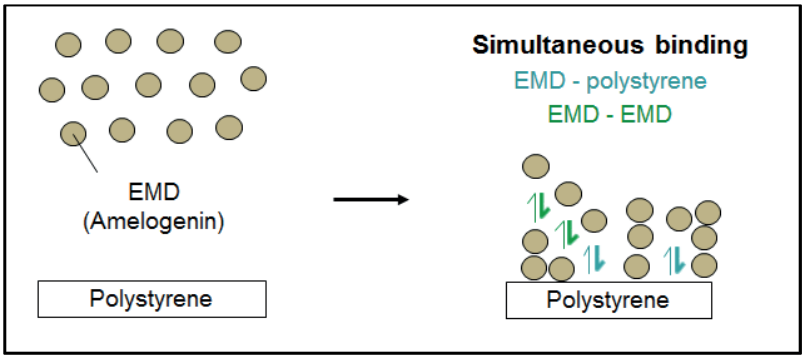


Figure 83 Amelogenin-amelogenin and amelogenin-polystyrene binding may occur simultaneously: hypothesis.

It was important to ensure that the initial phase of EMD-polystyrene binding was not confused with EMD binding to immobilised EMD for kinetic analysis. Passive adsorption by hydrophobic binding to polystyrene may cause denaturation of the proteins adsorbing, which is a well-known issue (Butler et al., 1993). and could alter subsequent amelogenin self-assembly.

To overcome this issue and avoid confounding the kinetics of EMD-EMD binding with those of EMD-polystyrene binding, the binding assay was carried out to include an initial incubation phase that saturated the polystyrene surface with EMD, followed by a second incubation with fresh EMD solution, allowing free EMD to bind to immobilised EMD (See Section 2.2.2.2, pp. 126 – 129 for detailed methodology), presumably representing EMD-EMD binding. The results obtained in Section 3.2.2.2 (Figure 64 p. 198) indicated that free EMD bound to immobilised EMD, with a greater propensity (in terms of rates of depletion from free solution) than that observed for EMD-polystyrene binding (Section 3.2.2.2, pp. 200 - 201).

The depletion kinetics at the 1st and 2nd phases of EMD binding (respectively EMD - polystyrene and EMD – EMD binding) were subjected to curve-fitting and the Hill sigmoid equation generated the best fit, encompassing all time-points (see Figure 65 p. 200) with $R^2 > 0.99$. The Hill equation originally described the saturation of haemoglobin with oxygen as a function of oxygen tension (Hill, 1910). It can be expressed as the following equation (as written in OriginPro 9.1 software (OriginLab, Northampton, MA)):

$$y = Vmax. (x^n/(K^n + x^n))$$

The variable **y** is the velocity of binding reaction and **x** is the concentration of ligand. The equation generates 3 parameters: **Vmax** (= maximum velocity), **K** (= concentration of ligand required to reach half of maximum velocity) and **n** (= cooperation or “Hill” coefficient, whose value indicates whether ligands positively or negatively cooperate while binding to the macromolecule.

The Hill equation is however not actually applicable to the data presented in the thesis, as it expresses the velocity of the reaction as a function of reactant concentration. Besides, the Hill equation is related to various equations in biology describing equilibria: Langmuir isotherms, used in adsorption studies, Michaelis-Menten equation in enzymology or EC_{50} calculations in pharmacology (Goutelle et al., 2008).
In the current work, the exponential functions would have been more relevant biologically to the second phase of depletion if this represented EMD-EMD binding only. As indicated in Figure 20 p. 87 (Section 1.3.6.1), the rate of protein-protein binding is a differential equation. Exponential equations were among the equations tested for curve-fitting (Section 2.2.2.2, p. 128) but did not generate a fit to $R^2 > 0.99$, comprising all time-points. The Hill equation was therefore chosen, because for all graphs obtained it proved to be the most convenient tool for interpretation and comparison. It generated the following parameters:

- Vmax = 'Maximum gain' of EMD to polystyrene surface predicted
- $K=T_{1/2}$ = time to reach half of maximum depletion. This allows to compare .paces of depletion.

4.2.2.2 EMD binding to EMD immobilised on the bottom of the microwells caused the decrease in UV absorbance as EMD was depleted as it bound to immobilised EMD

The absorbance measurements reported in Section 3.2.2.2 (Figure 64, p. 198), including the plateau region, were higher for EMD-EMD binding than EMD-polystyrene binding. At equal equilibrium concentrations (shown by analytical SDS PAGE in Figure 64, p. 198) the difference of absorbance observed (ΔAbs, in Figure 64, p. 198) probably reflected EMD binding to the EMD already immobilised to the bottom of the microwells so that the absorbance recorded was actually an aggregated value of this material and any free EMD left unbound in solution (as illustrated below in Figure 84A below). This was an unavoidable consequence of the UV light path passing through the bottom of the microwell. This was also observed in the initial test to determining sufficient starting EMD concentration to saturate the microwells, where the absorbance at equilibrium was higher at starting EMD concentration of 166 µg/mL (Section 3.2.2.1, p. 192) as illustrated in Figure 84B below.

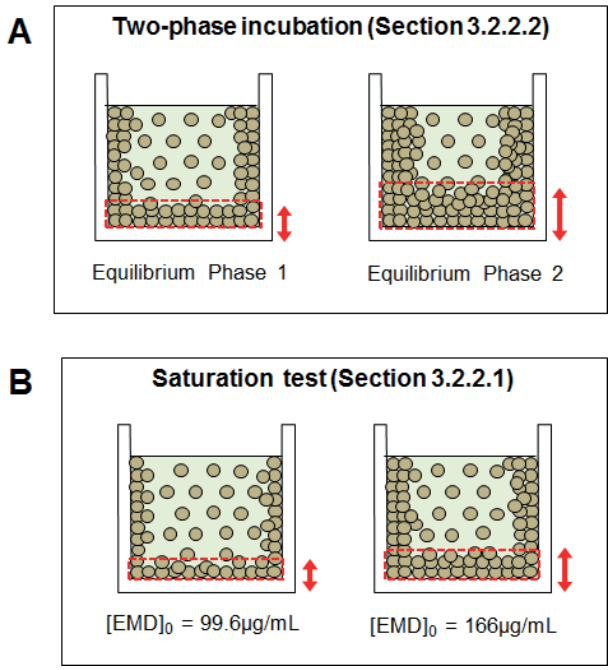


Figure 84 EMD bound to the bottom of the microwells affected the true absorbance reading associated with unbound EMD remining free in solution. Absorbances values were higher at equilibrium when more EMD had bound to the bottom surface of the microwells (highlighted by dashed red boxes and red double-headed arrows). This was observed (A) in the two-phase incubation assay (Section 3.2.2.2, p. 198) and (B) in the saturation test (Section 3.2.2.1, p. 192) where the EMD concentration in solution was the same at equilibrium but the absorbance values differed.

To confirm and further characterise the effect of proteins bound to the bottom of the microwells on the absorbance reading, a third test was carried out (Section 2.2.2.3, p. 130), to measure the absorbance of what remained bound to the bottom of the microwell after discarding the contents of the microwell once the 2nd incubation had reached equilibrium. The final absorbance read, post-incubation, was 0.24 which is not negligible (Section 3.2.2.3 p. 202). So, it is clear that the higher absorbance values at equilibrium (as discussed above, in Figure 84) were due to EMD proteins bound to the bottom of the microwells.

This implies that the interpretation of protein binding kinetics is complicated by proteins binding to the bottom of the microwells (which makes 22% of surface covered by 200 µL EMD solution) causing the depletion of free EMD in solution to be underestimated during the second incubation phase.

However, the results obtained (Section 3.2.2.3, pp. 202 - 203) opened up the possibility that the binding behaviour of WT and mutant p.Y64H r-amelogenins can be compared by measuring the direct accumulation of proteins on the bottom of the microwells, as the value 0.24 obtained with EMD is close to 0.3, which is where absorbance measurements can be read with the highest precision[1]. This method would provide end-point measurements but not direct real-time kinetics. To monitor the kinetics of EMD-EMD binding (shown as EMD depletion from solution) adjustments were needed to eliminate the bias caused by EMD binding to the bottom of the microwells. An option was to block the bottom of the microwell from binding to EMD as discussed in below in Section 4.2.2.3.

4.2.2.3 Blocking did not prevent EMD from binding to polystyrene

Microplate-based immunological assays use different blocking regimes to prevent non-specific adsorption of detection antibodies to the microwell surface. Ideally, antibodies will only be retained in the microwell if they recognise and bind specifically to the target antigen that has been immobilised to the microwell surface. BSA and NFDM are gold standard blockers in ELISA tests (Gibbs, 2001). BSA forms a monolayer when binding to polystyrene surfaces (Fair and Jamieson, 1980) and is predicted to interact with polystyrene via CH_3 groups, resisting washes in PBS (Jeyachandran et al., 2009). NFDM has proved a more effective blocker than BSA in western blotting and in ELISA, as well as being cheaper (Johnson et al., 1984) and like casein (which is its major

[1] To reiterate the interpretation of the results obtained in Section 3.2.2.3 (p. 202), an absorbance value of 0.3 corresponds to the point where 50% of the photons emitted by the light source are absorbed, which allows for reading at the highest precision.

component) was found to be the most effective blocker against peroxidase-conjugated immunoglobulin non-specific binding to polystyrene microwells (Vogt Jr et al., 1987).

In this thesis, the ability of BSA and NFDM to block EMD from binding to polystyrene was tested first, before any attempt to block the bottom of the microwells (See Section 2.2.2.4 pp. 131 – 133 for the method). If the blockers prevented EMD from binding to the polystyrene, the concentration of EMD added to a blocked microwell was expected to remain constant over the incubation period (illustrated in Figure 85A below), causing a constant absorbance reading over time to be obtained (illustrated by the orange dashed line in Figure 85B below). The absorbance pattern would be clearly different from when EMD was added to the "Control" microwells which were not blocked (black line in Figure 85B) .

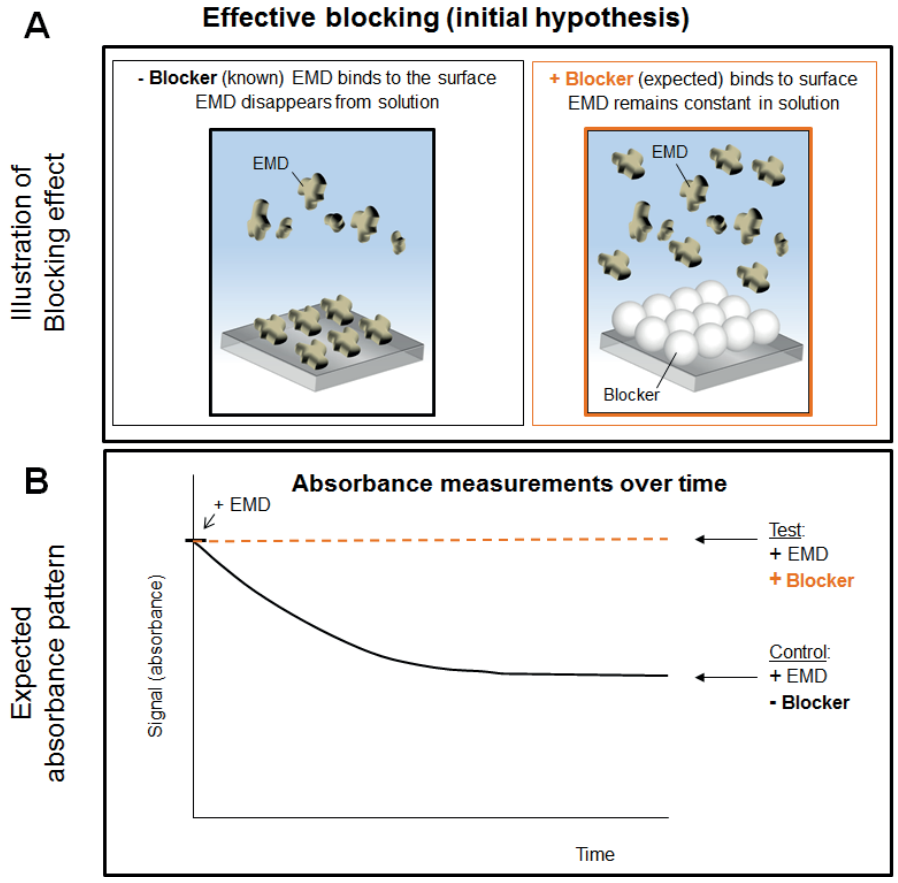


Figure 85 Absorbance pattern expected in case of successful blocking. Panel A illustrates the effect of blocking. (A) Without blocker (black box), EMD proteins bind to the microwell surface, so their concentration in solution decreases. In the presence of blocker (orange box), EMD is expected not to bind to the microwells surface and remain at constant concentration in solution. the absorbance of protein is expected to remain constant over time with blocker (dashed orange line in panel B).

The absorbance patterns obtained (Figure 86A below) however, did not show significant difference in the binding of EMD whether or not the microwells were blocked with BSA or NFDM before adding EMD solution. Three hypotheses to explain this can be drawn, as illustrated in Figure 86 B-D (below):

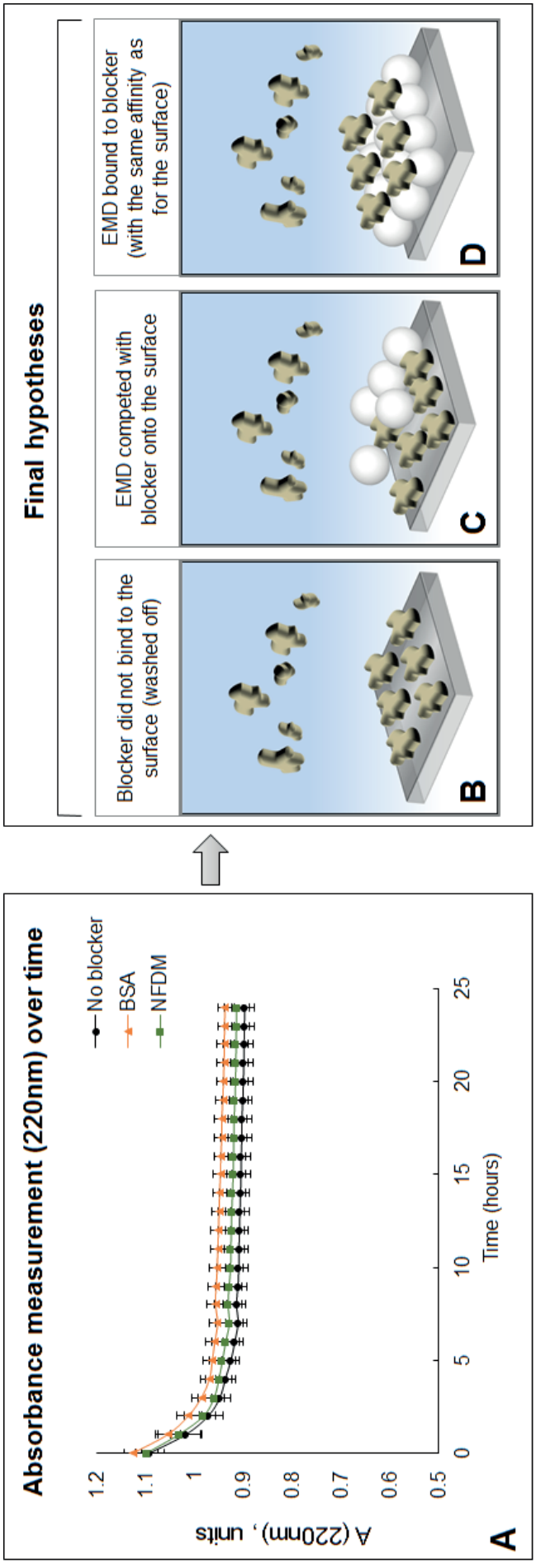


Figure 86 Hypotheses of the why blocking was apparently not effective. The absorbance pattern in panel A, shows that EMD disappeared from solution in nearly identical fashion whether the microwells was non-treated ("No blocker"), treated with BSA or treated with NFDM. This indicates that blocking did not properly occur. As illustrated in the next panels, it could be either: (B) the blockers were washed off the polystyrene surface, (C) EMD competed with the blocker, to bind to the polystyrene, or (D) the EMD bound to BSA or NFDM immobilised on the polystyrene. The hypotheses (B, C, D) are detailed in the text pp. 251 – 252.

- **The gold standard blockers were washed off before adding EMD solution (Figure 86B)**.

The washes were carried out three times simply with PBS (see Figure 34, p.133), but it is unlikely that BSA or NFDM were removed in these conditions, as both are common blockers in ELISA tests and can withstand high stringency washes with PBS even in the presence of 0.05% Tween 20 – a non-ionic detergent commonly added to ELISA wash buffers to minimise loose nonspecific binding . Both BSA and NFDM are "permanent blockers" that "only need to be added once" and can withstand multiple washes (Gibbs, 2001). For example, BSA withstands washes and phosphate-BSA complexes forming reportedly tend to favour BSA-polystyrene binding (Jeyachandran et al., 2009). The next hypothesis is that the blockers were not washed off during washes with PBS, but were removed by EMD acting as a binding competitor.

- **EMD had a stronger affinity than the blockers for the polystyrene (Figure 86C)**

Assuming that amelogenin may have a greater affinity for polystyrene surfaces than BSA or NFDM, then, its addition could quickly displace the adsorbed blocker by the "Vroman" effect. The Vroman effect originally described the displacement of fibrinogen initially adsorbed to a surface caused by plasma proteins having a greater affinity for that surface (Vroman et al., 1980) but is the general definition of the competitive desorption of one protein by another (Hirsh et al., 2013). It has been reported using QCM that BSA previously adsorbed to a surface was displaced by competitive binding of other proteins such as fibronectin (Felgueiras et al., 2016) Similarly, SPR studies suggested that albumin was displaced by competitive binding of fibrinogen and IgG (Green et al., 1999). As for NFDM, the Vroman effect has not been documented. This does not mean that NFDM cannot be displaced by competitive adsorption of EMD or other proteins. This hypothesis could be verified using AFM, QCM-D, Time-of-Flight Secondary Ion Mass Spectrometry, Time-of-flight mass spectrometry or SPR, which were employed to characterise competitive protein displacement on surfaces (Green et al., 1999, Hirsh et al., 2013). QCM is suitable technique in particular to characterise protein surface adsorption and blocking abilities (Reimhult et al., 2008). Alternatively, it would be a simple matter to run the contents of a blocked microwell following the addition of amelogenin on SDS PAGE to see if the blocking protein had been desorbed into solution.

Should the Vroman effect occur, one could expect that the kinetics of absorbance depletion would differ between conditions as the EMD would have to displace the blocker first before it could bind the surface itself (unless this process occurred very

quickly). How, then, can the fact that the absorbance patterns are similar be explained (Figure 86A, above)?
It maybe that EMD can displace the blocking protein very quickly; The time-point measurements were every 30 minutes, which may not provide a sufficient temporal resolution if EMD displaced the blocker very quickly.

- **EMD bound to BSA or NFDM with the same affinity as it bound to polystyrene (Figure 86D).**

BSA and NFDM are the most common blockers used in the laboratory as they prevent non-specific protein binding in ELISA studies (Gibbs, 2001) or western blotting. Naturally, they are only suitable providing the detection antibodies do not bind to them instead of the microwell surface. For BSA and NFDM, there is a risk of cross-reactivity with anti-phosphotyrosine antibodies while using BSA prepared from fraction V (which is the BSA used in this thesis, see Section 2.2.2.4, p. 131) as it may contain phosphotyrosine (Gibbs, 2001). In addition, in the case of BSA, some cross-reactivity was reported to human antibodies (Chart et al., 1998) or to vaccina virus complement control protein (Xiao and Isaacs, 2012). In contrast, the literature abounds with thousands of examples where BSA and NFDM have been used successfully as blockers suggesting that in general proteins do not interact with them. However, this is not to say EMD proteins follow this trend.
As blocking did not prove effective it could not be used to solve the issue around the undesirable binding of EMD to the bottom of the microwells and further development is required. At this stage then, the microplate binding assay can be used to provide an end-point measurement of EMD binding to the bottom of the microwells by using UV spectroscopy to directly measure the amount of proteins adsorbed to the bottom of the microwells after 24-hour incubation.

4.2.2.4 Achievements, future developments and prospectives

4.2.2.4.1 UV-transparent microplate-based depletion measurements provide a cost-effective and simple method to analyse amelogenin binding behaviour

To recap, the microplate binding assays developed and discussed previously in Section 4.2.2 provided reliable end-point measurements as a means to compare binding behaviours of WT and p.Y64H mutant r-amelogenins. This approach consisted of incubating proteins in polystyrene in microwells until saturation and protein-protein binding equilibriums are reached, and measuring the absorbance (after discarding the solutions) of what is left on the bottom of the microwell. Comparing the solution contents before and after incubation also confirmed protein depletion.
This assay provides an exciting prospective, as it allows high-throughput studies, the protocol is easy to carry out and is adaptable in that the solution conditions are easily modified. The use of 384 microwell microplates would further reduce the amounts of precious r-amelogenin even further as the microwell volume is only 131 µL compared to the 392 µL in a microwell from a 96 microwell plate.

4.2.2.4.2 Impact of the proposed method and future use

As it stands, the method developed here could also be used as a preliminary screening tool to determine conditions (buffer, pH, range of concentrations, temperature etc.) for subsequent binding studies using SPR, QCM and AFM.
As mentioned in Section 1.3.4 (p. 64) *in vitro* studies are useful adjuncts to *in cellulo or in vivo* studies, as they allow detailed understanding of biological mechanisms. On their own, *in vitro* studies rarely reflect *in vivo* events. Notably, common *in vitro* protein binding experiments may involve protein labelling and may be limited by protein solubility. Due to amelogenin's propensity to aggregate in physiological-like buffers (Tan et al., 1998), it has proved problematic to study amelogenin binding behaviour or assembly (see Sections 1.2.2.2.4 pp. 25 - 26 on amelogenin supramolecular assembly, and 1.3.3.2.1 p. 59 regarding the studies of p.P70T mutation).
Therefore, in techniques such as SPR (further details are provided in appendix E), which require that amelogenin should be solubilised (according to Biacore handbook (Biacore Assay Handbook 29-0194-00 Edition AA)), cost-effective pilot studies could be helpful to determine conditions that mimic as closely as possible those existing *in vivo* or *in cellulo* (eg. amelogenin concentrations, buffer, pH) while still being compatible with SPR. The microplate binding assay developed in this thesis (discussed throughout Section 4.2.2, pp. 241 - 247) can then be integrated into the pipeline below (Figure 87): following *in vivo* observations, it could provide a step that

would guide the choice for conditions to use in more elaborate binding studies using SPR or QCM. This would help dissecting *in vivo* events such as mechanisms underlying AI with greater accuracy, which itself would help characterising therapeutic targets.

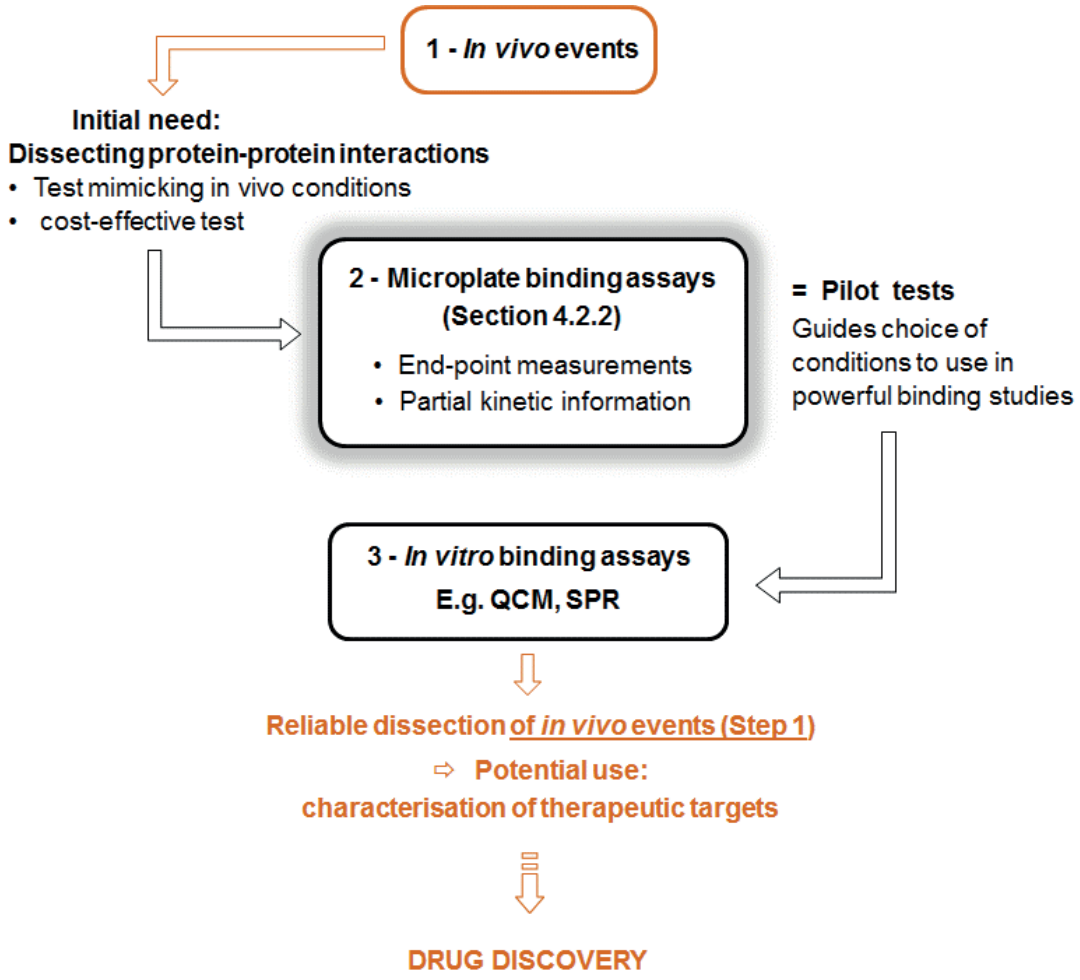


Figure 87 Experimental pipeline for dissecting protein binding behaviours *in* vivo ("*in vivo* events. Following initial *in vivo* observations (step 1) the microplate binding assay developed (Section 4.2.2) can provide pilot data (step 2) to guide the design of protein-protein interaction studies (step 3). Accurate characterisation of *in vivo* events is a first step for drug discovery. The interest of this pipeline is discussed in text, p. 253-254 regarding amelogenin-amelogenin binding studies (short term prospective) and pp. 263-264 regarding amelogenin binding to other proteins (middle-long term prospective).

However, the microplate binding test described here is open to further improvements that would allow it to be used to generate more exhaustive and accurate binding kinetics data such that it could replace the more expensive and lower throughput techniques such as SPR and QCM. The improvements to consider are discussed in the next Section (4.2.2.4.3).

4.2.2.4.3 Immediate possibilities for improvement of the microplate-based binding assay

The following improvements should be considered, to generate more accurate and reliable binding kinetics data using the microplate assay.

- **Measure accurately the initial rate of binding reaction**

The absorbance measurement recordings reported in Section 3.2.2 (Figures 61A, 64, 66, 68A, 69A) started within 4.5 to 10 minutes after preparing EMD solutions, which should be the actual starting point of the experiments. In this thesis, this was considered as negligible as it only represent 0.3 – 0.7% of the total duration of the absorbance measurements. Nonetheless further development would be needed to measure the initial reaction rates. In the short term, using QCM or SPR in complement could help to measure the initial binding reaction rates.

- **Prevent any amelogenin from binding to the bottom surface of the microwells and masking the solution depletion kinetics.**

To reiterate Section 4.2.2.2 (p. 247), all surfaces binding and immobilising EMD and r-amelogenin solutions were polystyrene, the bottom surface comprises 22% of the total surface exposed to the protein solutions which is not negligible. The rate at which proteins adsorb to the bottom surface of the microwell masked the rate at which they disappeared from solution due to binding to the rest of the microwell surface which is problematic. Initially, the approach taken was to block the bottom of the microwells with gold standard blockers BSA and NFDM, but these were not effective (Discussed in Section 4.2.2.3, pp. 247 - 252). Alternative 'anti-fouling' methods remain to be tested and investigated. Polystyrene surfaces can be transformed to prevent protein binding. For example, PLL-g-PEG adsorbed on to polystyrene prevented BSA adsorption to the surface (Hecker et al., 2018) and there are many other potential blockers that could be tested. Another approach to consider is chemically transforming the surfaces of the microwell. For example, treating the surface to make it hydrophilic by plasma irradiation or by corona discharge (Onyiriuka et al., 1991). This may inhibit amelogenin binding if the binding is hydrophobic in nature.

- **Test different concentrations of amelogenins to generate data relevant to the Hill equation.**

The Hill equation was chosen in EMD adsorption curve-fitting (see Figure 65, p. 200) as a function of time (as Discussed in Section 4.2.2.1, p. 244). However the Hill equation is (in simplified terms) the expression of a reaction velocity as a function of the initial reagent concentration. To be able to use the Hill equation model (with biological relevance), the next step would then consist of testing different starting concentrations of amelogenin/EMD and plot them against initial depletion rates.

4.2.2.4.4 General risks to consider with microplate-based studies: amelogenin is an aggregative protein

Amelogenin is prone to self-assembly and aggregation. It is also an intrinsically disordered protein . Although its best-known assembly mode is to form nanospheres, an ellipsometry and AFM study showed that when adsorbing to surfaces, it formed smaller structures, either monomers or reduced oligomers, and it may adopt various quaternary structures depending on the adsorption surface used (Tarasevich et al., 2009a, Tarasevich et al., 2009b). This effect may be an issue for any technique where amelogenin is adsorbed .

4.2.2.4.5 Relevance of EMD as a surrogate

The convenience of using EMD as a surrogate for mouse r-amelogenin is that it comprises mostly amelogenins and is available in abundant quantities for method development purposes (Sections 3.2.2.1 – 3.2.2.4 in Results chapter, and 4.2.2.1 – 4.2.2.4 in Discussion chapter) (Maycock et al., 2002). EMD contains a mixture of porcine amelogenins (Maycock et al., 2002) which share a high degree of homology with mouse amelogenins (Brookes et al., 1995). However, EMD is a heterogeneous mixture of amelogenins, comprising all of the amelogenin processing products, together with non-amelogenin components such as ameloblastin (Kuramitsu-Fujimoto et al., 2015) found in trace amounts in the pig secretory stage enamel used to produce EMD. Clearly, EMD is not comprised of 100% full length amelogenin, the amelogenin isotype comprising the r-amelogenins used here and therefore the following caveats need to be taken into account:

· **Caveat 1**: EMD It is a heterogeneous mixture. It comprises proteins that may exhibit a wide range of different binding affinities, giving absorbance data that reflect an average of all proteins rather than one specific species, unlike the r-amelogenins used in this thesis. This complicates the interpretation of protein binding kinetics.

· **Caveat 2**: Its major component is the 20 kDa “P148” amelogenin processing product, which raises 2 concerns:

(1) P148 does not contain the hydrophilic telopeptide. Studies comparing rP147 and the full length telopeptide-bearing rP172 showed that the telopeptide played a role in amelogenin assembly and amelogenin-mineral interactions (Wiedemann-Bidlack et al., 2007, Kwak et al., 2009). The ideal surrogate would have been full length porcine amelogenin P173 but this is not the most abundant isoform in EMD.

(2) P148 is found in the enamel matrix only since it is a proteolytic processing product of P173. As explained in the Sections 1.2.1.2.1 (p. 17) and 1.2.3.1.1 (pp. 30 – 31), P173 is the major isoform of parent amelogenin, which transits through the

ameloblast secretory pathway and towards secretion; given the conservation of amelogenin between species (Brookes et al., 1995) P173 would be the most appropriate surrogate for M180 and therefore, for the r-amelogenins used in this thesis.

In spite of the caveats mentioned above using EMD as a surrogate proved useful in these preliminary studies. It demonstrated that the method could be used to track the depletion of a relevant protein from solution as it bound to proteins adsorbed to the microplate microwell walls (discussed above in Section 4.2.2) and using microplate-based assay to compare the binding behaviours of WT and p.Y64H mutant r-amelogenins where feasible (discussed in the Section 4.2.3). Specifically, using EMD as surrogate informed on:

- Experimental design.
- Suitable starting concentrations (100 -160 µg/mL) which saturated the polystyrene surfaces and subsequently provided a sufficient negative change of absorbance that could be accurately tracked as the proteins were depleted from solution by interaction with the immobilised protein phase.
- The incubation time required to reach binding equilibrium (for saturating the microwell surfaces in order to establish the immobilised phase and for the subsequent binding of freshly added protein this immobilised phase; around 10 -15 hours).

4.2.3 Effect of the p.Y64H mutation on mouse amelogenin binding properties and the possibilities of dissecting molecular mechanisms underlying AI.

Using the experimental design developed with EMD as a surrogate (discussed above in Section 4.2.2), the binding behaviours of WT and mutant p.Y64H r-amelogenins were compared. This Section discusses the results obtained.

4.2.3.1 The hypothesis that the p.Y64H mutation causes amelogenin to become aggregative is supported by these data

The possibility that the p.Y64H amelogenin increases amelogenin-amelogenin binding propensity is supported by the results obtained in Section 3.2.2.5 (pp. 207 - 212):

- The absorbance measurements and analytical SDS PAGE gel densitometry analyses (Figure 69 p. 209) showed that WT^{-His} remained more or less in solution at constant concentration (excepting for a 8-10% minor decrease of signal). Mut^{-His} was depleted from solution by 45% over 24 hours incubation.
- Reading the absorbance of the proteins adsorbed onto the bottom of the microwells left after discarding the solutions (Figure 70 p. 210), the recombinant Mut^{-His} yielded an absorbance value of 0.29. In contrast, the absorbance values read after discarding WT^{-His} microwells were 5-7-fold lower (0.04) with a much smaller standard deviation which confirms that in contrast to Mut^{-His}, WT^{-His} was hardly bound to polystyrene. Mut^{-His} r-amelogenin appeared better at binding to the microwells than WT^{-His}, which did not show any obvious aggregation behaviour as it stayed in solution over the incubation period. This suggests that the p.Y64H mutation increased r-amelogenin apparent hydrophobicity and propensity to aggregate. This pilot experiment therefore provided with a further step in understanding the effect of the p.Y64H mutation, as discussed below in Section 4.2.3.2. However, the difference of binding behaviour was not observed with the His-tagged r-amelogenins, where WT^{+His} showed higher aggregation propensity than Mut^{+His}. It is clear that the presence of the His-tag appeared to make a huge difference on the behaviour of r-amelogenin in this experimental system. This point is addressed in greater detail later in Section 4.3.2 (p. 268).

It is worth noting that the r-amelogenins employed were expressed in *E. coli*, [which differs from the eukaryotic system of expression] and therefore lacked phosphorylation on Ser^{16} (See Section 4.1.3.2, p. 235). Although both WT and mutant p.Y64H r-amelogenins lacked this phosphorylation, this may have compromised the comparison between WT and mutant p.Y64H r-amelogenin. The effect of phosphorylation has

been itself a subject of great interest (see Section 1.2.1.3.1, p. 18 for detailed explanations).

4.2.3.2 Hypothesis regarding the effect of the mutation p.Y64H

As suggested in the Section 1.3.4 (pp. 64 - 65) understanding the effect of the mutation p.Y64H on amelogenin function, in the disruption of amelogenesis, requires a combination of *in vitro*, *in cellulo* and *in vivo* studies; to reiterate, in brief:

- *In vitro* and (if applicable) *in silico* data would dissect the mechanism underlying AI in fine details.
- *In cellulo* studies are also needed to dissect the underlying mechanisms of AI, while providing a broader picture.
- *In vivo* studies are needed as initial observations and also, at some point, to validate the data obtained *in vitro* and/or *in cellulo*.

Built on the literature review and on the pilot data discussed above in Section 4.2.3.1, there are various possible consequences of the p.Y64H mutation to consider, at different levels. The hypotheses are illustrated in the diagram below, adapted from Figure 13 (p. 65). The relevant sections of the literature review and thesis data are displayed on the Figure below.

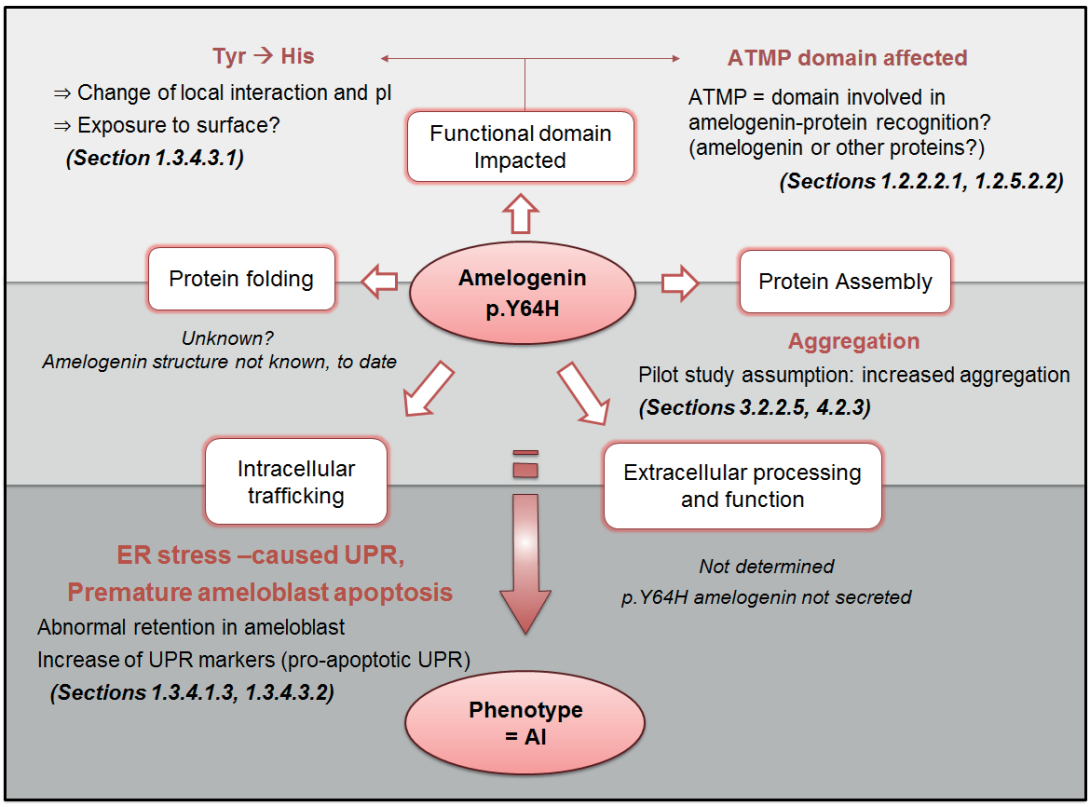


Figure 88 Hypothetical or identified effect(s) of the amelogenin p.Y64H mutation. This figure represents the scales which may be impacted by the mutation p.Y64H, i.e.: locally, in the amino acid sequence (box "functional domain impacted") at the amelogenin structure and function scale (boxes "folding" and "assembly") or at the cellular or extracellular scales (boxes"intracellular" and "extracellular trafficking"). Futher description can be found in the text below.

As shown in Figure 88 above, the p.Y64H mutation in amelogenin may have effects at different scales:

- **Locally, as a single point mutation**.

Tyrosine and Histidine have different propensities for exposure to aqueous surface and as well, can participate to different types of interactions (as reviewed in Section 1.3.4.3.1, pp. 79 - 82). At higher scales this single point mutation *per se* can hypothetically affect the conformation and binding properties of amelogenin.

- **Locally, affecting the ATMP domain**

As ATMP domain is reportedly involved in amelogenin interaction with itself and with other proteins (reviewed in Sections 1.2.2.2.1, p. 22 and 1.2.5.2.2, pp. 47 - 49), the mutation may affect amelogenin recognitions properties and interaction with itself and/or other proteins, an avenue to consider.

The two “local” effects, cited above, can be described *in silico* as they only concern a few amino acids, as such. However, *in silico* studies are limited at this level since amelogenin 3-dimensional structure has not been solved, to date (see Section 1.2.1.3.3, p. 20). Awaiting structural data, the effect of the p.Y64H mutation was studied *in vitro*, focusing on amelogenin binding properties.

- **Affecting amelogenin binding properties *in vitro***

The pilot data obtained with the microplate-based binding assay developed in this thesis (Sections 3.2.2.5, 4.2.3) suggested that the p.Y64H mutation increased amelogenin’s propensity to aggregate. This appears supported previously reported data, which found abnormal retention of mutant p.Y64H amelogenin in the ER, as detailed in the next paragraph.

- **Affecting amelogenin intracellular trafficking *in cellulo* and *in vivo*.**

As detailed in Section 1.3.4.1.3 (p. 70), ameloblasts expressing p.Y64H amelogenin displayed abnormally engorged ER vesicles, with mutant amelogenin, which strongly suggested ER stress (Brookes et al., 2014). Consistently, an increase of UPR markers was observed, including pro-apoptotic markers, which would explain the premature ameloblast apoptosis noticed *in vivo* by Brookes et al (2014). Abnormal retention of p.Y64H amelogenin was also observed *in cellulo* in COS-7 cells while co-transfected with ameloblastin (see Section 1.3.4.3.2, p. 82). These latter data are capital to validate any *in vitro* or *in silico* data, past, present or future.

To compare the binding behaviours of WT^{-His} and Mut^{-His} r-amelogenins with more *in vivo* relevance, the next step would be to obtain recombinants that carry phosphorylation on Ser^{16}. This can now be achieved since, as developed in this thesis, amelogenin can be purified using acetic acid extraction and preparative SDS PAGE (see Section 4.1.3.2, p. 235). Instead of *E. coli*, the expression system would need to be eukaryotic, which would include the relevant protein folding machinery operating in the ER during the secretory pathway (and overcome the problem of bacterial cleavage of the N-terminal methionine). Following purification however, a challenge would still potentially in whether the r-amelogenins can recover their native conformation.
A caveat in many *in vitro* experiments is that they rely on correct protein folding and conformation is achieved via the amino acid sequence alone- however, this is not always enough to fold a protein – as chaperones are often necessary to catalyse protein folding to their native functional conformations (Hartl et al., 2011).

4.2.3.3 Future tests to develop a therapeutic for AI as a conformational disease

Clearly, data obtained using the WT^{-His} and Mut^{-His} r-amelogenins suggested that their binding behaviour was affected by the p.Y64H mutation. The current results (Section 3.2.2.5, pp. 207 - 212, discussed in Section 4.2.3.1, p. 258) supported that r-amelogenin was more aggregative while carrying the mutation p.Y64H, which is consistent with its abnormal retention observed in the ameloblasts and the subsequent ER-stress and pro-apoptotic UPR (Brookes et al., 2014). However, the results obtained in this thesis (Section 3.2.2.5) were generated by one pilot experiment due to time constraints, so further work including the use of other techniques would be advisable to corroborate the data obtained and confirm their reproducibility. Characterising the way in which the p.Y64H mutation affects protein-protein interactions will be invaluable in helping to dissect ER-stress and UPR mechanisms driving AI in the presence of this mutation and help develop therapeutic strategies and targets.

This section proposes middle- and long-term strategies to study the pathogenesis of AI and other enamel diseases:

- In the middle-term, studying the effect of p.Y64H mutation on how amelogenin interacts with other proteins (e.g. chaperones and UPR activation trans-ER-membrane sensors) involved in the ameloblast secretory pathway (Discussed in Section 4.2.3.3.1 overleaf).
- In the long-term, the UPR has been associated with enamel fluorosis (Brookes et al., 2017a), so the scope of ER-stress and UPR- studies can be widened to understand and possibly treat other enamel diseases (Discussed in Section 4.2.3.3.2 overleaf).

4.2.3.3.1 Analysing impaired protein-protein interactions in the presence of the p.Y64H mutation

Section 3.2.2.5 showed that the effect of the p. Y64H mutation significantly affected amelogenin aggregation. Various binding partners for amelogenin, other than amelogenin itself, were previously identified *in vitro* (as reviewed in Section 1.2.5.2.1, pp. 45 - 47); for example ERAF components such as BiP/Grp78, calreticulin, protein disulphide isomerase precursor, apasin-ERP57 (Fukuda et al., 2013), cytoskeleton components such as actin, vimentin, tubulin (Fukuda et al., 2013), cytokeratins 5 and 14 (Ravindranath et al., 2001, Ravindranath et al., 2003), or other EMPs such as ameloblastin (Ravindranath et al., 2004), with which it colocalises in the intracellular secretory pathway (Mazumder et al., 2014). As suggested in Section 1.3.4.3.2 (pp. 82

- 83), the p.Y64H mutation could affect the binding equilibria related to amelogenin self-assembly, amelogenin-ameloblastin interactions, or amelogenin interactions with ER folding machinery. These interactions may be interdependent and necessary to the proper trafficking of amelogenin in the ameloblast and to its extracellular function (Figure 19, p. 83).

Therefore, in the middle-term, the effect of the p.Y64H mutation on the binding of amelogenin with other relevant proteins is worth investigating. This could be done using the microplate-based method developed in this thesis (discussed in Section 4.2.2), possibly in combination with other techniques. The pipeline proposed in Figure 87 (p. 254) and possible improvements to the microplate binding assay (discussed in the Section 4.2.2.4.3 pp. 255) would provide an efficient strategy to characterise the kinetics of protein-protein interactions involved.

The choice of amelogenin-binding partners to study can be guided by data from the literature, and ideally, the availability of an “interactome” database would be of great help (Shoemaker and Panchenko, 2007). The online “Search Tool for the Retrieval of Interacting Genes/Proteins (STRING)” (https://string-db.org) is a possible starting point. However, little information is currently available regarding the amelogenin interactome.

This research prospective mentioned, the studies on protein-protein interactions involved (and impaired) in the context of ER-stress and UPR (see Section 1.3.4.1.3 p. 70 on the mouse model carrying the mutation p.Y64H) could serve a wider investigation in the long-term, as UPR has been reported in other enamel pathologies (addressed below in Section 4.2.3.3.2);

4.2.3.3.2 Conformational diseases and therapy in amelogenesis

- **ER stress and UPR found in cases of AI**

The mouse model expressing the amelogenin p.Y64H mutation is the first model showing ER stress as mechanism driving AI (Barron et al., 2010, Brookes et al., 2014). A recent paper by Brookes et al, 2017, addressed the role of the UPR in enamel pathologies (Brookes et al., 2017a): AI and fluorosis. Following the publication on amelogenin p.Y64H , ameloblast ER stress was observed as well with enamelin mutations p.55I in mouse and p.L31R in human (Brookes et al., 2017b). The secretory stage ameloblasts secretion pathways were impaired as enamel matrix proteins were abnormally retained. Notably, UPR markers (e.g., Grp78, Xbp1, Grp94, and Atf4) were upregulated in ameloblasts expressing amelogenin p.Y64H and Enamelin p.S55I.

- **ER stress and UPR found in fluorosis**

The mechanisms underlying fluorosis have been studied for decades. Histological studies reported that fluoride disturbed intracellular protein trafficking (Matsuo et al., 1996). This supported the idea that the ameloblasts were suffering ER stress. Supporting this finding, a study using the ameloblast-derived cell line LS8 and porcine enamel organ epithelial cells showed that UPR components were up-regulated in response to fluoride. The researchers observed that fluoride promoted IRE1 activation, with up-regulation of BiP (Grp78), Xbp-1, and Chop expression (Kubota et al., 2005).

- **Therapeutic prospective**

Identifying the UPR as an aetiological factor in AI or fluorosis offered prospective for developing therapeutic interventions that could potentially abolish AI in affected familes or fluorosis. In the mouse model (X-linked AI) carrying the mutation amelogenin p.Y64H, treatment with 4-phenylbutyrate restored the phenotype in heterozygous mice for the mutation (Brookes et al., 2014). 4-Phenylbutyrate also protected the ameloblast-like cell line ALC from cytotoxicity caused by fluoride, again by impacting on elements of the UPR (Suzuki et al., 2017). The precise mechanism of action of 4-phenylburtyrate is not entirely known to date, although both reports showed that it mitigates UPR-pro-apoptotic outcome.

Conformational diseases in general are currently subject to intense research. The possibility of modulating intracellular protein-protein interactions with a view to reversing aggregation-linked ER stress is one major approach in attempts to identify a therapeutic intervention. As well as acting as an anti-apoptotic agent, 4-phenylbutyrate is also reported to act as a chemical chaperone and has been shown to restore a stalled secretory pathway in cells expressing the mutated sodium transporter responsible for cystic fibrosis (Rubenstein et al., 1997) or reduced the aggregation of mutant myocilin in a case of glaucoma (Yam et al., 2007). It had, though, no such effect on ameloblasts expressing p.Y64H amelogenin. Otherwise, BiP/Grp78 expression enhancer (BIX) has also been shown to restore correct protein folding and so potentially alleviate ER stress (Hetz et al., 2013). In the search to restore correct protein folding, a comprehensive characterisation of the effect of candidate drugs on protein-protein interactions is required. This will facilitate the design of therapeutics that can control pathological aggregation.

4.3 The presence of His-tag altered the binding/aggregation behaviour of amelogenin

In Section 3.2.2.5 (pp. 207 - 212), WT^{+His} and Mut^{+His} r-amelogenins clearly exhibited different binding behaviours when compared to their His-tag-free homologues. This section addresses (i) the predicted chemical properties of the His-tag (Section 4.3.1), then (ii) highlights the issues associated with using His-tagged amelogenin in binding assay reported in this thesis (Section 4.3.2) finally (iii), this section reviews the role of His-tag in purification methodologies and in protein binding studies (section 4.3.3).

4.3.1 Predicted and measured properties of fusion-His-tag

Histidine has well-known unique binding properties (Section 1.3.4.3.1, pp. 79 – 82) as its imidazole ring mediates various types of interactions including cation metal coordination via its electron-donating secondary amine. In immobilised metal affinity chromatography (in this thesis, a nickel column was used) the number of His residues comprising the His-tag affects the purification efficiency; as longer His-tags improve the protein purity (Hochuli et al., 1988). However, increasing the length of the His-tag increases the risk of altering the protein structure and function. Balancing both these factors (final protein purity and maintaining protein function), hexahistidine-tags are the preferred option (Bornhorst and Falke, 2000) and are very commonly used. Using hexahistidine tags, the binding affinity of the tag for immobilized Ni^{2+}-nitrilotriacetic acid ranges from K_D = 1-10 μM for fusion-tag proteins (Nieba et al., 1997, Lata et al., 2005) to K_D = 14 nM for oligopeptides (Knecht et al., 2009). Hexahistidine-tagged proteins can be eluted by competition with imidazole at 100 mM or higher concentrations (Bornhorst and Falke, 2000).

The His-tag employed in this thesis is 40 amino acids long (see sequence below) and includes a cleavage site for rhinovirus 3C protease (underlined) :
“MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSLEVLFQ#”
Using the on line tool ProtParam (https://web.expasy.org/protparam/), this His-tag has calculated a pI of 9.46 and a grand average of hydropathicity (GRAVY) index of -0.603, which suggests that it is hydrophilic (Kyte and Doolittle, 1982, Gasteiger et al., 2005). As summarised in Table 14 below, the inclusion of this His-tag in the r-amelogenin may not have significantly affected the Kyte and Doolittle GRAVY index, but would have increased the pI from 6.51 to 6.87; which is significantly closer to physiological pH. So, as the binding assays used PBS: 1% acetic acid at ratios 216.4:1 to 290:1, whose pH is within physiological values, which is closer to the His-tagged r-amelogenins’ pI; His-tagged r-amelogenins would be expected to be less

soluble and more aggregative than His-tag-free amelogenins. This may explain why both WT^{+His} and Mut^{+His} r-amelogenins used in this thesis were more aggregative, or were better able to bind polystyrene microwell surfaces, than WT^{-His} r-amelogenin . Crucially however, in contrast to WT^{-His} r-amelogenin, Mut^{-His} r-amelogenin appeared to exhibit significant aggregative or polystyrene binding properties. Whatever the case, the mutation was having a significant impact on r-amelogenin behaviour.

Table 14 Differences of pI and hydropathicity predicted by Protparam between WT^{-His} [1] and WT^{+His} [2] mouse amelogenins (Gasteiger et al., 2005): effect of the addition of His-tag.

	- His (M180)	+ His
Theoretical pI	6.51	6.87
Grand average of hydropathicity (GRAVY)	-0.696	-0.681

[1] The pI and GRAVY indexes were calculated from the sequence of His-tag-free mouse M180 amelogenin WT (WT^{-His}). The 'input' sequence is provided in Appendix A.

[2] The pI and GRAVY indexes were calculated from the sequence of recombinant His-tagged mouse M180 amelogenin WT (WT^{+His}) used in this thesis. The 'input' sequence is provided in Table 3 p. 92 (Section 2.1.1.1).

4.3.2 His-tag altered r-amelogenin function, masking the effect of mutation p.Y64H.

As mentioned earlier in Section 4.2.3.1 (p. 258), the presence of the His-tag significantly enhanced apparent binding of r-WT r-amelogenin to the plastic. The gain of proteins by the polystyrene surface was greatly enhanced and accelerated when His-tagged r-amelogenins were used. The models generated by curve-fitting with the Hill equation (Figure 71, p. 212) predicted that the depletion from solution of both WT^{+His} and Mut^{+His} r-amelogenins was 50% higher and faster that the depletion of WT^{-His} r-amelogenin , with a 'half-time' of 15 minutes (-vs- 40 minute predicted for Mut^{-His}). (To reiterate, these kinetics parameters were only a rough estimate and do not permit accurate comparison (Discussed in Section 4.2.2.1.2, pp. 244 - 245)).

If EMD binding to polystyrene is hydrophobic then the inclusion of the His-tag may have enhanced hydrophobic interaction; as shown in Table 14 (Section 4.3.1, p.267), the GRAVY index predicted on Protparam tool is less negative, which suggests an increase in hydrophobicity (Kyte and Doolittle, 1982). As mentioned previously in the Section 4.3.1, the change of pI (see Table 14), which is closer to the reaction pH, may play an important role.

Beyond these parameters (GRAVY, pI, etc.), which are predicted simply from amino acid sequences, the enhanced aggregative behaviour seen in the presence of the His-tag may be caused by the His-tag rendering the hydrophobic parts of the r-amelogenin more accessible to the surface. In some way this would alter the conformation of r-amelogenin, making it more aggregative. This is important in studies that use recombinants with the His tag still attached, which may possibly mislead the data reported previously in binding studies; this is addressed in the next Section.

4.3.3 r-amelogenin is often used in functionality studies with the His-tag still in place

The use of His-tagged proteins in functionality studies is widely practiced under the assumption that the tag has no effect on protein conformation or function (Chant et al., 2005). As His-tag affinity purification is a relatively simple and well-established method (Hochuli et al., 1988) the addition of a His-tag has been extensively used to purify proteins for crystallographic studies, for which the tag was not removed (Derewenda, 2004). There are several examples where His-tagged enamel proteins have been used in functionality studies, e.g. His-tagged amelogenin, ameloblastin and enamelin were used to investigate interaction of these enamel proteins with fibronectin (Beyeler et al., 2010) and His-tagged ameloblastin was used to investigate the growth factor-like activity of ameloblastin and its effect on cell attachment and proliferation of periodontal ligament cells (Zeichner‐David et al., 2006). A panel of *in vitro* studies of amelogenin self-assembly and mineral binding, used r-amelogenins carrying an N-terminal His-tag: “RGSHHHHHHGS”. The addition of this His-tag sequence, which is not cleavable, was initially for purification purposes (Moradian-Oldak et al., 2000, Tarasevich et al., 2007, Lakshminarayanan et al., 2010, Tao et al., 2015). Some reports estimated that the presence of His-tag in *in vitro* studies did not alter amelogenin nanosphere formation nor its calcium phosphate nucleation behaviour (Moradian-Oldak et al., 2000, Tarasevich et al., 2007), while other studies reported that the presence of the same His-tag (RGSHHHHHHGS) affected amelogenin assembly and adsorption behaviour to hydroxyapatite (Tao et al., 2015). Tao et al (2015) proposed that at pH 8, the His-tag promoted oligomer–oligomer interactions by lowering the net protein charge and promoting imidazole ring-stacking interactions. In this thesis, the His-tag clearly affected the binding of r-amelogenin to polystyrene surfaces at physiological pH and temperature. The mechanisms involved are unclear, though one may propose that the His-tag imparted some specific conformation on the r-amelogenin, which altered the availability of specific residues or domains to interact with the polystyrene.

Clearly, based on the pilot data presented in this thesis and the literature, it would be preferable to conduct functionality studies with r-amelogenins free of any tags in order to maximise confidence in the data obtained. The use of His-tags for functionality and structural studies does not guarantee that the results obtained will necessarily reflect *in vivo* situations. For example, a His-tag proved to be problematic when studying Gcn5-related *N*-acetyltransferase as the His-tag bound to the substrate-binding site and affected the enzyme’s active site conformation and activity (Majorek et al., 2014).

That the His-tag affected the behaviour of amelogenin-amelogenin interactions should be considered in the design of future functionality studies. For example, when carrying out SPR studies, sensor chips are available whose surfaces are functionalised with Ni^{2+} so that His-tagged protein can be captured prior to running a binding experiment (Biacore Series S NTA Sensor Chips). The resulting data would need to be interpreted in the knowledge that the His-tag could influence the results obtained.

Chapter 5 Conclusions and future work

To recap, the work achieved in this thesis responds to the initial aims as follows:

5.1 Aim 1: Production and purification of r-amelogenins

1) The previously reported method, using acetic acid for extracting His-tag-free r-amelogenins from *E. coli* at a high level of purity (Svensson Bonde and Bulow, 2012) was optimised for His-tagged recombinants prior to further purification using nickel column chromatography, the ”gold standard” purification methodology.
2) Nickel column chromatography proved less than ideal for purifying r-amelogenin since His-tag cleavage was not 100% efficient. This means that cleaved and uncleaved recombinants needed to be subjected to a second round of nickel column chromatography to separate them but this was confounded by the fact that amelogenin contains tri- and di histidine repeats which seem to act as pseudo His tags that made it impossible to produce a final product that exhibited single band purity using silver stained analytical SDS PAGE.
3) As an alternative, preparative SDS PAGE proved to be an effective means for purifying r-amelogenin and was able to produce a final product that exhibited single band purity using silver stained analytical SDS PAGE. Preparative SDS PAGE negated the need for nickel chromatography, therefore rendered the His-tag redundant, and provided the potential to purify His-tag-free r-amelogenins.

5.2 Aim 2: Investigation of the effect of p.Y64H mutation associated with AI in mouse, on amelogenin-amelogenin binding

1) A simple binding assay using UV-transparent microplate was developed with EMD as a surrogate with the aim to be used to compare WT and p.Y64H r-amelogenins in terms of their ability to adsorb to the polystyrene surface of the microwells (EMD-polystyrene binding) and, on top, to bind to the adsorbed EMD (EMD-EMD binding).
2) Given that the use of a His-tag is now redundant to purify r-amelogenin, plasmids coding His-tag-free WT^{-His} and Mut^{-His} r-amelogenins were produced to provide recombinants for use in the microplate-based binding assay. Time limitations only allowed for a pilot study to be carried out, comparing the behaviours of WT^{+His}, Mut^{+His}, WT^{-His} and Mut^{-His} r-amelogenins. The data obtained among His-tag-free r-amelogenins (WT^{-His} and Mut^{-His}) suggested that Mut^{-His} adsorbed strongly, whereas WT His-tag-free r-amelogenin showed weak adsorption. This suggests that the mutation had a significant impact on amelogenin behaviour in terms of its binding properties. In contrast, the His-tagged r-amelogenins (WT^{+His} and Mut^{+His}) both adsorbed strongly in the microplate assay.
3) Therefore, the comparative data between WT^{+His} and WT^{-His} r-amelogenins suggests that His-tags can have a significant impact on amelogenin behaviour and this should be considered when planning functional studies using r-amelogenins purified using His-tag-based methodologies.

5.3 Future work

Following these conclusions, the following prospective can be considered:

1) Given the purifying power of preparative SDS PAGE, r-amelogenins can be obtained, as close as possible to the native M180 amelogenin, for binding studies.

Time limitations allowed the use in the binding assay of r-amelogenins extracted in 3% acetic acid only, illustrated in Figure 89A (because the plasmids expressing His-tag-free r-amelogenins were delivered during the write-up phase of this thesis). So, next:

- the r-amelogenins will all be subjected to a round of preparative SDS PAGE to yield single band purity on silver staining analytical SDS PAGE (Figure 89B).
- Furthermore, r-amelogenins can be expressed in an eukaryotic system, which will allow, for example, phosphorylation on Ser^{16} (Figure 89B).

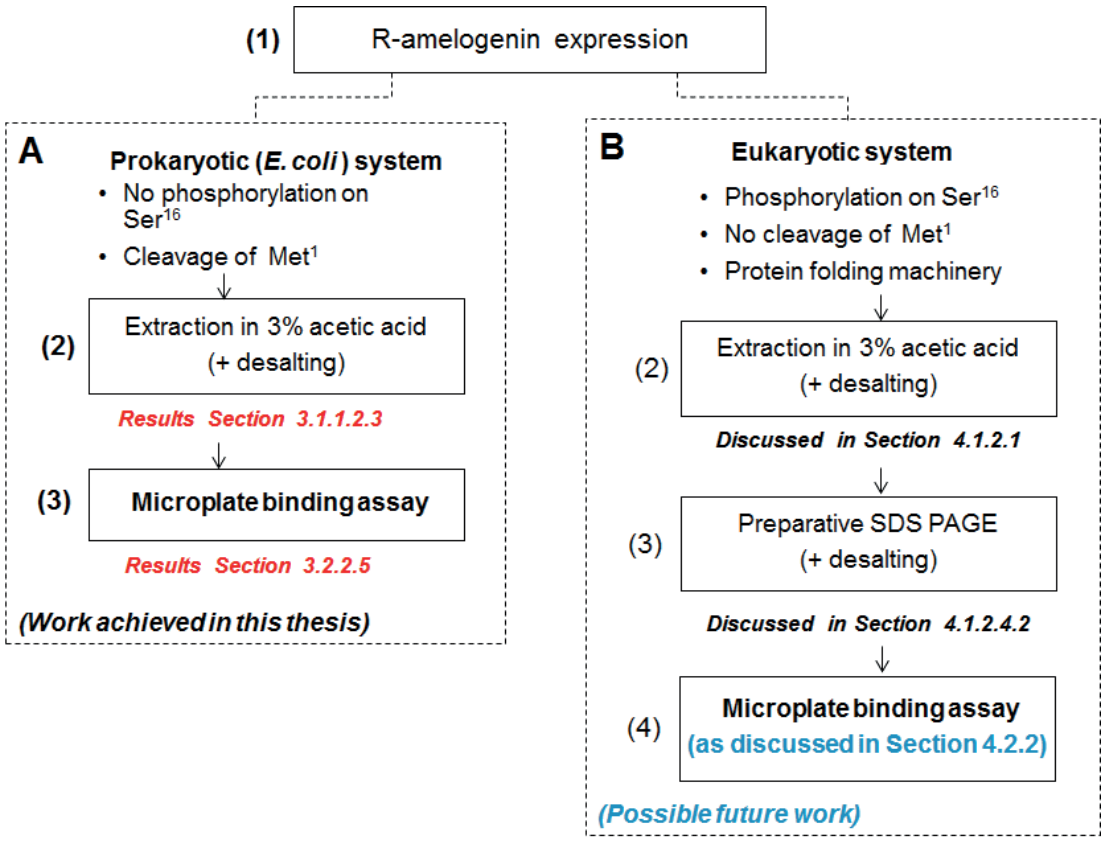


Figure 89 Production and purification of His-tag free r-amelogenins for microplate-based binding assays: current achievements and future prospectives. At this stage (box A), the WT and mutant p.Y64H r-amelogenins were expressed in *E. coli*, extracted in acetic acid, where they account for >90% of all proteins (see Section 4.1.2.1, p. 216) for further details). Due to time constraints, the acetic acid extracts (desalted) were tested directly in the microplate binding assays without further purification by preparative SDS PAGE. Future improvements would employ an eukaryotic system, so that the proteins produced are as close as possible to the native M180 amelogenin. They could be purified by preparative SDS PAGE and after desalting, used in microplate-based binding assay developed (Discussed in Section 4.2.2).

2) To investigate further the effect of the amelogenin mutation p.Y64H *in vitro*, the following prospective is to consider:

2.1) Data reproducibility.

Repeat the pilot binding assay carried out in this thesis, to corroborate the results obtained and confirm their repeatability.

2.2) UV-transparent microplate binding assay: method development

Pursue the development of the microplate-based binding assay, so that absorbance measurements reflect directly and accurately amelogenin adsorption to the sides of the microwells, i.e: Prevent any amelogenin from binding to the bottom surface of the microwells.

2.3) Studying the effects of p.Y64H mutation

- **Choice of techniques for studying protein-protein interactions.**

Considering the experimental pipeline proposed (Figure 87, reproduced below), other protein binding techniques can be employed to confirm or complete those obtained with the UV-transparent microplate-based assay.

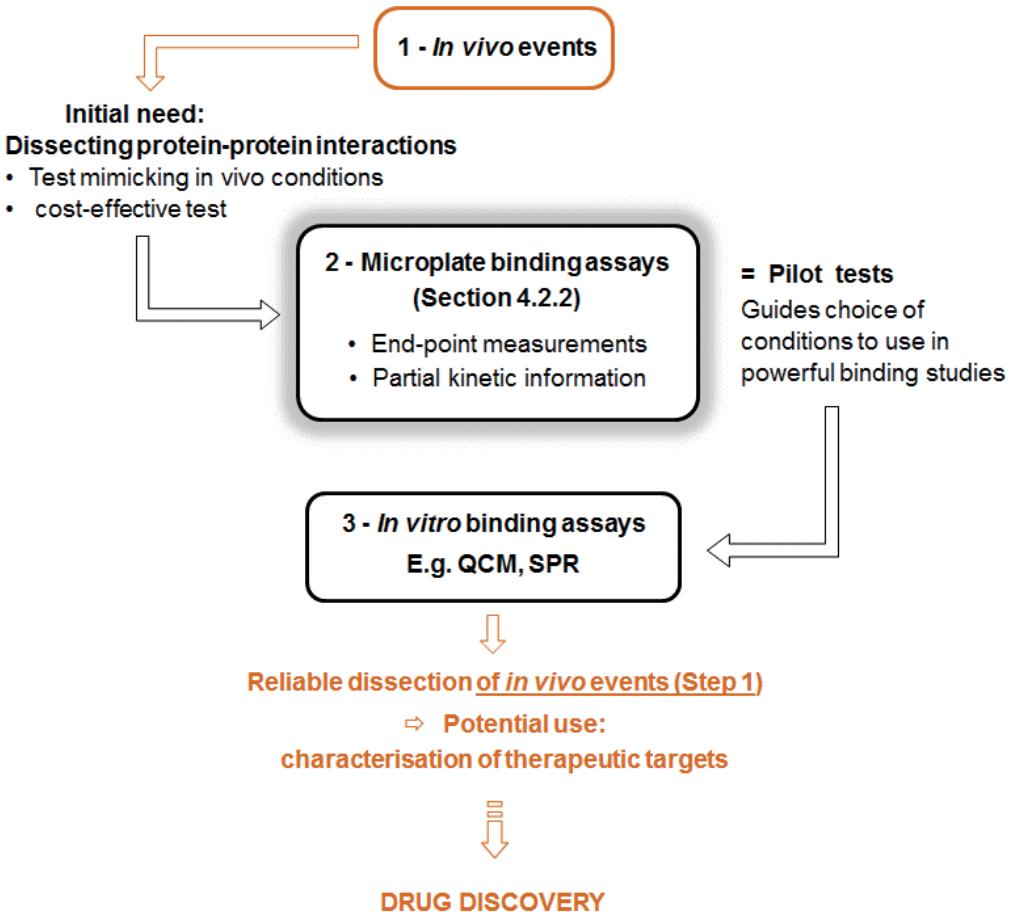


Figure 90 (Figure 88 reproduced) Experimental pipeline for dissecting protein binding behaviours in vivo.

- **Choice of proteins for protein binding assays.**

The results obtained in this thesis suggest that the p.Y64H mutation affected amelogenin-amelogenin binding and aggregation behaviour. The p.Y64H mutation may also affect amelogenin binding to other proteins (if these interactions indeed occur in WT), for instance chaperones, or EMP (such as ameloblastin) co-secreted with amelogenin in the secretory-stage ameloblasts. This point is worth considering further, in order to dissect the pathological mechanisms underlying ER stress- related AI:

.

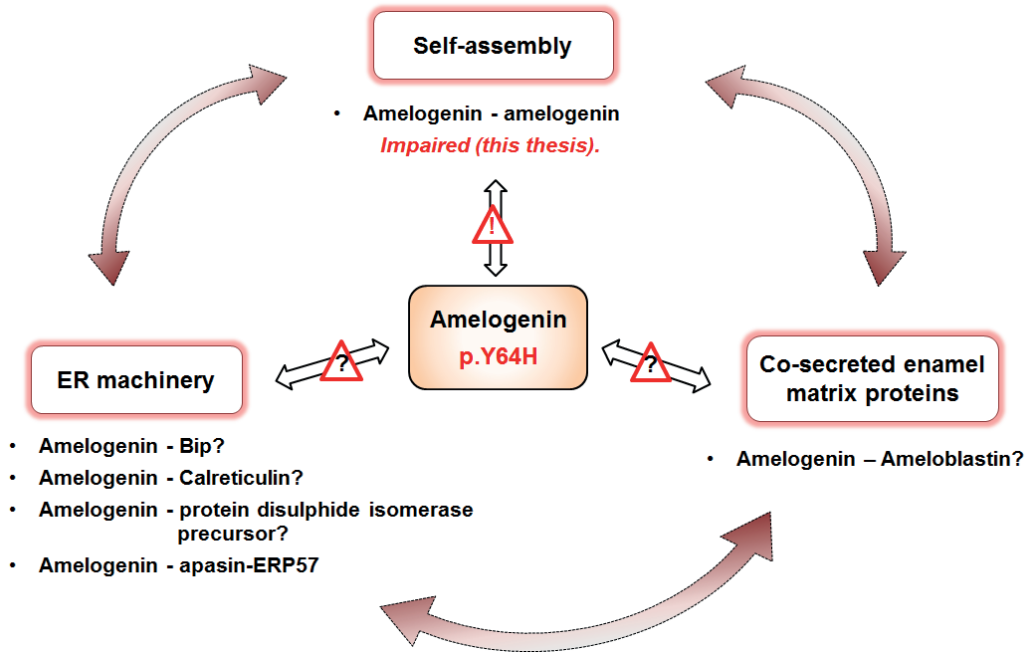


Figure 91 The amelogenin p.Y64H mutation may perturb the balance of protein-protein interactions involving amelogenin. Future routes to investigate for protein-protein interactions are indicated.

Appendix A Preparation of His-tag-free r-amelogenins WT and mutant p.Y64H for microplate-based binding studies

Recombinant His-tag-free amelogenins either WT (WT^{-His}) or carrying the p.Y64H mutation (Mut^{-His}), were expressed by *E. coli* BL21(DE3) competent cells previously transfected with vector pET30a /*Amelx*-WT^{-His} or *Amelx*-Mut^{-His}. The plasmid production and transfection were performed by a commercial company (Novoprotein Scientific, NJ, USA). Vector pET30a carried the T7 promoter, LacI operator, WT or p.Y64H mutant His-tag-free amelogenin gene (*Amelx*-WT^{-His} **or** *Amelx*-Mut^{-His}) and kanamycin resistance gene to allow for the selective growth of transfected cells. The predicted amino acid sequences of the r-amelogenins are shown below in Table 1.

Table 1 Primary sequences or r-amelogenins WT-His and Mut-His. The position of p.Y64H point mutation is indicated in red.

r-amelogenin	Primary sequence
WT-His	MPLPPHPGSPGYINLSYEVLTPLKWYQSMIRQP**Y**PSYGYEPM GGWLHHQIIPVLSQQHPPSHTLQPHHHLPVVPAQQPVAPQQ PMMPVPGHHSMTPTQHHQPNIPPSAQQPFQQPFQPQAIPPQ SHQPMQPQSPLHPMQPLAPQPPLPPLFSMQPLSPILPELPLE AWPATDKTKREEVD
Mut-His	MPLPPHPGSPGYINLSYEVLTPLKWYQSMIRQP**H**PSYGYEPM GGWLHHQIIPVLSQQHPPSHTLQPHHHLPVVPAQQPVAPQQ PMMPVPGHHSMTPTQHHQPNIPPSAQQPFQQPFQPQAIPPQ SHQPMQPQSPLHPMQPLAPQPPLPPLFSMQPLSPILPELPLE AWPATDKTKREEVD

E. coli cells storage and growth, induction of r-amelogenin expression, and extraction in acetic acid were kindly performed by Dr. Sarah Myers. The cell growth and r-amelogenin induction conditions were the same as those described in Section 2.1.1.1 (induction with IPTG) pp. 92 - 93. The extraction in acetic acid is the same as described in 2.1.1.2 p. 94. The acetic acid extracts were desalted against 125 mM formic acid (See section 2.1.1.3, p. 99)

Appendix B Preparation of His-tag-free WT and p.Y64H mutant r-amelogenins for microplate-based binding studies

Recombinant His-tag-free amelogenins either wild-type (WT^{-His}) or carrying the p.Y64H mutation (Mut^{-His}), were expressed by *E. coli* BL21 (DE3) and extracted in acetic acid were (See Appendix A for details). The supernatants of acetic acid extracts were subjected to mass spectrometry.

The results of mass spectrometry analysis are displayed below in Figure 1 (spectra for r-amelogenin WT^{-His}) and Figure 2 (spectra for r-amelogenin Mut^{-His}). This appendix supports the text in Section 4.1.2.1 (p. 217)

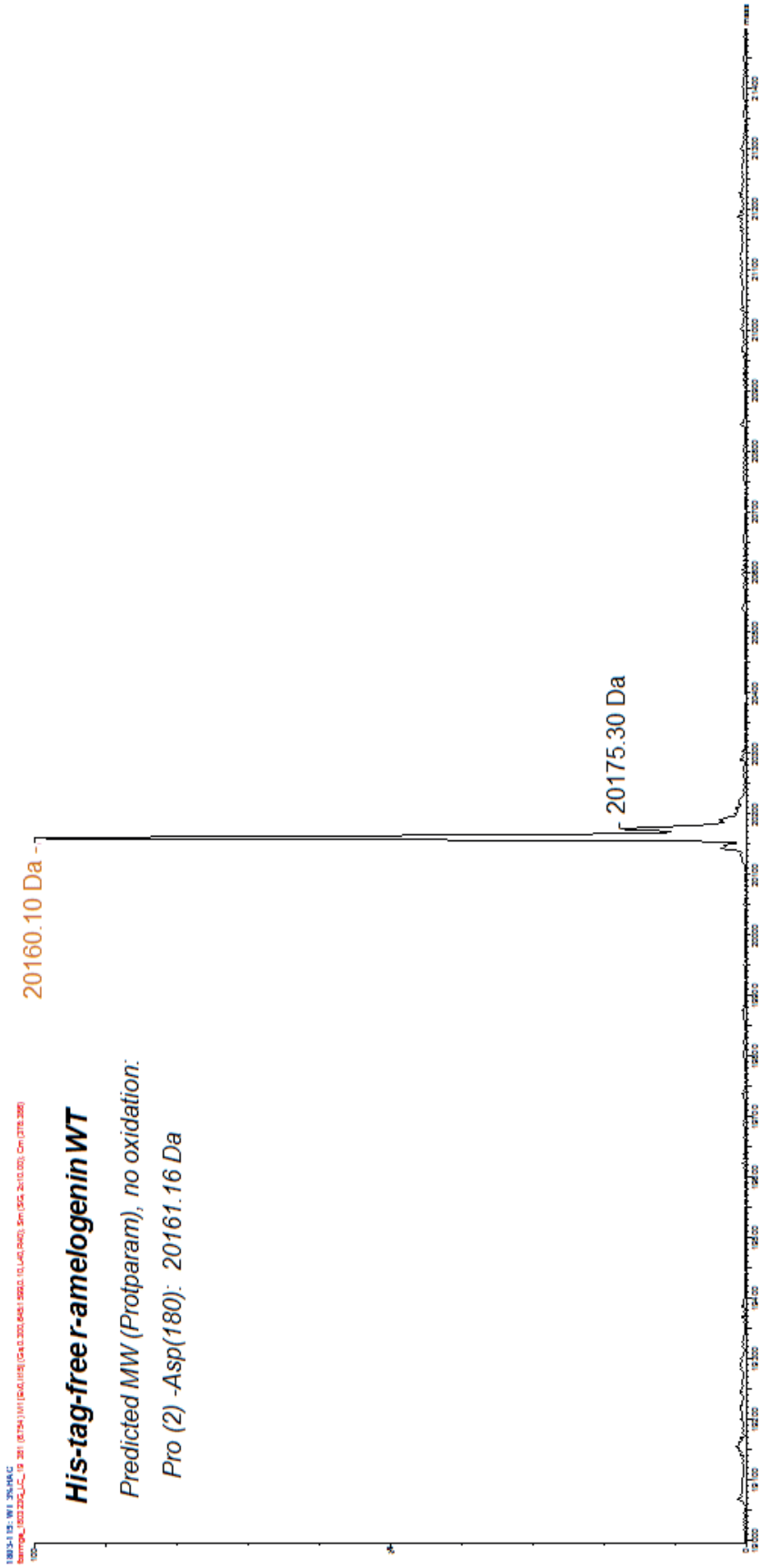


Figure 1 Mass spectrometry analysis of acetic acid extract of r-amelogenin (WT-^{His}) obtained in the supernatant from acetic acid extraction. See text in Dicussion Section 4.1.2.1 (p. 217) for further description of the spectrum. The numbers indicate the peaks corresponding to the different masses detected.

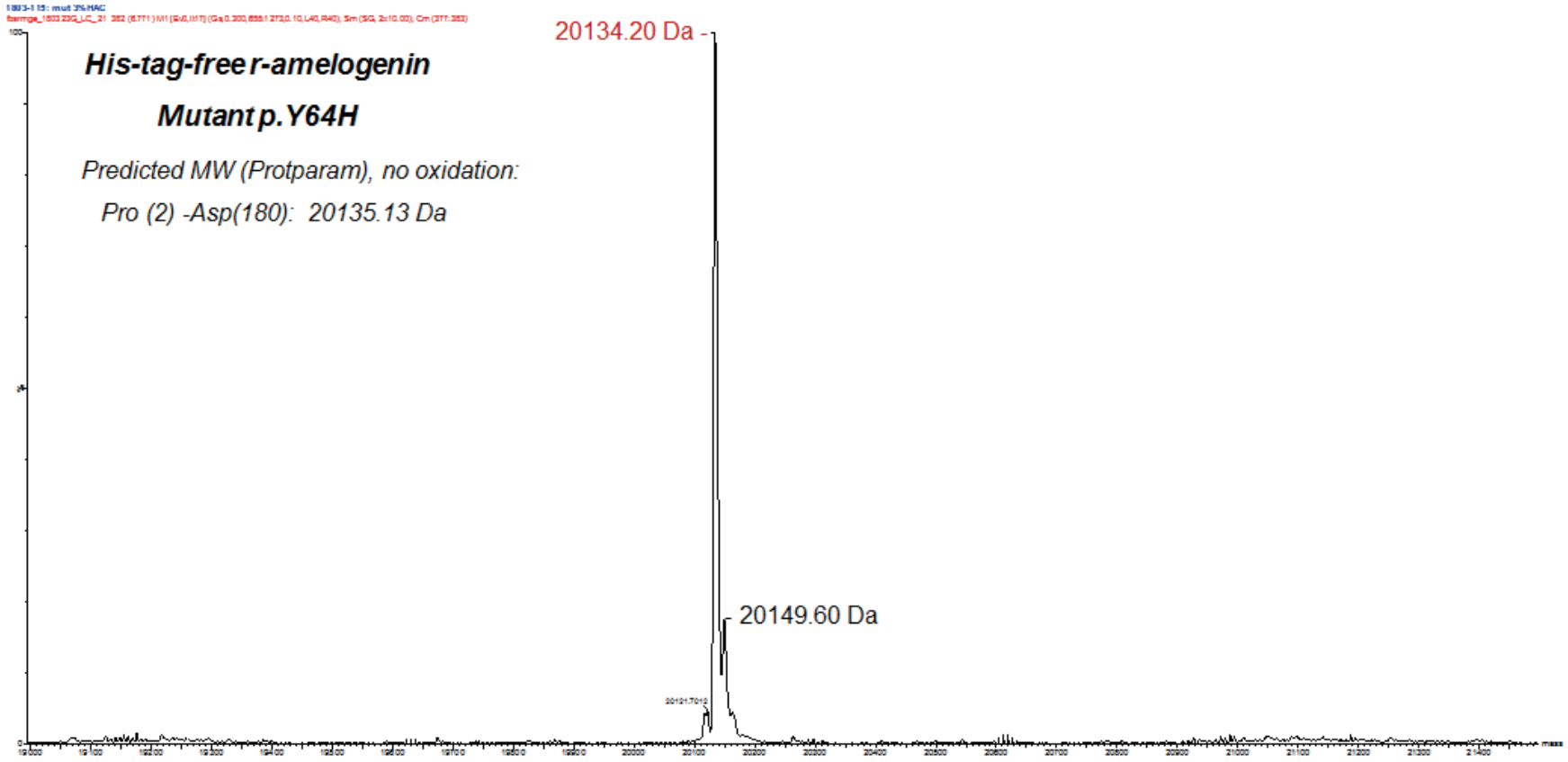


Figure 2 spectrometry analysis of acetic acid extract of r-amelogenin (Mut^{-His}) obtained in the supernatant from acetic acid extraction. See text in Section 4.1.2.1 (p. 217) for further description of the spectrum. The numbers indicate the peaks corresponding to the different masses detected.

Appendix C Determination of a minimum working enzyme/protein ratio

The N-terminal His-tag initially included in the sequence of r-amelogenin for purification purposes had to be removed using HRV3C (71493, Merck), a recombinant restriction-grade protease from human rhinovirus protease. HRV3C catalyses the hydrolysis between glutamine and glycine within the sequence LEVLFQ/GP. It is however an expensive reagent (currently priced at £226.10 / mL stock solution). To limit the need for costly amounts of HRV3C, the cleavage reaction was optimised to determine the most cost-effective enzyme-to-substrate ratio.

The ratios tested ranged from 2 µL enzyme solution per mg desalted acetic acid extract to 20 µL enzyme solution per mg desalted acetic acid extract. The acetic acid extract containing r-amelogenin WT^{+His} was resuspended at 2mg/mL into a solution of 1.5M NaCl, 0.5M Tris-HCl, pH = 7.5, which is the standard cleavage buffer recommended by the manufacturer for the restriction grade HRV3C protease (Millipore-71493, Herfordshire, UK). The protease (in its stock solution at 2U/µL) was added into the r-amelogenin solution, to 2, 2.4, 4, 10, or 20 µL per mg substrate protein. The mixture was left at 4°C over 72 hours and for each condition 50µL aliquots were taken to -80°C every 24 hours. The aliquots taken were lyophilised and resuspended to original concentration in 1X sample buffer for analytical SDS PAGE. The yields of cleaved r-amelogenin were measured by gel densitometry. The optimisation procedure is summarised in Figure 3 below and the results obtained are shown in Figure 4.

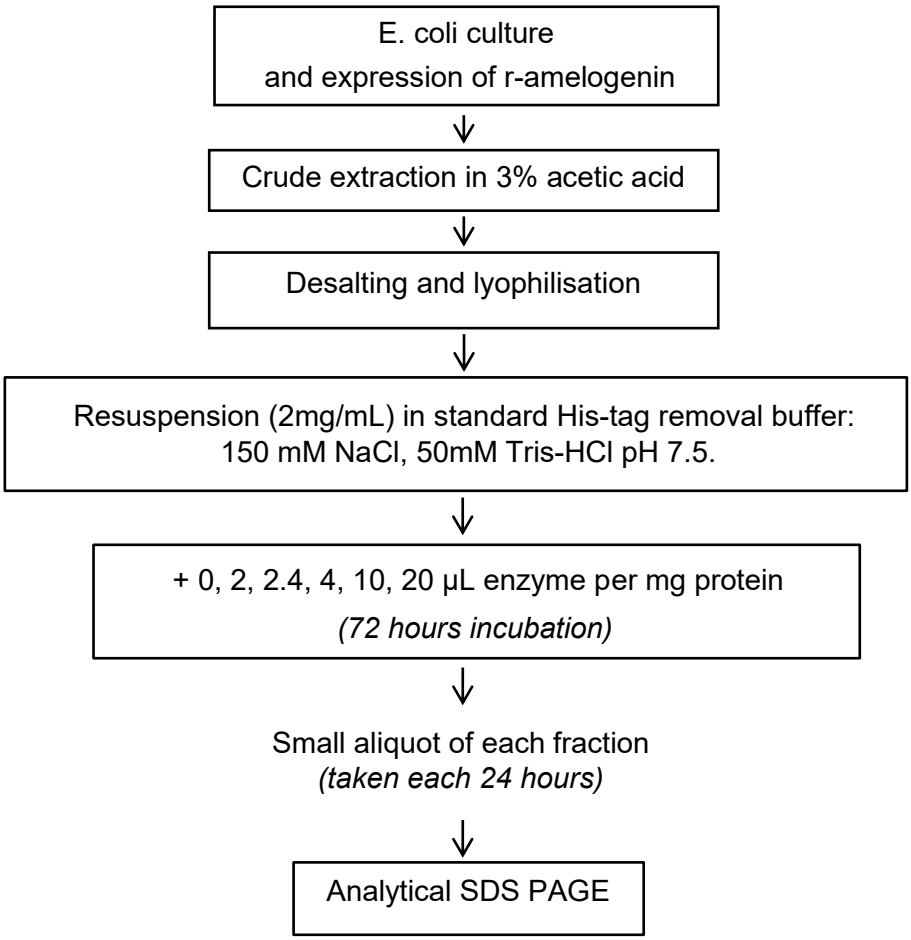


Figure 3 Summary of the procedure optimising the (enzyme-to-target) protein ratio, expressed as µL stock solutions HRV3C per mg desalted acetic acid extract. Acetic acid extracts of r-amelogenin were desalted and resuspended in His-tag removal buffer (recommended by the manufacturer for HRV3C) and stock solutions of HRV3C proteases were added as 0 (Control), 2, 2.4, 4, 10, 20 µL per mg proteins. The mixture was incubated 72 hours, with small aliquots taken each 24 hours for analytical SDS PAGE. The results obtained are shown in Figure 4.

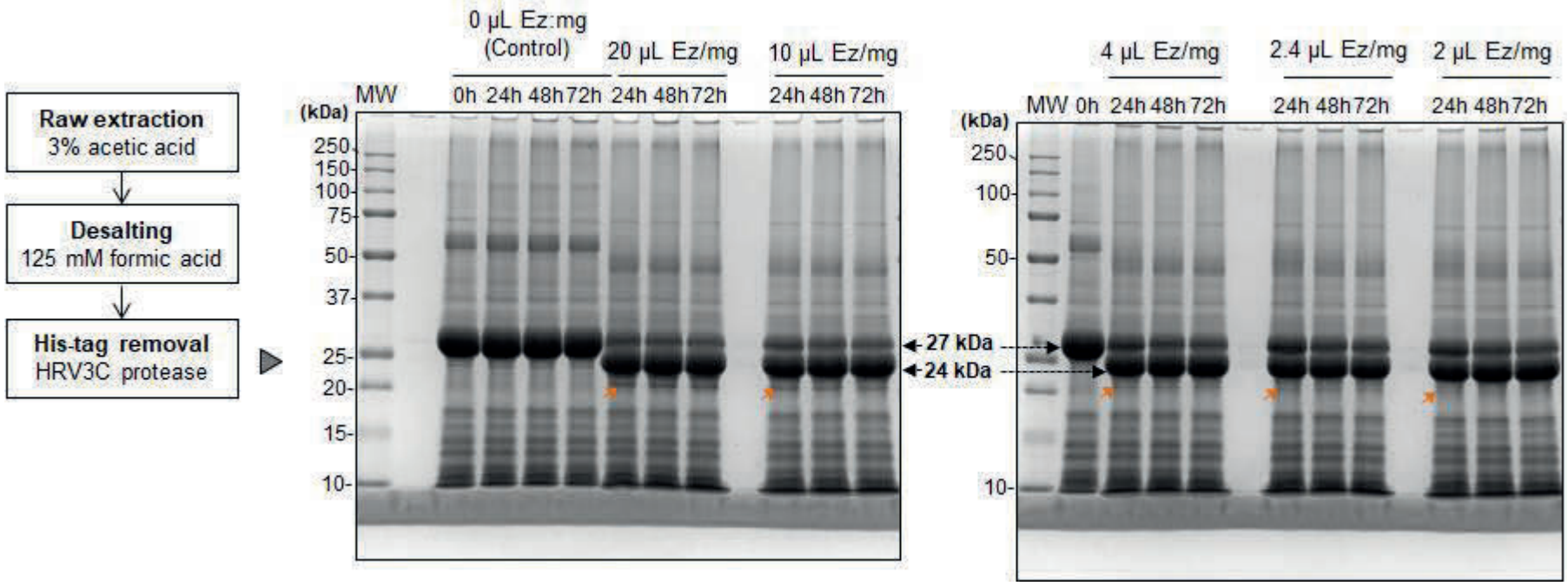


Figure 4 Analytical SDS PAGE with Coomassie Blue staining, comparing the yields of His-tag removal: for each of the (enzyme-to-target protein) ratio, and at 24, 48, 72 hours of incubation. The target protein is r-amelogenin in the acetic acid extract (prepared as shown on the diagram on the left). Without addition of HRV3C (0 µL Ez/mg substrate), the uncleaved 27kDa r-amelogenin remains intact over 72-hour incubation. The addition of HRV3C protease (orange arrows), at 2 - 20 µL per mg target protein, results in the production of a cleaved His-tag-free r-amelogenin (24kDa). However, cleavage was not 100% efficient, since the uncleaved 27kDa band is still present. The cleavage efficiency did not differ significantly between the lanes, indicating that neither the time of incubation nor the (enzyme-to-target protein) ratio affected the yield of His-tag removal. The (enzyme-to-target protein) ratio is referred to as "µL Ez/mg".

Figure 4 shows the effect of increasing incubation times and enzyme-to-substrate ratios on the efficiency of His-tag cleavage. In the control, containing no enzyme (“Control” in Figure 4 above), the uncleaved r-amelogenin at 27kDa is present through the incubation period indicating that there was no significant degradation over the 72-hour incubation time.

In the samples subjected to His-tag removal (i.e., where HRV3C was added to the mixture at 2 to 20 µL per mg protein extract), two additional bands were visible: one minor 22kDa band corresponding to the HRV3C enzyme (orange arrows in Figure 4), and one prominent 24kDa band, corresponding to the cleaved (His-tag-free) r-amelogenin. The 22kDa band corresponding to the HRV3C enzyme was more intense in the samples containing more enzyme. The fact that the band intensity of the ‘cleaved’ (His-tag-free) and ‘uncleaved (His-tagged) r-amelogenin are similar on all ‘test’ lanes indicated that neither the time of incubation, nor the enzyme: substrate ratio significantly affected the yield of His-tag removal. Given this, an enzyme-to-substrate ratio of 3µL of stock enzyme per mg of substrate was used for the cleavage reactions.

In all conditions tested, it was clear that the His-tag removal reaction never approached 100% efficiency.

Appendix D Effect of buffer composition and pH on efficiency of His-tag removal

As described in Appendix C it was decided to use the HRV3C enzyme at an enzyme-to-substrate ratio of 3µL per mg substrate protein mixture. This optimisation was carried out at pH 7.5 in the buffer suggested by the manufacturer -150 mM NaCl, 50mM Tris-HCl . However, in the course of the work, some cleavages were carried out at pH 8 in 50mM Tris-HCl and even at pH 9 in 0.1M Na_2CO_3

The rationale of carrying out the cleavage at pH 8 in 50mM Tris-HCl using was twofold. First the slight increase in pH and the reduction in ionic strength due to omission of the 150mM NaCl would be expected to aid r-amelogenin solubility and possibly allow it to be more accessible to the enzyme.

The rationale of carrying out the cleavage at pH 8 and 9 in 0.1M Na_2CO_3 is that this more extreme pH would be expected to favour further the solubility and reduce the aggregation of r-amelogenin (Tan et al., 1998), therefore increasing its accessibility to the protease HRV3C. The buffer 0.1M Na_2CO_3 at pH = 9 also allowed the cleavage products to be directly labelled with fluorescein at this point if required (as fluorescein labelling is carried out in this buffer system).
The freeze dried acetic acid extract was dissolved at 2mg/mL in the cleavage buffers and stock HRV3C solution was added to 3 µL per mg protein and incubated for 72 hours. The cleavage buffers were: Tris-HCl (pH=8), 0.1M Na_2CO_3 (pH=8), or 0.1M Na_2CO_3 (pH=9). As controls, small aliquots were taken before adding the HRV3C; then small aliquots were taken after adding HRV3C every 24 hours for analytical SDS PAGE. The aliquots taken for analytical SDS PAGE were lyophilised and resuspended to original concentration in 1X sample buffer.
The procedure is summarised in Figure 5 below and the results obtained are shown in Figure 6.

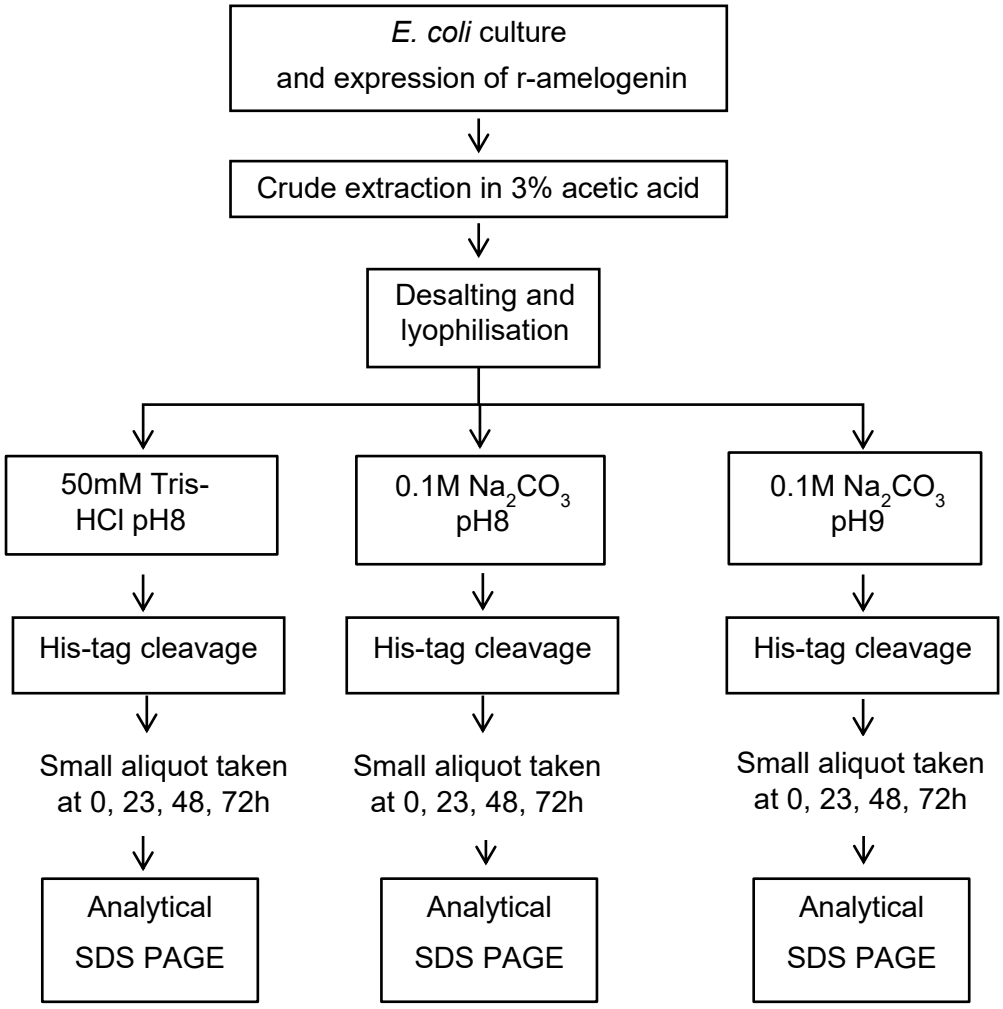


Figure 5 Summary of the procedure designed to test the effect of buffer composition and pH on r-amelogenin His-tag removal efficiency. The His-tag cleavage was carried out by incubating the acetic acid extract comprising r-amelogenin with HRV3C protease at 3 µL stock HRV3C solution per mg proteins for 72 hours. Incubation was done in either of these buffers: Tris-HCl (pH=8), 0.1M Na_2CO_3 (pH=8), or 0.1M Na_2CO_3 (pH=9). Small aliquots were taken before adding the HRV3C as controls and small aliquots were taken after adding HRV3C every 24 hours for SDS PAGE analyses. The results obtained are shown in Figure 6.

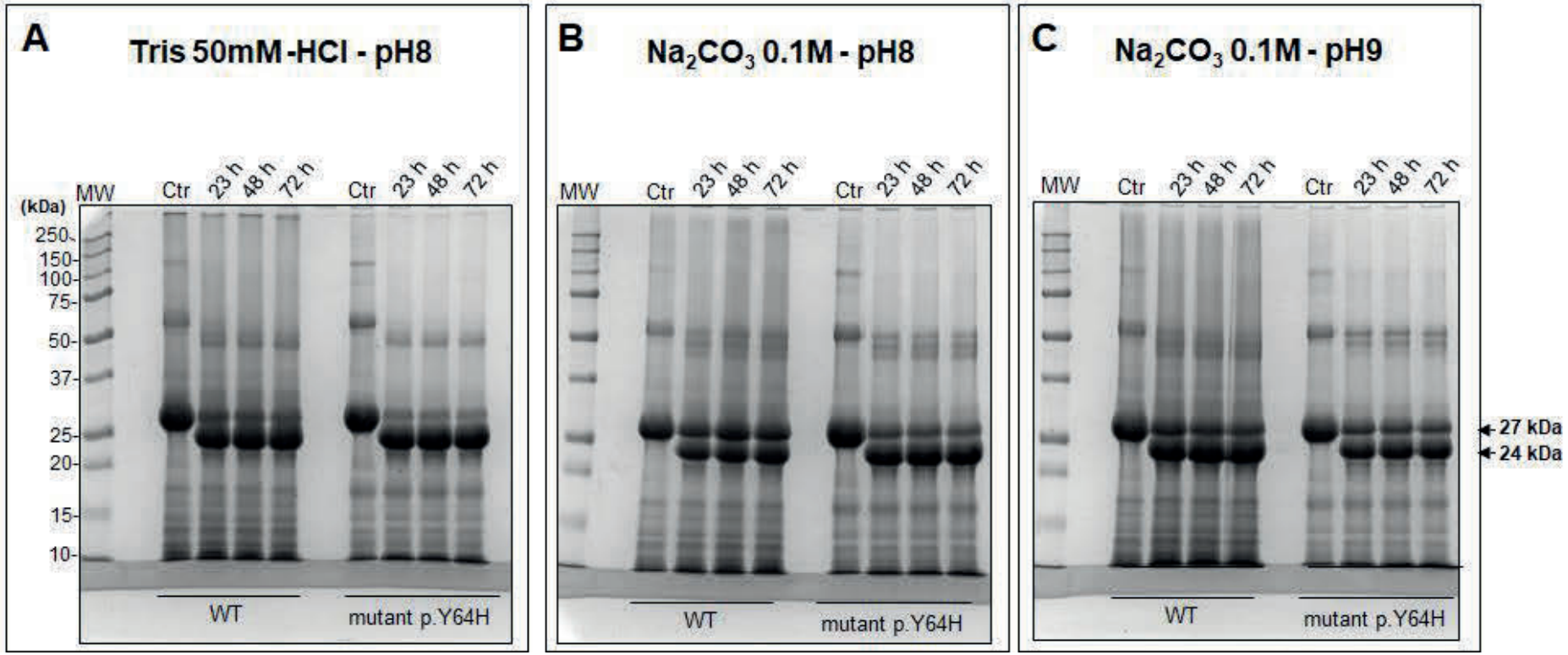


Figure 6 Analytical SDS PAGE with Commassie Blue staining, showing the results of His-tag removal (A) in 50mM Tris-HCl, pH=8 (a standard His-tag cleavage buffer), (B) in 0.1M Na_2CO_3 buffers, pH=8 and (C) in 0.1M Na_2CO_3 buffers, pH=9 (a standard FITC-labelling buffer). For all tested conditions, His-tag cleavage was effective, as a 24 kDa band is visible in all lanes except the controls ("Ctr" lanes), which have not been subjected to His-tag cleavage. Neither the incubation time nor the composition of buffers appeared to significantly affect the cleavage efficiency, whether it is the WT of the mutant p.Y64H r-amelogenin extract.

Analytical SDS PAGE in Figure 6 above showed that for all buffers, His-tag cleavage was effective, as a 24 kDa band is visible in all lanes except the controls, which have not been subjected to His-tag removal. However the cleavage yield never approached 100%.

Neither the composition of buffers nor the duration of incubation appeared to significantly affect the cleavage efficiency, since the intensities of the 'cleaved' and 'uncleaved' bands appeared similar between different lanes (except the controls, which did not contain HRV3C enzyme). Figure 6B shows that the cleavage in 0.1M Na_2CO_3 was as effective as that carried out in 50mM Tris-HCl (Figure 6A. Figure 6C shows that increasing the pH to 9 did not affect the rate of cleavage either.

It did not make significant difference either whether the r-amelogenin was WT or mutant p.Y64H.

Reference

TAN, J., LEUNG, W., MORADIAN-OLDAK, J., ZEICHNER-DAVID, M. & FINCHAM, A. G. (1998) The pH dependent amelogenin solubility and its biological significance. ***Connective tissue research,*** **38, 215-221.**

Appendix E Protein-protein interactions and existing analytical techniques

As mentioned in Section 1.3.6, “Understanding protein interactions is key to understanding normal amelogenesis as well as the aetiology of amelogenesis imperfectas driven by perturbed protein interactions”. To this end, the choice of protein binding studies to carry out is crucial. This appendix will complement Sections 1.3.6.2 -1.3.6.3 by providing further details on the principles, advantages and limitations of some existing *in vivo* and *in vitro* techniques.

E.1 Protein-protein interaction assays *in vivo*

This section reviews major *in vivo* protein-protein interactions studies. It details Y2H assay (Fields and Song, 1989) PCA (Johnsson and Varshavsky, 1994), FRET (Stryer, 1978) and cross-linking (Wong, 1991). Their principles, pros and cons are summarised in Table 2 in this appendix. Other techniques including immunofluorescence microscopy, which previously identified amelogenin binding partners (See Section 1.2.5.2, p. 45 - 49) are not detailed here.

Overall, *in vivo* protein-protein interactions studies are invaluable as they can generate the most relevant information on actual *in vivo* events. However, for detailed characterisation of protein-proteins interactions, *in vitro* studies are carried out as they can provide binding kinetics data, they are versatile and easier to handle.

E.2 Protein-protein interaction assays *in vitro*

There exists a plethora of *in vitro* methods to study protein-protein interactions. This Section addresses *in vitro* studies which do not require labelling or addition of a fusion protein or fusion tag. Popular methods include SPR (Karlsson et al., 1991), QCM (Sauerbrey, 1959), ITC (Freire et al., 1990),DLS (Pecora, 2013) or AFM (Binnig et al., 1986). Their principles, pros and cons are summarised in Table 3 in this appendix. Other techniques related to microplate-based binding assays, including ELISA, are addressed in the Section 1.3.6.3-4 of the thesis and are not detailed in this appendix.

Method	Principle	Advantages	Limitations	previous use in amelogenin studies
Y2H (1)	-Two binding partners respectively fused to a DNA binding domain and a transactivation domain. -Interaction of binding partners activates transcription of a reporter gene.	- high-throughput screening - protein interaction networks **(2)** - study protein domains involved in interactions - physiological environment	- uses fusion proteins - false positives or negatives **(3)** - limited to the nucleus environment **(4)**	-identification of domains and amino acids involved in amelogenin-amelogenin interactions **(5)** (see Sections 1.2.5.2)
PCA (6, 7)	-Two binding partners respectively fused to fragments of a protein (e.g. ubiquitin). - Interaction results in the cleavage or activation of a reporter protein	- see advantages Y2H; -not limited to the nucleus environment **(4)**	- uses fusion proteins - endogeneous protease or ubiquitin cause background **(4)** - time to generate signals	Not documented
FRET (8)	- Proteins fused to fluorophores - Interaction causes donor fluorophore to transfer its excitation energy to acceptor fluorophore	- lower background than PCA - real-time detection of protein-protein interactions in cells	-possible cell auto-fluorescence background	-amelogenin-ameloblastin N-terminal fragments interaction **(9)**
Cross-linking (10)	-A cross-linker binds covalently proteins interacting within a complex	- no labelling or tagging - direct detection	- captures one time-point of protein-protein interactions	characterise amelogenin self-assembly **(11, 12)**

Table 2 Protein-protein interactions studies in *vivo/in cellulo* major techniques: principles, highlights and application to amelogenin[1]

[1] **References (see next page)**

(1) FIELDS, S. & SONG, O.-K. 1989. A novel genetic system to detect protein–protein interactions. *Nature,* 340, 245.

(2) ITO, T., CHIBA, T., OZAWA, R., YOSHIDA, M., HATTORI, M. & SAKAKI, Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences,* 98**,** 4569-4574.

(3) BRÜCKNER, A., POLGE, C., LENTZE, N., AUERBACH, D. & SCHLATTNER, U. 2009. Yeast two-hybrid, a powerful tool for systems biology. *International journal of molecular sciences,* 10**,** 2763-2788.

(4) PIEHLER, J. 2005. New methodologies for measuring protein interactions in vivo and in vitro. *Current opinion in structural biology,* 15**,** 4-14.

(5) PAINE, M. L. & SNEAD, M. L. 1997. Protein interactions during assembly of the enamel organic extracellular matrix. *Journal of Bone and Mineral Research,* 12**,** 221-227.

(6) JOHNSSON, N. & VARSHAVSKY, A. 1994. Split ubiquitin as a sensor of protein interactions in vivo. *Proceedings of the National Academy of Sciences,* 91**,** 10340-10344.

(7) GALARNEAU, A., PRIMEAU, M., TRUDEAU, L.-E. & MICHNICK, S. W. 2002. β-Lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein–protein interactions. *Nature biotechnology,* 20**,** 619.

(8) STRYER, L. 1978. Fluorescence energy transfer as a spectroscopic ruler. *Annu Rev Biochem,* 47**,** 819-846.

(9) MAZUMDER, P., PRAJAPATI, S., BAPAT, R. & MORADIAN-OLDAK, J. 2016. Amelogenin-ameloblastin spatial interaction around maturing enamel rods. *J Dent Res,* 95**,** 1042-1048.

(10) WONG, S. S. 1991. *Chemistry of protein conjugation and cross-linking*, CRC press.

(11) BROOKES, S. J., KIRKHAM, J., LYNGSTADAAS, S. P., SHORE, R. C., WOOD, S. R. & ROBINSON, C. 2000. Spatially related amelogenin interactions in developing rat enamel as revealed by molecular cross-linking studies. *Arch Oral Biol,* 45**,** 937-43.

(12) BROOKES, S. J., LYNGSTADAAS, S. P., ROBINSON, C., SHORE, R. C. & KIRKHAM, J. 2006. Intracellular nanosphere subunit assembly as revealed by amelogenin molecular cross-linking studies. *Eur J Oral Sci,* 114 Suppl 1**,** 280-4; discussion 285-6, 382.

Method	Principle	Advantages	Limitations	Previous use in amelogenin studies
SPR (1)	-Proteins in solution bind to receptors or other proteins immobilised onto a surface of a sensor chip. ⇒ change of reflective index of a polarised light at the back of the sensor chip ⇒SPR signal	- label-free - real-time binding kinetics - versatile (choice of surfaces, buffer) - kinetics constants k_d, and k_a, strength and specificity of interaction	- requires high amounts of proteins (figure?) - requires "free" protein solubility in buffer - costly	- characterising amelogenin self-assembly **(2, 3)**
QCM (4)	- Signal: change of resonance frequency that is directly proportional to the mass of adsorbed proteins	- label-free and sensitive - real-time binding kinetics - versatile (choice of surfaces, buffer)	- requires high amounts of proteins (figure?)	adsorption of amelogenin onto hydroxyapatite surface **(5)**
ITC (6)	- measures the release of heat caused by proteins binding.	- generates thermodynamic parameters : affinity constant stoichiometry, binding enthalpy - no surface immobilisation	- requires large quantities sample - low-throughput, slow - cannot determine constants k_d, and k_a	Amelogenin self-assembly driven by hydrophobic bonding: hydrophobic core with water removal **(7)**
DLS (8)	- measures differences in the degree of light scattering of different populations - hydrodynamic radius) of proteins and oligomers and their size distribution	- high-throughput, fast - reaction stoichiometry, equilibrium dissociation constant, and thermodynamics **(9)** - requires small samples	- size range limited -does not provide precise kinetics binding parameters (unlike SPR)	characterisation amelogenin assembly (reviewed in the Introduction Section 1.2.2.2.4).
AFM	Measures the deflection of a cantilever caused by the	- structure of protein complexes,	- not fast scanning	amelogenin supramolecular assembly

(10)	interaction of its tip (probe) with the sample surface	- forces involved in proteins interactions - can analyse protein folding **(11, 12)**	- possible image artefacts	(reviewed in Section 1.2.2.2.4)

Table 3 Characterisation of protein-protein interaction studies *in vitro*, major techniques: principles, highlights and application to amelogenin[1]

[1] **References**

(1) KARLSSON, R., MICHAELSSON, A. & MATTSSON, L. 1991. Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system. *Journal of immunological methods,* 145**,** 229-240.4

(2) PAINE, M. L., LEI, Y.-P., DICKERSON, K. & SNEAD, M. L. 2002. Altered amelogenin self-assembly based on mutations observed in human X-linked amelogenesis imperfecta (AIH1). *Journal of Biological Chemistry,* 277**,** 17112-17116.

(3) WALD, T., SPOUTIL, F., OSICKOVA, A., PROCHAZKOVA, M., BENADA, O., KASPAREK, P., BUMBA, L., KLEIN, O. D., SEDLACEK, R., SEBO, P., PROCHAZKA, J. & OSICKA, R. 2017. Intrinsically disordered proteins drive enamel formation via an evolutionarily conserved self-assembly motif. *Proc Natl Acad Sci U S A,* 114**,** E1641-E1650.

(4) SAUERBREY, G. 1959. Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung. *Zeitschrift für physik,* 155**,** 206-222.

(5) CONNELLY, C., CICUTO, T., LEAVITT, J., PETTY, A., LITMAN, A., MARGOLIS, H. C. & GERDON, A. E. 2016. Dynamic interactions of amelogenin with hydroxyapatite surfaces are dependent on protein phosphorylation and solution pH. *Colloids and Surfaces B: Biointerfaces,* 148**,** 377-384.

(6) FREIRE, E., MAYORGA, O. L. & STRAUME, M. 1990. Isothermal titration calorimetry. *Analytical chemistry,* 62**,** 950A-959A.

(7) LAKSHMINARAYANAN, R., YOON, I., HEGDE, B. G., FAN, D., DU, C. & MORADIAN-OLDAK, J. 2009. Analysis of secondary structure and self-assembly of amelogenin by variable temperature circular dichroism and isothermal titration calorimetry. *Proteins,* 76**,** 560-9.

(8) PECORA, R. 2013. *Dynamic light scattering: applications of photon correlation spectroscopy*, Springer Science & Business Media.

(9) HANLON, A. D., LARKIN, M. I. & REDDICK, R. M. 2010. Free-solution, label-free protein-protein interactions characterized by dynamic light scattering. *Biophys J,* 98**,** 297-304.

(10) BINNIG, G., QUATE, C. F. & GERBER, C. 1986. Atomic force microscope. *Physical review letters,* 56**,** 930.

(11) FISHER, T. E., OBERHAUSER, A. F., CARRION-VAZQUEZ, M., MARSZALEK, P. E. & FERNANDEZ, J. M. 1999. The study of protein mechanics with the atomic force microscope. *Trends Biochem Sci,* 24**,** 379-384.

(12) YANG, Y., WANG, H. & ERIE, D. A. 2003. Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy. *Methods,* 29**,** 175-187.

References

2005. Corrections and Clarifications. *Science,* 309**,** 2166.

ABBARIN, N., SAN MIGUEL, S., HOLCROFT, J., IWASAKI, K. & GANSS, B. 2015. The enamel protein amelotin is a promoter of hydroxyapatite mineralization. *Journal of Bone and Mineral Research,* 30**,** 775-785.

ADACHI, Y., YAMAMOTO, K., OKADA, T., YOSHIDA, H., HARADA, A. & MORI, K. 2008. ATF6 Is a Transcription Factor Specializing in the Regulation of Quality Control Proteins in the Endoplasmic Reticulum. *Cell Structure and Function,* 33**,** 75-89.

ADZHUBEI, A. A., STERNBERG, M. J. & MAKAROV, A. A. 2013. Polyproline-II helix in proteins: structure and function. *Journal of Molecular Biology,* 425**,** 2100-2132.

AHIRWAR, R., BARIAR, S., BALAKRISHNAN, A. & NAHAR, P. 2015. BSA blocking in enzyme-linked immunosorbent assays is a non-mandatory step: a perspective study on mechanism of BSA blocking in common ELISA protocols. *RSC Advances,* 5**,** 100077-100083.

AICHMAYER, B., MARGOLIS, H. C., SIGEL, R., YAMAKOSHI, Y., SIMMER, J. P. & FRATZL, P. 2005. The onset of amelogenin nanosphere aggregation studied by small-angle X-ray scattering and dynamic light scattering. *Journal of Structural Biology,* 151**,** 239-49.

AICHMAYER, B., WIEDEMANN-BIDLACK, F. B., GILOW, C., SIMMER, J. P., YAMAKOSHI, Y., EMMERLING, F., MARGOLIS, H. C. & FRATZL, P. 2010. Amelogenin Nanoparticles in Suspension: Deviations from Spherical Shape and pH-Dependent Aggregation. *Biomacromolecules,* 11**,** 369-376.

ALALUUSUA, S. 2010. Aetiology of molar-incisor hypomineralisation: a systematic review. *European Archives of Paediatric Dentistry,* 11**,** 53-58.

ALDRED, M. J., SAVARIRAYAN, R. & CRAWFORD, P. J. M. 2003. Amelogenesis imperfecta: a classification and catalogue for the 21st century. *Oral Diseases,* 9**,** 19-23.

ANDREOTTI, A. H. 2003. Native state proline isomerization: an intrinsic molecular switch. *Biochemistry,* 42**,** 9515-9524.

AOBA, T., FUKAE, M., TANABE, T., SHIMIZU, M. & MORENO, E. 1987a. Selective adsorption of porcine-amelogenins onto hydroxyapatite and their inhibitory activity on hydroxyapatite growth in supersaturated solutions. *Calcified Tissue International,* 41**,** 281-289.

AOBA, T. & MORENO, E. C. 1987. The enamel fluid in the early secretory stage of porcine amelogenesis: chemical composition and saturation with respect to enamel mineral. *Calcified Tissue International,* 41**,** 86-94.

AOBA, T., TANABE, T. & MORENO, E. C. 1987b. Proteins in the enamel fluid of immature porcine teeth. *Journal of Dental Research,* 66**,** 1721-6.

BAGRAMIAN, R. A., GARCIA-GODOY, F. & VOLPE, A. R. 2009. The global increase in dental caries. A pending public health crisis. *American Journal of Dentistry,* 22**,** 3-8.

BAJAJ, D. & AROLA, D. 2009. Role of prism decussation on fatigue crack growth and fracture of human enamel. *Acta Biomaterialia,* 5**,** 3045-3056.

BAKER, E. N. & HUBBARD, R. E. 1984. Hydrogen bonding in globular proteins. *Progress in Biophysics and Molecular Biology,* 44**,** 97-179.

BARRON, M. J., BROOKES, S. J., KIRKHAM, J., SHORE, R. C., HUNT, C., MIRONOV, A., KINGSWELL, N. J., MAYCOCK, J., SHUTTLEWORTH, C. A. & DIXON, M. J. 2010. A mutation in the mouse Amelx tri-tyrosyl domain results in impaired secretion of amelogenin and phenocopies human X-linked amelogenesis imperfecta. *Human Molecular Genetics,* 19**,** 1230-47.

BARTLETT, J., BALL, R., KAWAI, T., TYE, C., TSUCHIYA, M. & SIMMER, J. 2006a. Origin, splicing, and expression of rodent amelogenin exon 8. *Journal of Dental Research,* 85**,** 894-899.

BARTLETT, J. D. 2013. Dental enamel development: proteinases and their enamel matrix substrates. *ISRN Dentistry,* 2013**,** 684607.

BARTLETT, J. D., GANSS, B., GOLDBERG, M., MORADIAN-OLDAK, J., PAINE, M. L., SNEAD, M. L., WEN, X., WHITE, S. N. & ZHOU, Y. L. 2006b. Protein–Protein Interactions of the Developing Enamel Matrix. *In:* GERALD, P. S. (ed.) *Current Topics in Developmental Biology.* Academic Press.

BARTLETT, J. D. & SIMMER, J. P. 2014. Kallikrein-related peptidase-4 (KLK4): role in enamel formation and revelations from ablated mice. *Frontiers in Physiology,* 5**,** 240.

BEN-BASSAT, A., BAUER, K., CHANG, S.-Y., MYAMBO, K., BOOSMAN, A. & CHANG, S. 1987. Processing of the initiation methionine from proteins: properties of the Escherichia coli methionine aminopeptidase and its gene structure. *Journal of bacteriology,* 169**,** 751-757.

BERKMAN, M. & SINGER, A. 1971. Demonstration of the lyon hypothesis in X-linked dominant hypoplastic amelogenesis imperfecta. *Birth Defects Original Article Series,* 7**,** 204-209.

BERTOLOTTI, A., ZHANG, Y., HENDERSHOT, L. M., HARDING, H. P. & RON, D. 2000. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nature Cell Biology,* 2**,** 326-32.

BETTS, M. J. & RUSSELL, R. B. 2003. Amino acid properties and consequences of substitutions. *Bioinformatics for Geneticists***,** 289-316.

BEYELER, M., SCHILD, C., LUTZ, R., CHIQUET, M. & TRUEB, B. 2010. Identification of a fibronectin interaction site in the extracellular matrix protein ameloblastin. *Exp Cell Res,* 316**,** 1202-1212.

BIESIADECKI, B. J. & JIN, J. P. 2011. A high-throughput solid-phase microplate protein-binding assay to investigate interactions between myofilament proteins. *J Biomed Biotechnol,* 2011**,** 421701.

BONASS, W. A., KIRKHAM, J., BROOKES, S. J., SHORE, R. C. & ROBINSON, C. 1994. Isolation and characterisation of an alternatively-spliced rat amelogenin cDNA: LRAP--a highly conserved, functional alternatively-spliced amelogenin? *Biochimica et Biophysica Acta,* 1219**,** 690-2.

BORNHORST, J. A. & FALKE, J. J. 2000. Purification of proteins using polyhistidine affinity tags. *Methods in enzymology,* 326**,** 245-254.

BRAAKMAN, I. & HEBERT, D. N. 2013. Protein folding in the endoplasmic reticulum. *Cold Spring Harbor perspectives in biology,* 5**,** a013201.

BREWER, J. W. 2014. Regulatory crosstalk within the mammalian unfolded protein response. *Cellular and Molecular Life Sciences,* 71**,** 1067-1079.

BROMLEY, K. M., KISS, A. S., LOKAPPA, S. B., LAKSHMINARAYANAN, R., FAN, D., NDAO, M., EVANS, J. S. & MORADIAN-OLDAK, J. 2011. Dissecting amelogenin protein nanospheres characterization of metastable oligomers. *Journal of Biological Chemistry,* 286**,** 34643-34653.

BROOKES, S. J., BARRON, M. J., BOOT-HANDFORD, R., KIRKHAM, J. & DIXON, M. J. 2014. Endoplasmic reticulum stress in amelogenesis imperfecta and phenotypic rescue using 4-phenylbutyrate. *Human Molecular Genetics,* 23**,** 2468-80.

BROOKES, S. J., BARRON, M. J., DIXON, M. J. & KIRKHAM, J. 2017a. The Unfolded Protein Response in Amelogenesis and Enamel Pathologies. *Frontiers in Physiology,* 8.

BROOKES, S. J., BARRON, M. J., SMITH, C. E., POULTER, J. A., MIGHELL, A. J., INGLEHEARN, C. F., BROWN, C. J., RODD, H., KIRKHAM, J. & DIXON, M. J. 2017b. Amelogenesis imperfecta caused by N-terminal enamelin point mutations in mice and men is driven by endoplasmic reticulum stress. *Human molecular genetics,* 26**,** 1863-1876.

BROOKES, S. J., BONASS, W. A., KIRKHAM, J. & ROBINSON, C. 1994. The human amelogenin C-terminal sequence is completely homologous to the C-terminal sequence of amelogenin in all species so far studied. *Journal of Dental Research,* 73**,** 716-7.

BROOKES, S. J., KIRKHAM, J., LYNGSTADAAS, S. P., SHORE, R. C., WOOD, S. R. & ROBINSON, C. 2000. Spatially related amelogenin interactions in developing rat enamel as revealed by molecular cross-linking studies. *Arch Oral Biol,* 45**,** 937-43.

BROOKES, S. J., KIRKHAM, J., SHORE, R. C., BONASS, W. A. & ROBINSON, C. 1998. Enzyme compartmentalization during biphasic enamel matrix processing. *Connective Tissue Research,* 39**,** 89-99; discussion 141-9.

BROOKES, S. J., KIRKHAM, J., SHORE, R. C., WOOD, S. R., SLABY, I. & ROBINSON, C. 2001. Amelin extracellular processing and aggregation during rat incisor amelogenesis. *Archives of Oral Biology,* 46**,** 201-208.

BROOKES, S. J., LYNGSTADAAS, S. P., ROBINSON, C., SHORE, R. C. & KIRKHAM, J. 2006. Intracellular nanosphere subunit assembly as revealed by amelogenin molecular cross-linking studies. *European Journal of Oral Sciences,* 114 Suppl 1**,** 280-4; discussion 285-6, 382.

BROOKES, S. J., LYNGSTADAAS, S. P., ROBINSON, C., SHORE, R. C., WOOD, S. R. & KIRKHAM, J. 2002. Enamelin compartmentalization in developing porcine enamel. *Connective Tissue Research,* 43**,** 477-81.

BROOKES, S. J., ROBINSON, C., KIRKHAM, J. & BONASS, W. A. 1995. Biochemistry and molecular biology of amelogenin proteins of developing dental enamel. *Archives of Oral Biology,* 40**,** 1-14.

BROWN, A. 2009. Analysis of cooperativity by isothermal titration calorimetry. *International journal of molecular sciences,* 10**,** 3457-3477.

BRÜCKNER, A., POLGE, C., LENTZE, N., AUERBACH, D. & SCHLATTNER, U. 2009. Yeast two-hybrid, a powerful tool for systems biology. *International Journal of Molecular Sciences,* 10**,** 2763-2788.

BRUNTON, P. A., DAVIES, R. P. W., BURKE, J. L., SMITH, A., AGGELI, A., BROOKES, S. J. & KIRKHAM, J. 2013. Treatment of early caries lesions using biomimetic self-assembling peptides – a clinical safety trial. *British Dental Journal,* 215**,** E6.

BUCHKO, G. W., LIN, G., TARASEVICH, B. J. & SHAW, W. J. 2013. A solution NMR investigation into the impaired self-assembly properties of two murine amelogenins containing the point mutations T21-->I or P41-->T. *Archives of Biochemistry and Biophysics,* 537**,** 217-24.

BUCHKO, G. W. & SHAW, W. J. 2015. Improved protocol to purify untagged amelogenin–Application to murine amelogenin containing the equivalent P70→ T point mutation observed in human amelogenesis imperfecta. *Protein expression and purification,* 105**,** 14-22.

BUTLER, J., NI, L., BROWN, W., JOSHI, K., CHANG, J., ROSENBERG, B. & VOSS JR, E. 1993. The immunochemistry of sandwich elisas—VI. Greater than 90% of monoclonal and 75% of polyclonal anti-fluorescyl capture antibodies (CAbs) are denatured by passive adsorption. *Molecular immunology,* 30**,** 1165-1175.

BYRD, K. E. 1997. Characterization of brux-like movements in the laboratory rat by optoelectronic mandibular tracking and electromyographic techniques. *Archives of Oral Biology,* 42**,** 33-43.

CARNEIRO, K. M., ZHAI, H., ZHU, L., HORST, J. A., SITLIN, M., NGUYEN, M., WAGNER, M., SIMPLICIANO, C., MILDER, M. & CHEN, C.-L. 2016. Amyloid-like ribbons of amelogenins in enamel mineralization. *Scientific Reports,* 6**,** 23105.

CATERINA, J. J., SKOBE, Z., SHI, J., DING, Y., SIMMER, J. P., BIRKEDAL-HANSEN, H. & BARTLETT, J. D. 2002. Enamelysin (matrix metalloproteinase 20)-deficient mice display an amelogenesis imperfecta phenotype. *Journal of Biological Chemistry,* 277**,** 49598-49604.

ČERNÝ, R., SLABY, I., HAMMARSTRÖM, L. & WURTZ, T. 1996. A novel gene expressed in rat ameloblasts codes for proteins with cell binding domains. *Journal of Bone and Mineral Research,* 11**,** 883-891.

CHALLENER, C. A. 2017. For Lyophilization, Excipients Really Do Matter.

CHANT, A., KRAEMER-PECORE, C. M., WATKIN, R. & KNEALE, G. G. 2005. Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. *Protein expression and purification,* 39**,** 152-159.

CHART, H., EVANS, J., CHALMERS, R. & SALMON, R. 1998. Escherichia coli O157 serology: false-positive ELISA results caused by human antibodies binding to bovine serum albumin. *Letters in applied microbiology,* 27**,** 76-78.

CHEN, A. K., CHENG, Z., BEHLKE, M. A. & TSOURKAS, A. 2008. Assessing the sensitivity of commercially available fluorophores to the intracellular environment. *Analytical chemistry,* 80**,** 7437-7444.

CHEN, Y. & BRANDIZZI, F. 2013. IRE1: ER stress sensor and cell fate executor. *Trends in Cell Biology,* 23**,** 547-55.

CHENG, L., LIN, Z. K., SHU, R., LIU, D. L., ZHANG, X. L., LIU, B., WANG, J. & TIAN, L. 2012. Analogous effects of recombinant human full-length amelogenin expressed by Pichia pastoris yeast and enamel matrix derivative in vitro. *Cell Proliferation,* 45**,** 456-65.

CHO, E. S., KIM, K.-J., LEE, K.-E., LEE, E.-J., YUN, C. Y., LEE, M.-J., SHIN, T. J., HYUN, H.-K., KIM, Y.-J., LEE, S.-H., JUNG, H.-S., LEE, Z. H. & KIM, J.-W. 2014. Alteration of Conserved Alternative Splicing in AMELX Causes Enamel Defects. *Journal of Dental Research,* 93**,** 980-987.

CHUN, K., CHOI, H. & LEE, J. 2014. Comparison of mechanical property and role between enamel and dentin in the human teeth. *Journal of Dental Biomechanics,* 5**,** 1758736014520809.

COFFIELD, K. D., PHILLIPS, C., BRADY, M., ROBERTS, M. W., STRAUSS, R. P. & WRIGHT, J. T. 2005. The psychosocial impact of developmental dental defects in people with hereditary amelogenesis imperfecta. *The Journal of the American Dental Association,* 136**,** 620-630.

COLLIER, P. M., SAUK, J. J., ROSENBLOOM, J., YUAN, Z. A. & GIBSON, C. W. 1997. An amelogenin gene defect associated with human X-linked amelogenesis imperfecta. *Archives of Oral Biology,* 42**,** 235-242.

CONNELLY, C., CICUTO, T., LEAVITT, J., PETTY, A., LITMAN, A., MARGOLIS, H. C. & GERDON, A. E. 2016. Dynamic interactions of amelogenin with hydroxyapatite surfaces are dependent on protein phosphorylation and solution pH. *Colloids and Surfaces B: Biointerfaces,* 148**,** 377-384.

COX, J. S., SHAMU, C. E. & WALTER, P. 1993. Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell,* 73**,** 1197-206.

CRAWFORD, P. J., ALDRED, M. & BLOCH-ZUPAN, A. 2007. Amelogenesis imperfecta. *Orphanet Journal of Rare Diseases,* 2**,** 17.

CREDLE, J. J., FINER-MOORE, J. S., PAPA, F. R., STROUD, R. M. & WALTER, P. 2005. On the mechanism of sensing unfolded protein in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 18773-84.

CROMBIE, F., MANTON, D. & KILPATRICK, N. 2009. Aetiology of molar–incisor hypomineralization: a critical review. *International Journal of Paediatric Dentistry,* 19**,** 73-83.

CULLINAN, S. B., ZHANG, D., HANNINK, M., ARVISAIS, E., KAUFMAN, R. J. & DIEHL, J. A. 2003. Nrf2 is a direct PERK substrate and effector of PERK-dependent cell survival. *Molecular and Cellular Biology,* 23**,** 7198-209.

CUY, J. L., MANN, A. B., LIVI, K. J., TEAFORD, M. F. & WEIHS, T. P. 2002. Nanoindentation mapping of the mechanical properties of human molar tooth enamel. *Archives of Oral Biology,* 47**,** 281-291.

DACULSI, G. & KEREBEL, B. 1978. High-resolution electron microscope study of human enamel crystallites: Size, shape, and growth. *Journal of Ultrastructure Research,* 65**,** 163-172.

DACULSI, G., MENANTEAU, J., KEREBEL, L. M. & MITRE, D. 1984. Length and shape of enamel crystals. *Calcified Tissue International,* 36**,** 550-5.

DE JONG, W. W., ZWEERS, A. & COHEN, L. H. 1978. Influence of single amino acid substitutions on electrophoretic mobility of sodium dodecyl sulfate-protein complexes. *Biochemical and Biophysical Research Communications,* 82**,** 532-539.

DE LA DURE-MOLLA, M., QUENTRIC, M., YAMAGUTI, P. M., ACEVEDO, A. C., MIGHELL, A. J., VIKKULA, M., HUCKERT, M., BERDAL, A. & BLOCH-ZUPAN, A. 2014. Pathognomonic oral profile of Enamel Renal Syndrome (ERS) caused by recessive FAM20A mutations. *Orphanet Journal of Rare Diseases,* 9**,** 84.

DE MENEZES OLIVEIRA, M. A., TORRES, C. P., GOMES-SILVA, J. M., CHINELATTI, M. A., DE MENEZES, F. C., PALMA-DIBB, R. G. & BORSATTO, M. C. 2010. Microstructure and mineral composition of

dental enamel of permanent and deciduous teeth. *Microscopy Research and Technique,* 73**,** 572-7.

DEAKINS, M. & BURT, R. L. 1944. The deposition of calcium, phosphorus, and carbon dioxide in calcifying dental enamel. *Journal of Biological Chemistry,* 156**,** 77-83.

DELAK, K., HARCUP, C., LAKSHMINARAYANAN, R., SUN, Z., FAN, Y., MORADIAN-OLDAK, J. & EVANS, J. S. 2009. The tooth enamel protein, porcine amelogenin, is an intrinsically disordered protein with an extended molecular configuration in the monomeric form. *Biochemistry,* 48**,** 2272-81.

DEREWENDA, Z. S. 2004. The use of recombinant methods and molecular engineering in protein crystallization. *Methods,* 34**,** 354-363.

DOBSON, C. M. 2003. Protein folding and misfolding. *Nature,* 426**,** 884-890.

DONG, J., AMOR, D., ALDRED, M. J., GU, T., ESCAMILLA, M. & MACDOUGALL, M. 2005. DLX3 mutation associated with autosomal dominant amelogenesis imperfecta with taurodontism. *American Journal of Medical Genetics Part A,* 133A**,** 138-141.

DU, C., FALINI, G., FERMANI, S., ABBOTT, C. & MORADIAN-OLDAK, J. 2005. Supramolecular assembly of amelogenin nanospheres into birefringent microribbons. *Science,* 307**,** 1450-1454.

DUMONT, E. R. 1995. Enamel Thickness and Dietary Adaptation among Extant Primates and Chiropterans. *Journal of Mammalogy,* 76**,** 1127-1136.

EASTOE, J. E. 1965. The Chemical Composition of Bone and Tooth. *Advances in Fluorine Research and Dental Caries Prevention,* 21**,** 5-17.

EGGERT, F. M., ALLEN, G. A. & BURGESS, R. C. 1973. Amelogenins. Purification and partial characterization of proteins from developing bovine dental enamel. *Biochemical journal,* 131**,** 471-484.

EL-SAYED, W., PARRY, D. A., SHORE, R. C., AHMED, M., JAFRI, H., RASHID, Y., AL-BAHLANI, S., AL HARASI, S., KIRKHAM, J. & INGLEHEARN, C. F. 2009. Mutations in the beta propeller WDR72 cause autosomal-recessive hypomaturation amelogenesis imperfecta. *The American Journal of Human Genetics,* 85**,** 699-705.

ELLIS, R. J. & HARTL, F. U. 1999. Principles of protein folding in the cellular environment. *Current Opinion in Structural Biology,* 9**,** 102-110.

FAIR, B. & JAMIESON, A. 1980. Studies of protein adsorption on polystyrene latex surfaces. *Journal of colloid and interface science,* 77**,** 525-534.

FAN, D., IIJIMA, M., BROMLEY, K. M., YANG, X., MATHEW, S. & MORADIAN-OLDAK, J. 2011. The cooperation of enamelin and amelogenin in controlling octacalcium phosphate crystal morphology. *Cells Tissues Organs,* 194**,** 194-198.

FANG, P. A., CONWAY, J. F., MARGOLIS, H. C., SIMMER, J. P. & BENIASH, E. 2011. Hierarchical self-assembly of amelogenin and the regulation of biomineralization at the nanoscale. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 14097-102.

FELGUEIRAS, H. P., MURTHY, N. S., SOMMERFELD, S. D., BRÁS, M. M., MIGONNEY, V. R. & KOHN, J. 2016. Competitive adsorption of plasma proteins using a quartz crystal microbalance. *ACS Appl Mater Interfaces,* 8**,** 13207-13217.

FERRARIO, V. F., SFORZA, C., ZANOTTI, G. & TARTAGLIA, G. M. 2004. Maximal bite forces in healthy young adults as predicted by surface electromyography. *Journal of Dentistry,* 32**,** 451-457.

FIELDS, S. & SONG, O.-K. 1989. A novel genetic system to detect protein–protein interactions. *Nature,* 340**,** 245.

FINCHAM, A. G., BELCOURT, A. B., LYARUU, D. M. & TERMINE, J. D. 1982. Comparative protein biochemistry of developing dental enamel matrix from five mammalian species. *Calcified Tissue International,* 34**,** 182-9.

FINCHAM, A. G., BELCOURT, A. B. & TERMINE, J. 1981a. The molecular composition of bovine fetal enamel matrix. *In:* VEIS, A. (ed.) *The chemistry and biology of mineralised connective tissues.* Amsterdam: Elsevier.

FINCHAM, A. G., BELCOURT, A. B., TERMINE, J. D., BUTLER, W. T. & COTHRAN, W. C. 1981b. Dental enamel matrix: Sequences of two amelogenin polypeptides. *Bioscience Reports,* 1**,** 771-778.

FINCHAM, A. G., BELCOURT, A. B., TERMINE, J. D., BUTLER, W. T. & COTHRAN, W. C. 1983. Amelogenins. Sequence homologies in enamel-matrix proteins from three mammalian species. *Biochemical Journal,* 211**,** 149-54.

FINCHAM, A. G., BESSEM, C. C., LAU, E. C., PAVLOVA, Z., SHULER, C., SLAVKIN, H. C. & SNEAD, M. L. 1991a. Human developing enamel proteins exhibit a sex-linked dimorphism. *Calcified Tissue International,* 48**,** 288-90.

FINCHAM, A. G., HU, Y., LAU, E. C., SLAVKIN, H. C. & SNEAD, M. L. 1991b. Amelogenin post-secretory processing during biomineralization in the postnatal mouse molar tooth. *Archives of Oral Biology,* 36**,** 305-17.

FINCHAM, A. G., HU, Y. Y., PAVLOVA, Z., SLAVKIN, H. C. & SNEAD, M. L. 1989. Human amelogenins: sequences of "TRAP" molecules. *Calcified Tissue International,* 45**,** 243-50.

FINCHAM, A. G. & MORADIAN-OLDAK, J. 1993. Amelogenin Post-translational Modifications: Carboxy-Terminal Processing and the Phosphorylation of Bovine and Porcine "TRAP" and "LRAP" Amelogenins. *Biochemical and Biophysical Research Communications,* 197**,** 248-255.

FINCHAM, A. G., MORADIAN-OLDAK, J., DIEKWISCH, T. G., LYARUU, D. M., WRIGHT, J. T., BRINGAS, P., JR. & SLAVKIN, H. C. 1995. Evidence for amelogenin "nanospheres" as functional components of secretory-stage enamel matrix. *Journal of Structural Biology,* 115**,** 50-9.

FINCHAM, A. G., MORADIAN-OLDAK, J. & SIMMER, J. P. 1999. The Structural Biology of the Developing Dental Enamel Matrix. *Journal of Structural Biology,* 126**,** 270-299.

FINCHAM, A. G., MORADIAN-OLDAK, J., SIMMER, J. P., SARTE, P., LAU, E. C., DIEKWISCH, T. & SLAVKIN, H. C. 1994. Self-Assembly of a Recombinant Amelogenin Protein Generates Supramolecular Structures. *Journal of Structural Biology,* 112**,** 103-109.

FISHER, T. E., OBERHAUSER, A. F., CARRION-VAZQUEZ, M., MARSZALEK, P. E. & FERNANDEZ, J. M. 1999. The study of protein mechanics with the atomic force microscope. *Trends Biochem Sci,* 24**,** 379-384.

FOUILLEN, A., NEVES, J. D. S., MARY, C., CASTONGUAY, J.-D., MOFFATT, P., BARON, C. & NANCI, A. 2017. Interactions of AMTN, ODAM and

SCPPPQ1 proteins of a specialized basal lamina that attaches epithelial cells to tooth mineral. *Scientific Reports,* 7**,** 46683.
FUKAE, M. & SHIMIZU, M. 1974. Studies on the proteins of developing bovine enamel. *Archives of Oral Biology,* 19**,** 381IN1-386.
FUKAE, M. & TANABE, T. 1998. Degradation of enamel matrix proteins in porcine secretory enamel. *Connective tissue research,* 39**,** 123-129.
FUKAE, M., TANABE, T., IJIRI, H. & SHIMIZU, M. 1980. Studies on porcine enamel proteins: a possible original enamel protein. *Tsurumi shigaku. Tsurumi University dental journal,* 6**,** 87.
FUKAE, M., YAMAMOTO, R., KARAKIDA, T., SHIMODA, S. & TANABE, T. 2007. Micelle Structure of Amelogenin in Porcine Secretory Enamel. *Journal of Dental Research,* 86**,** 758-763.
FUKUDA, T., SANUI, T., TOYODA, K., TANAKA, U., TAKETOMI, T., UCHIUMI, T. & NISHIMURA, F. 2013. Identification of novel amelogenin-binding proteins by proteomics analysis. *PLoS One,* 8**,** e78129.
FUKUMOTO, S., KIBA, T., HALL, B., IEHARA, N., NAKAMURA, T., LONGENECKER, G., KREBSBACH, P. H., NANCI, A., KULKARNI, A. B. & YAMADA, Y. 2004. Ameloblastin is a cell adhesion molecule required for maintaining the differentiation state of ameloblasts. *The Journal of Cell Biology,* 167**,** 973-983.
GABE, C. M., BROOKES, S. J. & KIRKHAM, J. 2017. Preparative SDS PAGE as an Alternative to His-Tag Purification of Recombinant Amelogenin. *Front Physiol,* 8**,** 424.
GADHIA, K., MCDONALD, S., ARKUTU, N. & MALIK, K. 2012. Amelogenesis imperfecta: an introduction. *British Dental Journal,* 212**,** 377.
GALARNEAU, A., PRIMEAU, M., TRUDEAU, L.-E. & MICHNICK, S. W. 2002. β-Lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein–protein interactions. *Nature biotechnology,* 20**,** 619.
GALLON, V., CHEN, L., YANG, X. & MORADIAN-OLDAK, J. J. J. O. S. B. 2013. Localization and quantitative co-localization of enamelin with amelogenin. *Journal of Structural Biology,* 183**,** 239-249.
GARDNER, B. M. & WALTER, P. 2011. Unfolded proteins are Ire1-activating ligands that directly induce the unfolded protein response. *Science,* 333**,** 1891-1894.
GASTEIGER, E., HOOGLAND, C., GATTIKER, A., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. 2005. Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook.* Springer.
GESTRELIUS, S., LYNGSTADAAS, S. & HAMMARSTRÖM, L. 2000. Emdogain–periodontal regeneration based on biomimicry. *Clinical oral investigations,* 4**,** 120-125.
GIBBS, J. 2001. Effective Blocking Procedures: ELISA Technical Bulletin–No. 3. *Corning Incorporated,* 13**,** 2016.
GIBSON, C., GOLUB, E., ABRAMS, W., SHEN, G., DING, W. & ROSENBLOOM, J. 1992. Bovine amelogenin message heterogeneity: alternative splicing and Y-chromosomal gene transcription. *Biochemistry,* 31**,** 8384-8388.
GIBSON, C. W., GOLUB, E., DING, W. D., SHIMOKAWA, H., YOUNG, M., TERMINE, J. & ROSENBLOOM, J. 1991. Identification of the Leucine-Rich Amelogenin Peptide (Lrap) as the Translation Product of an

Alternatively Spliced Transcript. *Biochemical and Biophysical Research Communications,* 174**,** 1306-1312.

GIBSON, C. W., LI, Y., SUGGS, C., KUEHL, M. A., PUGACH, M. K., KULKARNI, A. B. & WRIGHT, J. T. 2011. Rescue of the murine amelogenin null phenotype with two amelogenin transgenes. *European Journal of Oral Sciences,* 119**,** 70-74.

GIBSON, C. W., YUAN, Z.-A., HALL, B., LONGENECKER, G., CHEN, E., THYAGARAJAN, T., SREENATH, T., WRIGHT, J. T., DECKER, S. & PIDDINGTON, R. 2001. Amelogenin-deficient mice display an amelogenesis imperfecta phenotype. *Journal of Biological Chemistry,* 276**,** 31871-31875.

GIBSON, C. W., YUAN, Z. A., LI, Y., DALY, B., SUGGS, C., ARAGON, M. A., ALAWI, F., KULKARNI, A. B. & WRIGHT, J. T. 2007. Transgenic Mice that Express Normal and Mutated Amelogenins. *Journal of Dental Research,* 86**,** 331-335.

GOLDFARB, A. R., SAIDEL, L. J. & MOSOVICH, E. 1951. The ultraviolet absorption spectra of proteins. *Journal of Biological Chemistry,* 193**,** 397-404.

GONZALEZ, T. N., SIDRAUSKI, C., DORFLER, S. & WALTER, P. 1999. Mechanism of non-spliceosomal mRNA splicing in the unfolded protein response pathway. *The EMBO Journal,* 18**,** 3119-32.

GOUTELLE, S., MAURIN, M., ROUGIER, F., BARBAUT, X., BOURGUIGNON, L., DUCHER, M. & MAIRE, P. 2008. The Hill equation: a review of its capabilities in pharmacological modelling. *Fundamental & clinical pharmacology,* 22**,** 633-648.

GREEN, R., DAVIES, M., ROBERTS, C. & TENDLER, S. 1999. Competitive protein adsorption as observed by surface plasmon resonance. *Biomaterials,* 20**,** 385-391.

GREENE, S., YUAN, Z., WRIGHT, J., AMJAD, A., ABRAMS, W., BUCHANAN, J. & GIBSON, G. 2002. A gene deletion resulting in amelogenin with nine cysteine residues leads to amelogenesis imperfecta. *Archives of Oral Biology,* 47**,** 211-217.

GRUBISIC, Z., REMPP, P. & BENOIT, H. 1967. A universal calibration for gel permeation chromatography. *Journal of Polymer Science Part B: Polymer Letters,* 5**,** 753-759.

HANLON, A. D., LARKIN, M. I. & REDDICK, R. M. 2010. Free-solution, label-free protein-protein interactions characterized by dynamic light scattering. *Biophys J,* 98**,** 297-304.

HARADA, H., KETTUNEN, P., JUNG, H.-S., MUSTONEN, T., WANG, Y. A. & THESLEFF, I. 1999. Localization of putative stem cells in dental epithelium and their association with Notch and FGF signaling. *The Journal of Cell Biology,* 147**,** 105-120.

HARDING, M. W., GALAT, A., UEHLING, D. E. & SCHREIBER, S. L. 1989. A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase. *Nature,* 341**,** 758-60.

HARMER, I. & SAMUEL, D. 1989. The FITC-anti-FITC system is a sensitive alternative to biotin-streptavidin in ELISA. *Journal of immunological methods,* 122**,** 115-121.

HART, P., HART, T., MICHALEC, M., RYU, O., SIMMONS, D., HONG, S. & WRIGHT, J. 2004. Mutation in kallikrein 4 causes autosomal recessive

hypomaturation amelogenesis imperfecta. *Journal of Medical Genetics,* 41**,** 545-549.
HART, P. S., ALDRED, M. J., CRAWFORD, P. J. M., WRIGHT, N. J., HART, T. C. & WRIGHT, J. T. 2002. Amelogenesis imperfecta phenotype–genotype correlations with two amelogenin gene mutations. *Archives of Oral Biology,* 47**,** 261-265.
HARTL, F. U. 1996. Molecular chaperones in cellular protein folding. *Nature,* 381**,** 571-9.
HARTL, F. U., BRACHER, A. & HAYER-HARTL, M. 2011. Molecular chaperones in protein folding and proteostasis. *Nature,* 475**,** 324.
HAZE, K., YOSHIDA, H., YANAGI, H., YURA, T. & MORI, K. 1999. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Molecular Biology of the Cell,* 10**,** 3787-99.
HE, L.-H., YIN, Z.-H., JANSEN VAN VUUREN, L., CARTER, E. A. & LIANG, X.-W. 2013. A natural functionally graded biocomposite coating – Human enamel. *Acta Biomaterialia,* 9**,** 6330-6337.
HE, L. H. & SWAIN, M. V. 2008. Understanding the mechanical behaviour of human enamel from its structural and compositional characteristics. *Journal of the Mechanical Behavior of Biomedical Materials,* 1**,** 18-29.
HECKER, M., TING, M. S. H. & MALMSTRÖM, J. 2018. Simple Coatings to Render Polystyrene Protein Resistant. *Coatings,* 8**,** 55.
HELENIUS, A. & AEBI, M. 2001. Intracellular functions of N-linked glycans. *Science,* 291**,** 2364-2369.
HETZ, C. 2012. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nature Reviews Molecular Cell Biology,* 13**,** 89-102.
HETZ, C., CHEVET, E. & HARDING, H. P. 2013. Targeting the unfolded protein response in disease. *Nature Reviews Drug Discovery,* 12**,** 703-719.
HETZ, C. & GLIMCHER, L. H. 2009. Fine-tuning of the unfolded protein response: Assembling the IRE1alpha interactome. *Molecular Cell,* 35**,** 551-61.
HILL, A. V. 1910. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *The Journal of Physiology,* 40**,** 4-7.
HIRSH, S. L., MCKENZIE, D. R., NOSWORTHY, N. J., DENMAN, J. A., SEZERMAN, O. U. & BILEK, M. M. 2013. The Vroman effect: competitive protein exchange with dynamic multilayer protein aggregates. *Colloids and Surfaces B: Biointerfaces,* 103**,** 395-404.
HOCHULI, E., BANNWARTH, W., DÖBELI, H., GENTZ, R. & STÜBER, D. 1988. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Nature biotechnology,* 6**,** 1321-1325.
HOLCROFT, J. & GANSS, B. 2011. Identification of Amelotin-and ODAM-interacting enamel matrix proteins using the yeast two-hybrid system. *European Journal of Oral Sciences,* 119**,** 301-306.
HU, C.-C., FUKAE, M., UCHIDA, T., QIAN, Q., ZHANG, C., RYU, O., TANABE, T., YAMAKOSHI, Y., MURAKAMI, C. & DOHI, N. J. J. O. D. R. 1997. Cloning and characterization of porcine enamelin mRNAs. *Journal of Dental Research,* 76**,** 1720-1729.

HU, C.-D., CHINENOV, Y. & KERPPOLA, T. K. 2002a. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol Cell,* 9**,** 789-798.

HU, J.-C., RYU, O., CHEN, J., UCHIDA, T., WAKIDA, K., MURAKAMI, C., JIANG, H., QIAN, Q., ZHANG, C. & CTTMERS, V. 2000. Localization of EMSP1 expression during tooth formation and cloning of mouse cDNA. *Journal of Dental Research,* 79**,** 70-76.

HU, J. C., CHUN, Y. H., AL HAZZAZZI, T. & SIMMER, J. P. 2007. Enamel formation and amelogenesis imperfecta. *Cells Tissues Organs,* 186**,** 78-85.

HU, J. C., HU, Y., SMITH, C. E., MCKEE, M. D., WRIGHT, J. T., YAMAKOSHI, Y., PAPAGERAKIS, P., HUNTER, G. K., FENG, J. Q., YAMAKOSHI, F. & SIMMER, J. P. 2008. Enamel defects and ameloblast-specific expression in Enam knock-out/lacz knock-in mice. *Journal of Biological Chemistry,* 283**,** 10858-71.

HU, J. C. C., SUN, X., ZHANG, C., LIU, S., BARTLETT, J. D. & SIMMER, J. P. 2002b. Enamelysin and kallikrein-4 mRNA expression in developing mouse molars. *European Journal of Oral Sciences,* 110**,** 307-315.

HU, J. C. C., SUN, X., ZHANG, C. & SIMMER, J. P. 2001. A comparison of enamelin and amelogenin expression in developing mouse molars. *European Journal of Oral Sciences,* 109**,** 125-132.

HU, J. K.-H., MUSHEGYAN, V. & KLEIN, O. D. 2014. On the cutting edge of organ renewal: Identification, regulation, and evolution of incisor stem cells. *Genesis,* 52**,** 79-92.

HU, Y., SMITH, C. E., RICHARDSON, A. S., BARTLETT, J. D., HU, J. C. & SIMMER, J. P. 2016. MMP20, KLK4, and MMP20/KLK4 double null mice define roles for matrix proteases during dental enamel formation. *Molecular Genetics & Genomic Medicine,* 4**,** 178-196.

HUNTER, T. 2014. The genesis of tyrosine phosphorylation. *Cold Spring Harbor perspectives in biology,* 6**,** a020644.

IANNITTI, T. & PALMIERI, B. 2011. Clinical and experimental applications of sodium phenylbutyrate. *Drugs in R & D,* 11**,** 227-249.

IIJIMA, M., FAN, D., BROMLEY, K. M., SUN, Z., MORADIAN-OLDAK, J. J. C. G. & DESIGN 2010. Tooth enamel proteins enamelin and amelogenin cooperate to regulate the growth morphology of octacalcium phosphate crystals. *Crystal Growth & Design,* 10**,** 4815-4822.

IKAWA, T., KAKEGAWA, A., NAGANO, T., ANDO, H., YAMAKOSHI, Y., TANABE, T., SIMMER, J., HU, C.-C., FUKAE, M. & OIDA, S. 2005. Porcine amelogenin is expressed from the X and Y chromosomes. *Journal of Dental Research,* 84**,** 144-148.

INGHAM, K. C. 1984. Protein precipitation with polyethylene glycol. *Methods in Enzymology.* Elsevier.

ITO, T., CHIBA, T., OZAWA, R., YOSHIDA, M., HATTORI, M. & SAKAKI, Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences,* 98**,** 4569-4574.

IWASAKI, K., BAJENOVA, E., SOMOGYI-GANSS, E., MILLER, M., NGUYEN, V., NOURKEYHANI, H., GAO, Y., WENDEL, M. & GANSS, B. 2005. Amelotin—a novel secreted, ameloblast-specific protein. *Journal of Dental Research,* 84**,** 1127-1132.

IWATA, T., YAMAKOSHI, Y., HU, J.-C., ISHIKAWA, I., BARTLETT, J., KREBSBACH, P. & SIMMER, J. 2007. Processing of ameloblastin by MMP-20. *Journal of Dental Research,* 86**,** 153-157.

JAUREGUIBERRY, G., DE LA DURE-MOLLA, M., PARRY, D., QUENTRIC, M., HIMMERKUS, N., KOIKE, T., POULTER, J., KLOOTWIJK, E., ROBINETTE, S. L., HOWIE, A. J., PATEL, V., FIGUERES, M. L., STANESCU, H. C., ISSLER, N., NICHOLSON, J. K., BOCKENHAUER, D., LAING, C., WALSH, S. B., MCCREDIE, D. A., POVEY, S., ASSELIN, A., PICARD, A., COULOMB, A., MEDLAR, A. J., BAILLEUL-FORESTIER, I., VERLOES, A., LE CAIGNEC, C., ROUSSEY, G., GUIOL, J., ISIDOR, B., LOGAN, C., SHORE, R., JOHNSON, C., INGLEHEARN, C., AL-BAHLANI, S., SCHMITTBUHL, M., CLAUSS, F., HUCKERT, M., LAUGEL, V., GINGLINGER, E., PAJAROLA, S., SPARTA, G., BARTHOLDI, D., RAUCH, A., ADDOR, M. C., YAMAGUTI, P. M., SAFATLE, H. P., ACEVEDO, A. C., MARTELLI-JUNIOR, H., DOS SANTOS NETOS, P. E., COLETTA, R. D., GRUESSEL, S., SANDMANN, C., RUEHMANN, D., LANGMAN, C. B., SCHEINMAN, S. J., OZDEMIR-OZENEN, D., HART, T. C., HART, P. S., NEUGEBAUER, U., SCHLATTER, E., HOUILLIER, P., GAHL, W. A., VIKKULA, M., BLOCH-ZUPAN, A., BLEICH, M., KITAGAWA, H., UNWIN, R. J., MIGHELL, A., BERDAL, A. & KLETA, R. 2012. Nephrocalcinosis (enamel renal syndrome) caused by autosomal recessive FAM20A mutations. *Nephron Physiology,* 122**,** 1-6.

JAYASUDHA, B., NAVIN, H. & PRASANNA, K. 2014. Enamel regeneration-current progress and challenges. *Journal of Clinical and Diagnostic Research: JCDR,* 8**,** ZE06.

JEREMIAS, F., KORUYUCU, M., KÜCHLER, E. C., BAYRAM, M., TUNA, E. B., DEELEY, K., PIERRI, R. A., SOUZA, J. F., FRAGELLI, C. M. & PASCHOAL, M. A. 2013. Genes expressed in dental enamel development are associated with molar-incisor hypomineralization. *Archives of oral biology,* 58**,** 1434-1442.

JEYACHANDRAN, Y., MIELCZARSKI, E., RAI, B. & MIELCZARSKI, J. 2009. Quantitative and qualitative evaluation of adsorption/desorption of bovine serum albumin on hydrophilic and hydrophobic surfaces. *Langmuir,* 25**,** 11614-11620.

JI, J. A., ZHANG, B., CHENG, W. & WANG, Y. J. 2009. Methionine, tryptophan, and histidine oxidation in a model protein, PTH: mechanisms and stabilization. *Journal of pharmaceutical sciences,* 98**,** 4485-4500.

JIANG, L. & LAI, L. 2002. CH··· O hydrogen bonds at protein-protein interfaces. *Journal of Biological Chemistry,* 277**,** 37732-37740.

JOHNSON, D. A., GAUTSCH, J. W., SPORTSMAN, J. R. & ELDER, J. H. 1984. Improved technique utilizing nonfat dry milk for analysis of proteins and nucleic acids transferred to nitrocellulose. *Gene Analysis Techniques,* 1**,** 3-8.

JOHNSSON, N. & VARSHAVSKY, A. 1994. Split ubiquitin as a sensor of protein interactions in vivo. *Proceedings of the National Academy of Sciences,* 91**,** 10340-10344.

JONES, S. & THORNTON, J. M. 1996. Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences,* 93**,** 13-20.

JOSEPHSEN, K. & FEJERSKOV, O. 1977. Ameloblast modulation in the maturation zone of the rat incisor enamel organ. A light and electron microscopic study. *Journal of Anatomy,* 124**,** 45.

KALLENBACH, E. 1973. The fine structure of Tomes' process of rat incisor ameloblasts and its relationship to the elaboration of enamel. *Tissue and Cell,* 5**,** 501-524.

KANG, S. W., YOON, I., LEE, H. W. & CHO, J. 2011. Association between AMELX polymorphisms and dental caries in Koreans. *Oral Diseases,* 17**,** 399-406.

KASPER, M., KARSTEN, U., STOSLEK, P. & MOLL, R. 1989. Distribution of intermediate-filament proteins in the human enamel organ: unusually complex pattern of coexpression of cytokeratin polypeptides and vimentin*. *Differentiation,* 40**,** 207-214.

KATSURA, K., HORST, J., CHANDRA, D., LE, T., NAKANO, Y., ZHANG, Y., HORST, O., ZHU, L., LE, M. & DENBESTEN, P. 2014. WDR72 models of structure and function: a stage-specific regulator of enamel mineralization. *Matrix Biology,* 38**,** 48-58.

KENSLER, T. W., WAKABAYASHI, N. & BISWAL, S. 2007. Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway. *Annual Review of Pharmacology and Toxicology,* 47**,** 89-116.

KESTLER, D. P., FOSTER, J. S., MACY, S. D., MURPHY, C. L., WEISS, D. T. & SOLOMON, A. 2008. Expression of odontogenic ameloblast-associated protein (ODAM) in dental and other epithelial neoplasms. *Molecular Medicine,* 14**,** 318.

KIDA, M., SAKIYAMA, Y., MATSUDA, A., TAKABAYASHI, S., OCHI, H., SEKIGUCHI, H., MINAMITAKE, S. & ARIGA, T. 2007. A novel missense mutation (p. P52R) in amelogenin gene causing X-linked amelogenesis imperfecta. *Journal of Dental Research,* 86**,** 69-72.

KIM, J.-W., LEE, S.-K., LEE, Z. H., PARK, J.-C., LEE, K.-E., LEE, M.-H., PARK, J.-T., SEO, B.-M., HU, J. C.-C. & SIMMER, J. P. 2008. FAM83H mutations in families with autosomal-dominant hypocalcified amelogenesis imperfecta. *The American Journal of Human Genetics,* 82**,** 489-494.

KIM, J.-W., SIMMER, J., HU, Y., LIN, B.-L., BOYD, C., WRIGHT, J., YAMADA, C., RAYES, S., FEIGAL, R. & HU, J.-C. 2004. Amelogenin p. M1T and p. W4S mutations underlying hypoplastic X-linked amelogenesis imperfecta. *Journal of Dental Research,* 83**,** 378-383.

KIM, J.-W., SIMMER, J. P., HART, T. C., HART, P. S., RAMASWAMI, M. D., BARTLETT, J. D. & HU, J. C.-C. 2005. MMP-20 mutation in autosomal recessive pigmented hypomaturation amelogenesis imperfecta. *Journal of Medical Genetics,* 42**,** 271-275.

KIM, T. D., RYU, H. J., CHO, H. I., YANG, C.-H. & KIM, J. 2000. Thermal behavior of proteins: heat-resistant proteins and their heat-induced secondary structural changes. *Biochemistry,* 39**,** 14839-14846.

KIM, Y.-J., KIM, Y. J., KANG, J., SHIN, T. J., HYUN, H.-K., LEE, S.-H., LEE, Z. H. & KIM, J.-W. 2017. A novel AMELX mutation causes hypoplastic amelogenesis imperfecta. *Archives of Oral Biology,* 76**,** 61-65.

KINDELAN, S., BROOK, A., GANGEMI, L., LENCH, N., WONG, F., FEARNE, J., JACKSON, Z., FOSTER, G. & STRINGER, B. 2000. Detection of a novel mutation in X-linked amelogenesis imperfecta. *Journal of Dental Research,* 79**,** 1978-1982.

KIRKHAM, J., ANDREEV, I., ROBINSON, C., BROOKES, S. J., SHORE, R. C. & SMITH, D. A. 2006. Evidence for direct amelogenin–target cell interactions using dynamic force spectroscopy. *Eur J Oral Sci,* 114**,** 219-224.

KIRKHAM, J., BROOKES, S. J., DIEKWISCH, T. G. H., MARGOLIS, H. C., BERDAL, A. & HUBBARD, M. J. 2017. Enamel Research: Priorities and Future Directions. *Frontiers in Physiology,* 8.

KIRKHAM, J., BROOKES, S. J., SHORE, R. C., WOOD, S. R., SMITH, D. A., ZHANG, J., CHEN, H. & ROBINSON, C. 2002. Physico-chemical properties of crystal surfaces in matrix–mineral interactions during mammalian biomineralisation. *Current Opinion in Colloid & Interface Science,* 7**,** 124-132.

KIRKHAM, J., FIRTH, A., VERNALS, D., BODEN, N., ROBINSON, C., SHORE, R., BROOKES, S. & AGGELI, A. 2007. Self-assembling peptide scaffolds promote enamel remineralization. *Journal of Dental Research,* 86**,** 426-430.

KNECHT, S., RICKLIN, D., EBERLE, A. N. & ERNST, B. 2009. Oligohis-tags: mechanisms of binding to Ni2+-NTA surfaces. *Journal of Molecular Recognition,* 22**,** 270-279.

KOCH, G., HALLONSTEN, A. L., LUDVIGSSON, N., HANSSON, B. O., HOIST, A. & ULLBRO, C. 1987. Epidemiologic study of idiopathic enamel hypomineralization in permanent teeth of Swedish children. *Community Dentistry and Oral Epidemiology,* 15**,** 279-285.

KÖBLINGER, C., UTTENTHALER, E., DROST, S., ABERL, F., WOLF, H., BRINK, G., STANGLMAIER, A. & SACKMANN, E. 1995. Comparison of the QCM and the SPR method for surface studies and immunological applications. *Sensors and Actuators B: Chemical,* 24**,** 107-112.

KREBSBACH, P. H., LEE, S. K., MATSUKI, Y., KOZAK, C. A., YAMADA, K. M. & YAMADA, Y. 1996. Full-length sequence, localization, and chromosomal mapping of ameloblastin a novel tooth-specific gene. *Journal of Biological Chemistry,* 271**,** 4431-4435.

KUBOTA, K., LEE, D. H., TSUCHIYA, M., YOUNG, C. S., EVERETT, E. T., MARTINEZ-MIER, E. A., SNEAD, M. L., NGUYEN, L., URANO, F. & BARTLETT, J. D. 2005. Fluoride induces endoplasmic reticulum stress in ameloblasts responsible for dental enamel formation. *J Biol Chem,* 280**,** 23194-202.

KÜHNISCH, J., THIERING, E., HEITMÜLLER, D., TIESLER, C. M., GRALLERT, H., HEINRICH-WELTZIEN, R., HICKEL, R., HEINRICH, J., GROUP, G.-P. S. & GROUP, L.-P. S. 2014. Genome-wide association study (GWAS) for molar–incisor hypomineralization (MIH). *Clinical Oral Investigations,* 18**,** 677-682.

KURAMITSU-FUJIMOTO, S., ARIYOSHI, W., SAITO, N., OKINAGA, T., KAMO, M., ISHISAKI, A., TAKATA, T., YAMAGUCHI, K. & NISHIHARA, T. 2015. Novel biological activity of ameloblastin in enamel matrix derivative. *Journal of Applied Oral Science,* 23**,** 49-55.

KWAK, S.-Y., KIM, S., YAMAKOSHI, Y., SIMMER, J. P., BENIASH, E. & MARGOLIS, H. C. 2014. Regulation of calcium phosphate formation by native amelogenins in vitro. *Connective Tissue Research,* 55**,** 21-24.

KWAK, S.-Y., WIEDEMANN-BIDLACK, F. B., BENIASH, E., YAMAKOSHI, Y., SIMMER, J. P., LITMAN, A. & MARGOLIS, H. C. 2009. Role of 20-kDa

amelogenin (P148) phosphorylation in calcium phosphate formation in vitro. *Journal of Biological Chemistry,* 284**,** 18972-18979.

KWAK, S., YAMAKOSHI, Y., SIMMER, J. & MARGOLIS, H. 2016. MMP20 proteolysis of native amelogenin regulates mineralization in vitro. *Journal of Dental Research,* 95**,** 1511-1517.

KWAK, S. Y., GREEN, S., WIEDEMANN-BIDLACK, F. B., BENIASH, E., YAMAKOSHI, Y., SIMMER, J. P. & MARGOLIS, H. C. 2011. Regulation of calcium phosphate formation by amelogenins under physiological conditions. *European Journal of Oral Sciences,* 119 Suppl 1**,** 103-11.

KYTE, J. & DOOLITTLE, R. F. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology,* 157**,** 105-132.

LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227**,** 680-5.

LAGERSTRÖM-FERMÉR, M., NILSSON, M., BÄCKMAN, B., SALIDO, E., SHAPIRO, L., PETTERSSON, U. & LANDEGREN, U. 1995. Amelogenin signal peptide mutation: correlation between mutations in the amelogenin gene (AMGX) and manifestations of X-linked amelogenesis imperfecta. *Genomics,* 26**,** 159-162.

LAGERSTROM-FERMER, M., PETTERSSON, U. & LANDEGREN, U. 1993. Molecular basis and consequences of a deletion in the amelogenin gene, analyzed by capture PCR. *Genomics,* 17**,** 89-92.

LAGERSTROM, M., DAHL, N., NAKAHORI, Y., NAKAGOME, Y., BACKMAN, B., LANDEGREN, U. & PETTERSSON, U. 1991. A deletion in the amelogenin gene (AMG) causes X-linked amelogenesis imperfecta (AIH1). *Genomics,* 10**,** 971-5.

LAKSHMINARAYANAN, R., BROMLEY, K. M., LEI, Y.-P., SNEAD, M. L. & MORADIAN-OLDAK, J. 2010. Perturbed amelogenin secondary structure leads to uncontrolled aggregation in amelogenesis imperfecta mutant proteins. *Journal of Biological chemistry,* 285**,** 40593-40603.

LAKSHMINARAYANAN, R., FAN, D., DU, C. & MORADIAN-OLDAK, J. 2007. The Role of Secondary Structure in the Entropically Driven Amelogenin Self-Assembly. *Biophysical Journal,* 93**,** 3664-3674.

LAKSHMINARAYANAN, R., YOON, I., HEGDE, B. G., FAN, D., DU, C. & MORADIAN-OLDAK, J. 2009. Analysis of secondary structure and self-assembly of amelogenin by variable temperature circular dichroism and isothermal titration calorimetry. *Proteins,* 76**,** 560-9.

LAMBERT, J. E., CHAPMAN, C. A., WRANGHAM, R. W. & CONKLIN-BRITTAIN, N. L. 2004. Hardness of cercopithecine foods: implications for the critical function of enamel thickness in exploiting fallback foods. *American Journal of Physical Anthropology,* 125**,** 363-368.

LANG, K., SCHMID, F. X. & FISCHER, G. 1987. Catalysis of protein folding by prolyl isomerase. *Nature,* 329**,** 268.

LATA, S., REICHEL, A., BROCK, R., TAMPÉ, R. & PIEHLER, J. 2005. High-affinity adaptors for switchable recognition of histidine-tagged proteins. *Journal of the American Chemical Society,* 127**,** 10205-10215.

LAU, E. C., MOHANDAS, T. K., SHAPIRO, L. J., SLAVKIN, H. C. & SNEAD, M. L. 1989. Human and mouse amelogenin gene loci are on the sex chromosomes. *Genomics,* 4**,** 162-168.

LAU, E. C., SIMMER, J. P., BRINGAS, P., JR., HSU, D. D., HU, C. C., ZEICHNER-DAVID, M., THIEMANN, F., SNEAD, M. L., SLAVKIN, H. C.

& FINCHAM, A. G. 1992. Alternative splicing of the mouse amelogenin primary RNA transcript contributes to amelogenin heterogeneity. *Biochemical and Biophysical Research Communications,* 188**,** 1253-60.

LAYNE, E. 1957. Spectrophotometric and turbidometric methods for measuring proteins. *Methods in Enzymology,* 3**,** 447-454.

LE NORCY, E., KWAK, S. Y., WIEDEMANN-BIDLACK, F. B., BENIASH, E., YAMAKOSHI, Y., SIMMER, J. P. & MARGOLIS, H. C. 2011. Leucine-rich Amelogenin Peptides Regulate Mineralization in vitro. *Journal of Dental Research,* 90**,** 1091-1097.

LEE, S. K., KIM, S. M., LEE, Y. J., YAMADA, K. M., YAMADA, Y. & CHI, J. G. 2003. The structure of the rat ameloblastin gene and its expression in amelogenesis. *Molecules & Cells (Springer Science & Business Media BV),* 15.

LEE, S. K., KREBSBACH, P. H., MATSUKI, Y., NANCI, A., YAMADA, K. M. & YAMADA, Y. 1996. Ameloblastin expression in rat incisors and human tooth germs. *International Journal of Developmental Biology,* 40**,** 1141-1150.

LEHMANN, K., HOFFMANN, S., NEUDECKER, P., SUHR, M., BECKER, W.-M. & RÖSCH, P. 2003. High-yield expression in Escherichia coli, purification, and characterization of properly folded major peanut allergen Ara h 2. *Protein Expression and Purification,* 31**,** 250-259.

LENCH, N., BROOK, A. & WINTER, G. 1994. SSCP detection of a nonsense mutation in exon 5 of the amelogenin gene (AMGX) causing X-linked amelogenesis imperfecta (AIH1). *Human Molecular Genetics,* 3**,** 827-828.

LENCH, N. J. & WINTER, G. B. 1995. Characterisation of molecular defects in X-linked amelogenesis imperfecta (AIH1). *Human Mutation,* 5**,** 251-9.

LI, B. & FIELDS, S. 1993. Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system. *The FASEB Journal,* 7**,** 957-963.

LI, W., GIBSON, C. W., ABRAMS, W. R., ANDREWS, D. W. & DENBESTEN, P. K. 2001. Reduced hydrolysis of amelogenin may result in X-linked amelogenesis imperfecta. *Matrix Biology,* 19**,** 755-760.

LI, W., MATHEWS, C., GAO, C. & DENBESTEN, P. 1998. Identification of two additional exons at the 3′ end of the amelogenin gene. *Archives of Oral Biology,* 43**,** 497-504.

LIAO, S.-M., DU, Q.-S., MENG, J.-Z., PANG, Z.-W. & HUANG, R.-B. 2013. The multiple roles of histidine in protein interactions. *Chemistry Central Journal,* 7**,** 44.

LICHTY, J. J., MALECKI, J. L., AGNEW, H. D., MICHELSON-HOROWITZ, D. J. & TAN, S. 2005. Comparison of affinity tags for protein purification. *Protein expression and purification,* 41**,** 98-105.

LIMEBACK, H. 1987. Isolation and characterization of pig enamelins. *Biochemical journal,* 243**,** 385-390.

LIN, J. H., LI, H., YASUMURA, D., COHEN, H. R., ZHANG, C., PANNING, B., SHOKAT, K. M., LAVAIL, M. M. & WALTER, P. 2007. IRE1 signaling affects cell fate during the unfolded protein response. *Science,* 318**,** 944-9.

LINDHOLM, D., WOOTZ, H. & KORHONEN, L. 2006. ER stress and neurodegenerative diseases. *Cell Death & Differentiation,* 13**,** 385-392.

LIPSON, K. L., FONSECA, S. G., ISHIGAKI, S., NGUYEN, L. X., FOSS, E., BORTELL, R., ROSSINI, A. A. & URANO, F. 2006. Regulation of insulin biosynthesis in pancreatic beta cells by an endoplasmic reticulum-resident protein kinase IRE1. *Cell metabolism,* 4**,** 245-254.

LIU, C. Y., SCHRÖDER, M. & KAUFMAN, R. J. 2000. Ligand-independent dimerization activates the stress response kinases IRE1 and PERK in the lumen of the endoplasmic reticulum. *Journal of Biological Chemistry,* 275**,** 24881-24885.

LU, X., ITO, Y., ATSAWASUWAN, P., DANGARIA, S., YAN, X., WU, T., EVANS, C. A. & LUAN, X. J. B. 2013. Ameloblastin modulates osteoclastogenesis through the integrin/ERK pathway. *Bone,* 54**,** 157-168.

LU, X., LI, W., FUKUMOTO, S., YAMADA, Y., EVANS, C. A., DIEKWISCH, T. & LUAN, X. J. M. B. 2016. The ameloblastin extracellular matrix molecule enhances bone fracture resistance and promotes rapid bone fracture healing. *Matrix Biology,* 52**,** 113-126.

MACDOUGALL, M., SIMMONS, D., GU, T. T., FORSMAN-SEMB, K., KÄRRMAN MÅRDH, C., MESBAH, M., FOREST, N., KREBSBACH, P. H., YAMADA, Y. & BERDAL, A. 2000. Cloning, characterization and immunolocalization of human ameloblastin. *European Journal of Oral Sciences,* 108**,** 303-310.

MAEDA, H., ISHIDA, N., KAWAUCHI, H. & TUZIMURA, K. 1969. Reaction of Fluorescein-Isothiocyanate with Proteins and Amino Acids
I. Covalent and Non-Covalent Binding of Fluorescein-Isothiocyanate and Fluorescein to Proteins. *The Journal of Biochemistry,* 65**,** 777-783.

MAJOREK, K. A., KUHN, M. L., CHRUSZCZ, M., ANDERSON, W. F. & MINOR, W. 2014. Double trouble—Buffer selection and H is-tag presence may be responsible for nonreproducibility of biomedical experiments. *Protein Science,* 23**,** 1359-1368.

MARDH, C. K., BÄCKMAN, B., HOLMGREN, G., HU, J. C.-C., SIMMER, J. P. & FORSMAN-SEMB, K. 2002. A nonsense mutation in the enamelin gene causes local hypoplastic autosomal dominant amelogenesis imperfecta (AIH2). *Human Molecular Genetics,* 11**,** 1069-1074.

MARGOLIS, H. C., BENIASH, E. & FOWLER, C. E. 2006. Role of macromolecular assembly of enamel matrix proteins in enamel formation. *Journal of Dental Research,* 85**,** 775-93.

MARX, K. A. 2003. Quartz crystal microbalance: a useful tool for studying thin polymer films and complex biomolecular systems at the solution–surface interface. *Biomacromolecules,* 4**,** 1099-1120.

MATSUO, S., INAI, T., KURISU, K., KIYOMIYA, K.-I. & KUREBE, M. 1996. Influence of fluoride on secretory pathway of the secretory ameloblast in rat incisor tooth germs exposed to sodium fluoride. *Archives of Toxicology,* 70**,** 420-429.

MAYCOCK, J., WOOD, S., BROOKES, S., SHORE, R., ROBINSON, C. & KIRKHAM, J. 2002. Characterization of a Procine Amelogenin Preparation, EMADOGAIN, a Biological Treatment for Periodontal Disease. *Connective Tissue Research,* 43**,** 472-476.

MAZUMDER, P., PRAJAPATI, S., LOKAPPA, S. B., GALLON, V. & MORADIAN-OLDAK, J. 2014. Analysis of co-assembly and co-localization of ameloblastin and amelogenin. *Frontiers in Physiology,* 5**,** 274.

MCDOWALL, F., KENNY, K., MIGHELL, A. J. & BALMER, R. C. 2018. Genetic testing for amelogenesis imperfecta: knowledge and attitudes of paediatric dentists. *British Dental Journal,* 225**,** 335.

MECHANIC, G. L., KATZ, E. P. & GLIMCHER, M. J. 1967. The sephadex gel filtration characteristics of the neutral soluble proteins of embryonic bovine enamel. *Biochimica et Biophysica Acta (BBA) - Protein Structure,* 133**,** 97-113.

MERMUT, O., PHILLIPS, D. C., YORK, R. L., MCCREA, K. R., WARD, R. S. & SOMORJAI, G. A. 2006. In situ adsorption studies of a 14-amino acid leucine-lysine peptide onto hydrophobic polystyrene and hydrophilic silica surfaces using quartz crystal microbalance, atomic force microscopy, and sum frequency generation vibrational spectroscopy. *J Am Chem Soc,* 128**,** 3598-3607.

MÉSZÁROS, B., TOMPA, P., SIMON, I. & DOSZTÁNYI, Z. 2007. Molecular principles of the interactions of disordered proteins. *Journal of Molecular Biology,* 372**,** 549-561.

MIDDELBERG, A. P. J. 1995. Process-scale disruption of microorganisms. *Biotechnology Advances,* 13**,** 491-551.

MOFFATT, P., SMITH, C. E., SOOKNANAN, R., ST-ARNAUD, R. & NANCI, A. 2006a. Identification of secreted and membrane proteins in the rat incisor enamel organ using a signal-trap screening approach. *European Journal of Oral Sciences,* 114**,** 139-146.

MOFFATT, P., SMITH, C. E., ST-ARNAUD, R., SIMMONS, D., WRIGHT, J. T. & NANCI, A. 2006b. Cloning of rat amelotin and localization of the protein to the basal lamina of maturation stage ameloblasts and junctional epithelium. *Biochemical Journal,* 399**,** 37-46.

MOFFATT, P., SMITH, C. E., ST-ARNAUD, R. & NANCI, A. 2008. Characterization of Apin, a secreted protein highly expressed in tooth-associated epithelia. *Journal of Cellular Biochemistry,* 103**,** 941-956.

MOFFATT, P., WAZEN, R. M., NEVES, J. D. S. & NANCI, A. 2014. Characterisation of secretory calcium-binding phosphoprotein-proline-glutamine-rich 1: a novel basal lamina component expressed at cell-tooth interfaces. *Cell and Tissue Research,* 358**,** 843-855.

MORADIAN-OLDAK, J. 2001. Amelogenins: assembly, processing and control of crystal morphology. *Matrix Biology,* 20**,** 293-305.

MORADIAN-OLDAK, J. 2009. The regeneration of tooth enamel. *Dimensions of Dental Hygiene,* 7**,** 12-15.

MORADIAN-OLDAK, J., BOUROPOULOS, N., WANG, L. & GHARAKHANIAN, N. 2002. Analysis of self-assembly and apatite binding properties of amelogenin proteins lacking the hydrophilic C-terminal. *Matrix Biology,* 21**,** 197-205.

MORADIAN-OLDAK, J., LEUNG, W. & FINCHAM, A. 1998. Temperature and pH-dependent supramolecular self-assembly of amelogenin molecules: a dynamic light-scattering analysis. *Journal of Structural Biology,* 122**,** 320-327.

MORADIAN-OLDAK, J., PAINE, M., LEI, Y., FINCHAM, A. & SNEAD, M. 2000. Self-assembly properties of recombinant engineered amelogenin proteins analyzed by dynamic light scattering and atomic force microscopy. *Journal of Structural Biology,* 131**,** 27-37.

MORADIAN-OLDAK, J., SIMMER, J. P., LAU, E. C., DIEKWISCH, T., SLAVKIN, H. C. & FINCHAM, A. G. 1995. A review of the aggregation

properties of a recombinant amelogenin. *Connect Tissue Res,* 32**,** 125-30.

MORADIAN-OLDAK, J., SIMMER, J. P., LAU, E. C., SARTE, P. E., SLAVKIN, H. C. & FINCHAM, A. G. 1994. Detection of monodisperse aggregates of a recombinant amelogenin by dynamic light scattering. *Biopolymers,* 34**,** 1339-1347.

MORADIAN-OLDAK, J., DU, C. & FALINI, G. 2006. On the formation of amelogenin microribbons. *European Journal of Oral Sciences,* 114**,** 289-296.

MORASSO, M. I., GRINBERG, A., ROBINSON, G., SARGENT, T. D. & MAHON, K. A. 1999. Placental failure in mice lacking the homeobox gene Dlx3. *Proceedings of the National Academy of Sciences,* 96**,** 162-167.

MORRISSEY, J. H. 1981. Silver stain for proteins in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity. *Anal Biochem,* 117**,** 307-310.

MUKHERJEE, K., RUAN, Q., NUTT, S., TAO, J., DE YOREO, J. J. & MORADIAN-OLDAK, J. 2018. Peptide-Based Bioinspired Approach to Regrowing Multilayered Aprismatic Enamel. *ACS Omega,* 3**,** 2546-2557.

NAGANO, T., IWATA, T., OGATA, Y., TANABE, T., GOMI, K., FUKAE, M., ARAI, T. & OIDA, S. 2004. Effect of heat treatment on bioactivities of enamel matrix derivatives in human periodontal ligament (HPDL) cells. *Journal of periodontal research,* 39**,** 249-256.

NAGANO, T., KAKEGAWA, A., YAMAKOSHI, Y., TSUCHIYA, S., HU, J. C., GOMI, K., ARAI, T., BARTLETT, J. D. & SIMMER, J. P. 2009. Mmp-20 and Klk4 cleavage site preferences for amelogenin sequences. *Journal of Dental Research,* 88**,** 823-8.

NANCI, A., BENDAYAN, M. & SLAVKIN, H. C. 1985. Enamel protein biosynthesis and secretion in mouse incisor secretory ameloblasts as revealed by high-resolution immunocytochemistry. *Journal of Histochemistry & Cytochemistry,* 33**,** 1153-1160.

NANCI, A. & TEN CATE, A. R. 2013. *Ten Cate's oral histology : development, structure, and function,* St. Louis, Mo., Elsevier.

NANCI, A. & WARSHAWSKY, H. 1984. Characterization of putative secretory sites on ameloblasts of the rat incisor. *American Journal of Anatomy,* 171**,** 163-189.

NANCI, A., ZALZAL, S., LAVOIE, P., KUNIKATA, M., CHEN, W.-Y., KREBSBACH, P., YAMADA, Y., HAMMARSTRÖM, L., SIMMER, J. & FINCHAM, A. 1998. Comparative immunochemical analyses of the developmental expression and distribution of ameloblastin and amelogenin in rat incisors. *Journal of Histochemistry & Cytochemistry,* 46**,** 911-934.

NIEBA, L., NIEBA-AXMANN, S. E., PERSSON, A., HÄMÄLÄINEN, M., EDEBRATT, F., HANSSON, A., LIDHOLM, J., MAGNUSSON, K., KARLSSON, Å. F. & PLÜCKTHUN, A. 1997. BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. *Analytical biochemistry,* 252**,** 217-228.

NIKIFORUK, G. & SIMMONS, N. 1965. Purification and properties of protein from embryonic bovine enamel. *Journal of Dental Research,* 44**,** 1119-1122.

NORDE, W., GONZALEZ, F. G. & HAYNES, C. A. 1995. Protein adsorption on polystyrene latex particles. *Polymers for advanced technologies,* 6**,** 518-525.

NÚÑEZ, S. M., CHUN, Y.-H. P., GANSS, B., HU, Y., RICHARDSON, A. S., SCHMITZ, J. E., FAJARDO, R., YANG, J., HU, J. C.-C. & SIMMER, J. P. 2016. Maturation stage enamel malformations in Amtn and Klk4 null mice. *Matrix Biology,* 52**,** 219-233.

NYLEN, M. U., EANES, E. D. & OMNELL, K. A. 1963. Crystal growth in rat enamel. *Journal of Cell Biology,* 18**,** 109-23.

OGATA, M., HINO, S., SAITO, A., MORIKAWA, K., KONDO, S., KANEMOTO, S., MURAKAMI, T., TANIGUCHI, M., TANII, I., YOSHINAGA, K., SHIOSAKA, S., HAMMARBACK, J. A., URANO, F. & IMAIZUMI, K. 2006. Autophagy is activated for cell survival after endoplasmic reticulum stress. *Molecular and Cellular Biology,* 26**,** 9220-31.

ONYIRIUKA, E. C., HERSCH, L. S. & HERTL, W. 1991. Solubilization of corona discharge-and plasma-treated polystyrene. *Journal of colloid and interface science,* 144**,** 98-102.

OU, W.-J., CAMERON, P. H., THOMAS, D. Y. & BERGERON, J. J. M. 1993. Association of folding intermediates of glycoproteins with calnexin during protein maturation. *Nature,* 364**,** 771.

OZDEMIR, D., HART, P. S., FIRATLI, E., AREN, G., RYU, O. H. & HART, T. C. 2005. Phenotype of ENAM Mutations is Dosage-dependent. *Journal of Dental Research,* 84**,** 1036-1041.

PAINE, M. L., KREBSBACH, P. H., CHEN, L. S., PAINE, C. T., YAMADA, Y., DEUTSCH, D. & SNEAD, M. L. 1998. Protein-to-Protein Interactions: Criteria Defining the Assembly of the Enamel Organic Matrix. *Journal of Dental Research,* 77**,** 496-502.

PAINE, M. L., LEI, Y.-P., DICKERSON, K. & SNEAD, M. L. 2002. Altered amelogenin self-assembly based on mutations observed in human X-linked amelogenesis imperfecta (AIH1). *Journal of Biological Chemistry,* 277**,** 17112-17116.

PAINE, M. L. & SNEAD, M. L. 1997. Protein interactions during assembly of the enamel organic extracellular matrix. *Journal of Bone and Mineral Research,* 12**,** 221-227.

PAINE, M. L., WANG, H.-J., LUO, W., KREBSBACH, P. H. & SNEAD, M. L. 2003a. A transgenic animal model resembling amelogenesis imperfecta related to ameloblastin overexpression. *Journal of Biological Chemistry,* 278**,** 19447-19452.

PAINE, M. L., WANG, H.-J. & SNEAD, M. L. 2003b. Amelogenin self-assembly and the role of the proline located within the carboxyl-teleopeptide. *Connective Tissue Research,* 44**,** 52-57.

PAINE, M. L., ZHU, D. H., LUO, W., BRINGAS, P., JR., GOLDBERG, M., WHITE, S. N., LEI, Y. P., SARIKAYA, M., FONG, H. K. & SNEAD, M. L. 2000. Enamel biomineralization defects result from alterations to amelogenin self-assembly. *Journal of Structural Biology,* 132**,** 191-200.

PALMER, L. C., NEWCOMB, C. J., KALTZ, S. R., SPOERKE, E. D. & STUPP, S. I. 2008. Biomimetic systems for hydroxyapatite mineralization inspired by bone and enamel. *Chemical Reviews,* 108**,** 4754-83.

PAPAGERAKIS, P., IBARRA, J., INOZENTSEVA, N., DENBESTEN, P. & MACDOUGALL, M. 2005. Mouse amelogenin exons 8 and 9: sequence

analysis and protein distribution. *Journal of Dental Research,* 84**,** 613-617.

PARK, J. C., PARK, J. T., SON, H. H., KIM, H. J., JEONG, M. J., LEE, C. S., DEY, R. & CHO, M. L. 2007. The amyloid protein APin is highly expressed during enamel mineralization and maturation in rat incisors. *European Journal of Oral Sciences,* 115**,** 153-160.

PARK, S., QUINN, J. B., ROMBERG, E. & AROLA, D. 2008. On the brittleness of enamel and selected dental materials. *Dental Materials,* 24**,** 1477-85.

PARRY, D. A., BROOKES, S. J., LOGAN, C. V., POULTER, J. A., EL-SAYED, W., AL-BAHLANI, S., AL HARASI, S., SAYED, J., RAIF EL, M., SHORE, R. C., DASHASH, M., BARRON, M., MORGAN, J. E., CARR, I. M., TAYLOR, G. R., JOHNSON, C. A., ALDRED, M. J., DIXON, M. J., WRIGHT, J. T., KIRKHAM, J., INGLEHEARN, C. F. & MIGHELL, A. J. 2012. Mutations in C4orf26, encoding a peptide with in vitro hydroxyapatite crystal nucleation and growth activity, cause amelogenesis imperfecta. *American Journal of Human Genetics,* 91**,** 565-71.

PARRY, D. A., POULTER, J. A., LOGAN, C. V., BROOKES, S. J., JAFRI, H., FERGUSON, C. H., ANWARI, B. M., RASHID, Y., ZHAO, H. & JOHNSON, C. A. 2013. Identification of mutations in SLC24A4, encoding a potassium-dependent sodium/calcium exchanger, as a cause of amelogenesis imperfecta. *The American Journal of Human Genetics,* 92**,** 307-312.

PARRY, D. A., SMITH, C. E., EL-SAYED, W., POULTER, J. A., SHORE, R. C., LOGAN, C. V., MOGI, C., SATO, K., OKAJIMA, F. & HARADA, A. 2016. Mutations in the pH-sensing G-protein-coupled receptor GPR68 cause amelogenesis imperfecta. *The American Journal of Human Genetics,* 99**,** 984-990.

PHAM, C.-D., SMITH, C. E., HU, Y., HU, J. C., SIMMER, J. P. & CHUN, Y.-H. P. 2017. Endocytosis and Enamel Formation. *Frontiers in Physiology,* 8**,** 529.

PIEHLER, J. 2005. New methodologies for measuring protein interactions in vivo and in vitro. *Current opinion in structural biology,* 15**,** 4-14.

POULTER, J. A., MURILLO, G., BROOKES, S. J., SMITH, C. E., PARRY, D. A., SILVA, S., KIRKHAM, J., INGLEHEARN, C. F. & MIGHELL, A. J. 2014. Deletion of ameloblastin exon 6 is associated with amelogenesis imperfecta. *Human Molecular Genetics,* 23**,** 5317-24.

PRASAD, M. K., GEOFFROY, V., VICAIRE, S., JOST, B., DUMAS, M., LE GRAS, S., SWITALA, M., GASSE, B., LAUGEL-HAUSHALTER, V., PASCHAKI, M., LEHEUP, B., DROZ, D., DALSTEIN, A., LOING, A., GROLLEMUND, B., MULLER-BOLLA, M., LOPEZ-CAZAUX, S., MINOUX, M., JUNG, S., OBRY, F., VOGT, V., DAVIDEAU, J.-L., DAVIT-BEAL, T., KAISER, A.-S., MOOG, U., RICHARD, B., MORRIER, J.-J., DUPREZ, J.-P., ODENT, S., BAILLEUL-FORESTIER, I., ROUSSET, M. M., MERAMETDIJAN, L., TOUTAIN, A., JOSEPH, C., GIULIANO, F., DAHLET, J.-C., COURVAL, A., EL ALLOUSSI, M., LAOUINA, S., SOSKIN, S., GUFFON, N., DIEUX, A., DORAY, B., FEIERABEND, S., GINGLINGER, E., FOURNIER, B., DE LA DURE MOLLA, M., ALEMBIK, Y., TARDIEU, C., CLAUSS, F., BERDAL, A., STOETZEL, C., MANIÈRE, M. C., DOLLFUS, H. & BLOCH-ZUPAN, A. 2016. A targeted next-generation sequencing assay for the molecular diagnosis of genetic

disorders with orodental involvement. *Journal of Medical Genetics,* 53**,** 98-110.

QIANG, X., SUN, K., XING, L., XU, Y., WANG, H., ZHOU, Z., ZHANG, J., ZHANG, F., CALISKAN, B. & WANG, M. 2017. Discovery of a polystyrene binding peptide isolated from phage display library and its application in peptide immobilization. *Scientific Reports,* 7**,** 2673.

QING, G., MA, L.-C., KHORCHID, A., SWAPNA, G. V. T., MAL, T. K., TAKAYAMA, M. M., XIA, B., PHADTARE, S., KE, H., ACTON, T., MONTELIONE, G. T., IKURA, M. & INOUYE, M. 2004. Cold-shock induced high-yield protein production in Escherichia coli. *Nature Biotechnology,* 22**,** 877.

QU, Q., HAITINA, T., ZHU, M. & AHLBERG, P. E. 2015. New genomic and fossil data illuminate the origin of enamel. *Nature,* 526**,** 108.

RAJPAR, M. H., HARLEY, K., LAING, C., DAVIES, R. M. & DIXON, M. J. 2001. Mutation of the gene encoding the enamel-specific protein, enamelin, causes autosomal-dominant amelogenesis imperfecta. *Human Molecular Genetics,* 10**,** 1673-1677.

RAVINDRANATH, H. H., CHEN, L.-S., ZEICHNER-DAVID, M., ISHIMA, R. & RAVINDRANATH, R. M. H. 2004. Interaction between the enamel matrix proteins amelogenin and ameloblastin. *Biochemical and Biophysical Research Communications,* 323**,** 1075-1083.

RAVINDRANATH, R. M., BASILROSE, R. M., RAVINDRANATH, N. H. & VAITHEESVARAN, B. 2003. Amelogenin interacts with cytokeratin-5 in ameloblasts during enamel growth. *Journal of Biological Chemistry,* 278**,** 20293-20302.

RAVINDRANATH, R. M., MORADIAN-OLDAK, J. & FINCHAM, A. G. 1999. Tyrosyl Motif in Amelogenins BindsN-Acetyl-d-glucosamine. *Journal of Biological Chemistry,* 274**,** 2464-2471.

RAVINDRANATH, R. M., TAM, W.-Y., BRINGAS, P., SANTOS, V. & FINCHAM, A. G. 2001. Amelogenin-cytokeratin 14 interaction in ameloblasts during enamel formation. *Journal of Biological Chemistry,* 276**,** 36586-36597.

RAVINDRANATH, R. M., TAM, W.-Y., NGUYEN, P. & FINCHAM, A. G. 2000. The enamel protein amelogenin binds to the N-acetyl-D-glucosamine-mimicking peptide motif of cytokeratins. *Journal of Biological Chemistry,* 275**,** 39654-39661.

REIMHULT, K., PETERSSON, K. & KROZER, A. 2008. QCM-D analysis of the performance of blocking agents on gold and polystyrene surfaces. *Langmuir,* 24**,** 8695-8700.

REITH, E. J. 1961. The ultrastructure of ameloblasts during matrix formation and the maturation of enamel. *The Journal of Cell Biology,* 9**,** 825-839.

REITH, E. J. 1967. The early stage of amelogenesis as observed in molar teeth of young rats. *Journal of Ultrastructure Research,* 17**,** 503-526.

REITH, E. J. 1970. The stages of amelogenesis as observed in molar teeth of young rats. *Journal of Ultrastructure Research,* 30**,** 111-151.

REITH, E. J. & BOYDE, A. 1981. The arrangement of ameloblasts on the surface of maturing enamel of the rat incisor tooth. *Journal of Anatomy,* 133**,** 381.

RENUGOPALAKRISHNAN, V., PRABHAKARAN, M., HUANG, S.-G., BALASUBRAMANIAM, A., STRAWICH, E. & GLIMCHER, M. 1989.

Secondary structure and limited three-dimensional structure of bovine amelogenin. *Connective Tissue Research,* 22**,** 757-764.

RENUGOPALAKRISHNAN, V., STRAWICH, E., HOROWITZ, P. & GLIMCHER, M. J. 1986. Studies of the secondary structures of amelogenin from bovine tooth enamel. *Biochemistry,* 25**,** 4879-4887.

REYNOLDS, J. A. & TANFORD, C. 1970a. Binding of dodecyl sulfate to proteins at high binding ratios. Possible implications for the state of proteins in biological membranes. *Proceedings of the National Academy of Sciences,* 66**,** 1002-1007.

REYNOLDS, J. A. & TANFORD, C. 1970b. The gross conformation of protein-sodium dodecyl sulfate complexes. *Journal of Biological Chemistry,* 245**,** 5161-5165.

RIDDICK, J. A., BUNGER, W. B. & SAKANO, T. K. 1986. Organic solvents: physical properties and methods of purification.

ROBINSON, C., BROOKES, S. J., SHORE, R. C. & KIRKHAM, J. 1998. The developing enamel matrix: nature and function. *European Journal of Oral Sciences,* 106 Suppl 1**,** 282-91.

ROBINSON, C., FUCHS, P. & WEATHERELL, J. A. 1981. The appearance of developing rat incisor enamel using a freeze fracturing technique. *Journal of Crystal Growth,* 53**,** 160-165.

ROBINSON, C., HILLER, C. R. & WEATHERELL, J. A. 1974. Uptake of32P-labelled phosphate into developing rat incisor enamel. *Calcified Tissue Research,* 15**,** 143-152.

RON, D. & WALTER, P. 2007. Signal integration in the endoplasmic reticulum unfolded protein response. *Nature Reviews Molecular Cell Biology,* 8**,** 519-29.

ROTH, C. M., NEAL, B. L. & LENHOFF, A. M. 1996. Van der Waals interactions involving proteins. *Biophys J,* 70**,** 977-987.

RUAN, Q. & MORADIAN-OLDAK, J. 2015. Amelogenin and Enamel Biomimetics. *Journal of materials chemistry. B, Materials for biology and medicine,* 3**,** 3112-3129.

RUBENSTEIN, R. C., EGAN, M. E. & ZEITLIN, P. L. 1997. In vitro pharmacologic restoration of CFTR-mediated chloride transport with sodium 4-phenylbutyrate in cystic fibrosis epithelial cells containing delta F508-CFTR. *The Journal of clinical investigation,* 100**,** 2457-2465.

RYU, O., HU, J. C. C., YAMAKOSHI, Y., VILLEMAIN, J. L., CAO, X., ZHANG, C., BARTLETT, J. D. & SIMMER, J. P. 2002. Porcine kallikrein-4 activation, glycosylation, activity, and expression in prokaryotic and eukaryotic hosts. *European Journal of Oral Sciences,* 110**,** 358-365.

RYU, O. H., FINCHAM, A. G., HU, C. C., ZHANG, C., QIAN, Q., BARTLETT, J. D. & SIMMER, J. P. 1999. Characterization of recombinant pig enamelysin activity and cleavage of recombinant pig and mouse amelogenins. *Journal of Dental Research,* 78**,** 743-50.

SALIDO, E. C., YEN, P., KOPRIVNIKAR, K., YU, L. & SHAPIRO, L. 1992. The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. *American Journal of Human Genetics,* 50**,** 303.

SASAKI, T., SEGAWA, K., TAKIGUCHI, R. & HIGASHI, S. 1984. Intercellular junctions in the cells of the human enamel organ as revealed by freeze-fracture. *Archives of Oral Biology,* 29**,** 275-286.

SCHEUNER, D. & KAUFMAN, R. J. 2008. The unfolded protein response: a pathway that links insulin demand with β-cell failure and diabetes. *Endocrine reviews,* 29**,** 317-333.

SCHEUNER, D., SONG, B., MCEWEN, E., LIU, C., LAYBUTT, R., GILLESPIE, P., SAUNDERS, T., BONNER-WEIR, S. & KAUFMAN, R. J. 2001. Translational control is required for the unfolded protein response and in vivo glucose homeostasis. *Molecular Cell,* 7**,** 1165-76.

SCHUBERT, U., ANTÓN, L. C., GIBBS, J., NORBURY, C. C., YEWDELL, J. W. & BENNINK, J. R. 2000. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature,* 404**,** 770-774.

SEKIGUSHI, H. 2001. A new mutation in the amelogenin gene causes X-linked amelogenesis imperfecta. *Journal of Dental Research,* 80**,** 617.

SEYMEN, F., KIM, Y. J., LEE, Y. J., KANG, J., KIM, T.-H., CHOI, H., KORUYUCU, M., KASIMOGLU, Y., TUNA, E. B. & GENCAY, K. 2016. Recessive mutations in ACPT, encoding testicular acid phosphatase, cause hypoplastic amelogenesis imperfecta. *The American Journal of Human Genetics,* 99**,** 1199-1205.

SHAMU, C. E. & WALTER, P. 1996. Oligomerization and phosphorylation of the Ire1p kinase during intracellular signaling from the endoplasmic reticulum to the nucleus. *The EMBO Journal,* 15**,** 3028-39.

SHAW, W. J., CAMPBELL, A. A., PAINE, M. L. & SNEAD, M. L. 2004. The COOH terminus of the amelogenin, LRAP, is oriented next to the hydroxyapatite surface. *Journal of Biological Chemistry,* 279**,** 40263-40266.

SHAW, W. J., FERRIS, K., TARASEVICH, B. & LARSON, J. L. 2008. The Structure and Orientation of the C-Terminus of LRAP. *Biophysical Journal,* 94**,** 3247-3257.

SHEINERMAN, F. B., NOREL, R. & HONIG, B. 2000. Electrostatic aspects of protein–protein interactions. *Current opinion in structural biology,* 10**,** 153-159.

SHEN, J., CHEN, X., HENDERSHOT, L. & PRYWES, R. 2002. ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of Golgi localization signals. *Developmental Cell,* 3**,** 99-111.

SHEVCHENKO, A., WILM, M., VORM, O. & MANN, M. 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical chemistry,* 68**,** 850-858.

SHIKHMAN, A. R., GREENSPAN, N. S. & CUNNINGHAM, M. W. 1994. Cytokeratin peptide SFGSGFGGGY mimics N-acetyl-beta-D-glucosamine in reaction with antibodies and lectins, and induces in vivo anti-carbohydrate antibody response. *Journal of Immunology,* 153**,** 5593-606.

SHIMIZU, D. & MACHO, G. A. 2007. Functional significance of the microstructural detail of the primate dentino-enamel junction: a possible example of exaptation. *Journal of Human Evolution,* 52**,** 103-11.

SHIMIZU M, F. M. 1983. Enamel Proteins. *In:* S, S. (ed.) *Mechanisms of Tooth Enamel Formation.* Quintessence Publishing Co Ltd.

SHIMOKAWA, H., SOBEL, M. E., SASAKI, M., TERMINE, J. D. & YOUNG, M. F. 1987. Heterogeneity of amelogenin mRNA in the bovine tooth germ. *Journal of Biological Chemistry,* 262**,** 4042-7.

SHOEMAKER, B. A. & PANCHENKO, A. R. 2007. Deciphering protein–protein interactions. Part I. Experimental techniques and databases. *PLoS computational biology,* 3**,** e42.

SIMMELINK, J. 1982. Mode of enamel matrix secretion. *Journal of Dental Research***,** 1483-1489.

SIMMER, J. & FINCHAM, A. 1995. Molecular mechanisms of dental enamel formation. *Critical Reviews in Oral Biology & Medicine,* 6**,** 84-108.

SIMMER, J., LAU, E., HU, C., AOBA, T., LACEY, M., NELSON, D., ZEICHNER-DAVID, M., SNEAD, M., SLAVKIN, H. & FINCHAM, A. 1994. Isolation and characterization of a mouse amelogenin expressed in Escherichia coli. *Calcified tissue international,* 54**,** 312-319.

SIMMER, J. P., HU, Y., LERTLAM, R., YAMAKOSHI, Y. & HU, J. C.-C. 2009. Hypomaturation enamel defects in Klk4 knockout/LacZ knockin mice. *Journal of Biological Chemistry,* 284**,** 19110-19121.

SIRE, J.-Y., DAVIT-BÉAL, T., DELGADO, S. & GU, X. 2007. The origin and evolution of enamel mineralization genes. *Cells Tissues Organs,* 186**,** 25-48.

SIRE, J.-Y., HUANG, Y., LI, W., DELGADO, S., GOLDBERG, M. & DENBESTEN, P. K. 2012. Evolutionary Story of Mammalian-specific Amelogenin Exons 4, “4b”, 8, and 9. *Journal of Dental Research,* 91**,** 84-89.

SIVASHANMUGAM, A., MURRAY, V., CUI, C., ZHANG, Y., WANG, J. & LI, Q. 2009. Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein Science : A Publication of the Protein Society,* 18**,** 936-948.

SMITH, C. 1998. Cellular and chemical events during enamel maturation. *Critical Reviews in Oral Biology & Medicine,* 9**,** 128-161.

SMITH, C., POMPURA, J., BORENSTEIN, S., FAZEL, A. & NANCI, A. 1989. Degradation and loss of matrix proteins from developing enamel. *The Anatomical Record,* 224**,** 292-316.

SMITH, C. E. & NANCI, A. 2003. Overview of morphological changes in enamel organ cells associated with major events in amelogenesis. *International Journal of Developmental Biology,* 39**,** 153-161.

SMITH, C. E., WAZEN, R., HU, Y., ZALZAL, S. F., NANCI, A., SIMMER, J. P. & HU, J. C. C. 2009. Consequences for enamel development and mineralization resulting from loss of function of ameloblastin or enamelin. *European Journal of Oral Sciences,* 117**,** 485-497.

SMITH, C. E., WHITEHOUSE, L. L., POULTER, J. A., BROOKES, S. J., DAY, P. F., SOLDANI, F., KIRKHAM, J., INGLEHEARN, C. F. & MIGHELL, A. J. 2017a. Defects in the acid phosphatase ACPT cause recessive hypoplastic amelogenesis imperfecta. *European Journal of Human Genetics,* 25**,** 1015.

SMITH, C. E. L., MURILLO, G., BROOKES, S. J., POULTER, J. A., SILVA, S., KIRKHAM, J., INGLEHEARN, C. F. & MIGHELL, A. J. 2016. Deletion of amelotin exons 3–6 is associated with amelogenesis imperfecta. *Human Molecular Genetics,* 25**,** 3578-3587.

SMITH, C. E. L., POULTER, J. A., ANTANAVICIUTE, A., KIRKHAM, J., BROOKES, S. J., INGLEHEARN, C. F. & MIGHELL, A. J. 2017b. Amelogenesis Imperfecta; Genes, Proteins, and Pathways. *Frontiers in Physiology,* 8.

SNEAD, M. L., LAU, E. C., ZEICHNER-DAVID, M., FINCHAM, A. G., WOO, S. L. & SLAVKIN, H. C. 1985. DNA sequence for cloned cDNA for murine amelogenin reveal the amino acid sequence for enamel-specific protein. *Biochemical and Biophysical Research Communications,* 129**,** 812-818.

SOMOGYI-GANSS, E., NAKAYAMA, Y., IWASAKI, K., NAKANO, Y., STOLF, D., MCKEE, M. D. & GANSS, B. 2012. Comparative temporospatial expression profiling of murine amelotin protein during amelogenesis. *Cells Tissues Organs,* 195**,** 535-549.

STEPHAN, A. B., TOBOCHNIK, S., DIBATTISTA, M., WALL, C. M., REISERT, J. & ZHAO, H. 2012. The Na+/Ca 2+ exchanger NCKX4 governs termination and adaptation of the mammalian olfactory response. *Nature Neuroscience,* 15**,** 131.

STRANGE, R. W., BLACKBURN, N. J., KNOWLES, P. & HASNAIN, S. S. 1987. X-ray absorption spectroscopy of metal-histidine coordination in metalloproteins. Exact simulation of the EXAFS of tetrabis (imidazole) copper (II) nitrate and other copper-imidazole complexes by the use of a multiple-scattering treatment. *Journal of the American Chemical Society,* 109**,** 7157-7162.

STRYER, L. 1978. Fluorescence energy transfer as a spectroscopic ruler. *Annu Rev Biochem,* 47**,** 819-846.

SU, J., CHANDRABABU, K. B. & MORADIAN-OLDAK, J. 2016. Ameloblastin peptide encoded by exon 5 interacts with amelogenin N-terminus. *Biochemistry and Biophysics Reports,* 7**,** 26-32.

SUN, G. & ANDERSON, V. E. 2004. Prevention of artifactual protein oxidation generated during sodium dodecyl sulfate-gel electrophoresis. *Electrophoresis,* 25**,** 959-965.

SUZUKI, M., EVERETT, E. T., WHITFORD, G. M. & BARTLETT, J. D. 2017. 4-phenylbutyrate mitigates fluoride-induced cytotoxicity in ALC cells. *Front Physiol,* 8**,** 302.

SVENSSON BONDE, J. & BULOW, L. 2012. One-step purification of recombinant human amelogenin and use of amelogenin as a fusion partner. *PloS one,* 7**,** e33269.

SVENSSON, J., ANDERSSON, C., RESELAND, J. E., LYNGSTADAAS, P. & BÜLOW, L. 2006. Histidine tag fusion increases expression levels of active recombinant amelogenin in Escherichia coli. *Protein expression and purification,* 48**,** 134-141.

TABAS, I. & RON, D. 2011. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nature Cell Biology,* 13**,** 184-90.

TAKAGI, T., SUZUKI, M., BABA, T., MINEGISHI, K. & SASAKI, S. 1984. Complete amino acid sequence of amelogenin in developing bovine enamel. *Biochemical and Biophysical Research Communications,* 121**,** 592-597.

TAKANO, Y., OZAWA, H. & CRENSHAW, M. 1986. Ca-ATPase and ALPase activities at the initial calcification sites of dentin and enamel in the rat incisor. *Cell and Tissue Research,* 243**,** 91-99.

TAMBURSTUEN, M. V., REPPE, S., SPAHR, A., SABETRASEKH, R., KVALHEIM, G., SLABY, I., SYVERSEN, U., LYNGSTADAAS, S. P. & RESELAND, J. E. J. E. J. O. O. S. 2010. Ameloblastin promotes bone growth by enhancing proliferation of progenitor cells and by stimulating immunoregulators. *European Journal of Oral Sciences,* 118**,** 451-459.

TAN, J., LEUNG, W., MORADIAN-OLDAK, J., ZEICHNER-DAVID, M. & FINCHAM, A. 1998a. Quantitative analysis of amelogenin solubility. *Journal of Dental Research,* 77**,** 1388-1396.

TAN, J., LEUNG, W., MORADIAN-OLDAK, J., ZEICHNER-DAVID, M. & FINCHAM, A. G. 1998b. The pH dependent amelogenin solubility and its biological significance. *Connect Tissue Res,* 38**,** 215-21; discussion 241-6.

TANABE, T. 1984. Purification and characterization of proteolytic enzymes in porcine immature enamel. *Tsurumi shigaku. Tsurumi University dental journal,* 10**,** 443-452.

TANIMOTO, K., LE, T., ZHU, L., WITKOWSKA, H. E., ROBINSON, S., HALL, S., HWANG, P., DENBESTEN, P. & LI, W. 2008. Reduced Amelogenin-MMP20 Interactions in Amelogenesis Imperfecta. *Journal of Dental Research,* 87**,** 451-455.

TAO, J., BUCHKO, G. W., SHAW, W. J., DE YOREO, J. J. & TARASEVICH, B. J. 2015. Sequence-Defined Energetic Shifts Control the Disassembly Kinetics and Microstructure of Amelogenin Adsorbed onto Hydroxyapatite (100). *Langmuir*.

TARASEVICH, B. J., HOWARD, C. J., LARSON, J. L., SNEAD, M. L., SIMMER, J. P., PAINE, M. & SHAW, W. J. 2007. The nucleation and growth of calcium phosphate by amelogenin. *Journal of Crystal Growth,* 304**,** 407-415.

TARASEVICH, B. J., LEA, S., BERNT, W., ENGELHARD, M. & SHAW, W. J. 2009a. Adsorption of amelogenin onto self-assembled and fluoroapatite surfaces. *The Journal of Physical Chemistry B,* 113**,** 1833-1842.

TARASEVICH, B. J., LEA, S., BERNT, W., ENGELHARD, M. H. & SHAW, W. J. 2009b. Changes in the quaternary structure of amelogenin when adsorbed onto surfaces. *Biopolymers: Original Research on Biomolecules,* 91**,** 103-107.

TARASEVICH, B. J., PHILO, J. S., MALUF, N. K., KRUEGER, S., BUCHKO, G. W., LIN, G. & SHAW, W. J. 2015. The leucine-rich amelogenin protein (LRAP) is primarily monomeric and unstructured in physiological solution. *Journal of Structural Biology,* 190**,** 81-91.

TAYLOR, A. L., HAZE-FILDERMAN, A., BLUMENFELD, A., SHAY, B., DAFNI, L., ROSENFELD, E., LEISER, Y., FERMON, E., GRUENBAUM-COHEN, Y. & DEUTSCH, D. 2006. High yield of biologically active recombinant human amelogenin using the baculovirus expression system. *Protein Expression and Purification,* 45**,** 43-53.

TEAFORD, M. F. & UNGAR, P. S. 2000. Diet and the evolution of the earliest human ancestors. *Proceedings of the National Academy of Sciences,* 97**,** 13506-13511.

TEIXEIRA, R. J. P. B., ANDRADE, N. S., QUEIROZ, L. C. C., MENDES, F. M., MOURA, M. S., MOURA, L. D. F. A. D. D. & LIMA, M. D. M. 2018. Exploring the association between genetic and environmental factors and molar incisor hypomineralization: evidence from a twin study. *International Journal of Paediatric Dentistry,* 28**,** 198-206.

TERMINE, J. D., BELCOURT, A. B., CHRISTNER, P. J., CONN, K. M. & NYLEN, M. U. 1980. Properties of dissociatively extracted fetal tooth matrix proteins. I. Principal molecular species in developing bovine enamel. *Journal of Biological Chemistry,* 255**,** 9760-8.

TERPE, K. 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology,* 60**,** 523-533.

THE UNIPROT CONSORTIUM 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Research,* 45**,** D158-D169.

THEILLET, F. X., KALMAR, L., TOMPA, P., HAN, K. H., SELENKO, P., DUNKER, A. K., DAUGHDRILL, G. W. & UVERSKY, V. N. 2013. The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disordered Proteins,* 1**,** e24360.

THESLEFF, I. 2003. Epithelial-mesenchymal signalling regulating tooth morphogenesis. *Journal of Cell Science,* 116**,** 1647-1648.

TOMPA, P. 2002. Intrinsically unstructured proteins. *Trends in Biochemical Sciences,* 27**,** 527-33.

TOWBIN, H., STAEHELIN, T. & GORDON, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences,* 76**,** 4350-4354.

TRAVERS, K. J., PATIL, C. K., WODICKA, L., LOCKHART, D. J., WEISSMAN, J. S. & WALTER, P. 2000. Functional and Genomic Analyses Reveal an Essential Coordination between the Unfolded Protein Response and ER-Associated Degradation. *Cell,* 101**,** 249-258.

TSUCHIYA, M., TYE, C. E., SHARMA, R., SMITH, C. E. & BARTLETT, J. D. 2008. XBP1 may determine the size of the ameloblast endoplasmic reticulum. *Journal of Dental Research,* 87**,** 1058-62.

UCHIDA, K. 2003. Histidine and lysine as targets of oxidative modification. *Amino acids,* 25**,** 249-257.

UCHIDA, T., MURAKAMI, C., DOHI, N., WAKIDA, K., SATODA, T. & TAKAHASHI, O. 1997. Synthesis, secretion, degradation, and fate of ameloblastin during the matrix formation stage of the rat incisor as shown by immunocytochemistry and immunochemistry using region-specific antibodies. *Journal of Histochemistry & Cytochemistry,* 45**,** 1329-1340.

UCHIDA, T., MURAKAMI, C., WAKIDA, K., DOHI, N., IWAI, Y., SIMMER, J. P., FUKAE, M., SATODA, T. & TAKAHASHI, O. 1998. Sheath proteins: synthesis, secretion, degradation and fate in forming enamel. *European Journal of Oral Sciences,* 106**,** 308-314.

UCHIDA, T., TANABE, T., FUKAE, M. & SHIMIZU, M. 1991a. Immunocytochemical and immunochemical detection of a 32kDa nonamelogenin and related proteins in porcine tooth germs. *Archives of Histology and Cytology,* 54**,** 527-538.

UCHIDA, T., TANABE, T., FUKAE, M., SHIMIZU, M., YAMADA, M., MIAKE, K. & KOBAYASHI, S. 1991b. Immunochemical and immunohistochemical studies, using antisera against porcine 25 kDa amelogenin, 89 kDa enamelin and the 13-17 kDa nonamelogenins, on immature enamel of the pig and rat. *Histochemistry,* 96**,** 129-38.

UEMURA, A., OKU, M., MORI, K. & YOSHIDA, H. 2009. Unconventional splicing of XBP1 mRNA occurs in the cytoplasm during the mammalian unfolded protein response. *Journal of Cell Science,* 122**,** 2877-86.

VEIS, A. 2003. Amelogenin gene splice products: potential signaling molecules. *Cellular and Molecular Life Sciences CMLS,* 60**,** 38-55.

VIEIRA, A. R. & KUP, E. 2016. On the etiology of molar-incisor hypomineralization. *Caries Research,* 50**,** 166-169.

VOGEL, P., HANSEN, G., READ, R., VANCE, R., THIEL, M., LIU, J., WRONSKI, T., SMITH, D., JETER-JONES, S. & BROMMAGE, R. 2012. Amelogenesis imperfecta and other biomineralization defects in Fam20a and Fam20c null mice. *Veterinary pathology,* 49**,** 998-1017.

VOGT JR, R. V., PHILLIPS, D. L., HENDERSON, L. O., WHITFIELD, W. & SPIERTO, F. W. 1987. Quantitative differences among various proteins as blocking agents for ELISA microtiter plates. *Journal of immunological methods,* 101**,** 43-50.

VROMAN, L., ADAMS, A., FISCHER, G. & MUNOZ, P. 1980. Interaction of high molecular weight kininogen, factor XII, and fibrinogen in plasma at interfaces. *Blood,* 55**,** 156-159.

WALD, T., SPOUTIL, F., OSICKOVA, A., PROCHAZKOVA, M., BENADA, O., KASPAREK, P., BUMBA, L., KLEIN, O. D., SEDLACEK, R., SEBO, P., PROCHAZKA, J. & OSICKA, R. 2017. Intrinsically disordered proteins drive enamel formation via an evolutionarily conserved self-assembly motif. *Proceedings of the National Academy of Sciences of the United States of America,* 114**,** E1641-E1650.

WALLWORK, M. L., KIRKHAM, J., ZHANG, J., SMITH, D. A., BROOKES, S. J., SHORE, R. C., WOOD, S. R., RYU, O. & ROBINSON, C. 2001. Binding of matrix proteins to developing enamel crystals: an atomic force microscopy study. *Langmuir,* 17**,** 2508-2513.

WANG, C., ZHANG, J., YANG, F. & DU, C. 2018. One-shot method for purification of multiple natural amelogenin isoforms. *Journal of materials science & technology,* 34**,** 481-487.

WANG, H., TANNUKIT, S., ZHU, D., SNEAD, M. L. & PAINE, M. L. 2005. Enamel matrix protein interactions. *Journal of Bone and Mineral Research,* 20**,** 1032-1040.

WANG, M. & KAUFMAN, R. J. 2016. Protein misfolding in the endoplasmic reticulum as a conduit to human disease. *Nature,* 529**,** 326.

WANG, S.-K., CHOI, M., RICHARDSON, A. S., REID, B. M., LIN, B. P., WANG, S. J., KIM, J.-W., SIMMER, J. P. & HU, J. C. C. 2014a. ITGB6 loss-of-function mutations cause autosomal recessive amelogenesis imperfecta. *Human Molecular Genetics,* 23**,** 2157-2163.

WANG, S., CHOI, M., RICHARDSON, A., REID, B., SEYMEN, F., YILDIRIM, M., TUNA, E., GENCAY, K., SIMMER, J. & HU, J. 2014b. STIM1 and SLC24A4 are critical for enamel maturation. *Journal of Dental Research,* 93**,** 94S-100S.

WANG, S. K., AREF, P., HU, Y., MILKOVICH, R. N., SIMMER, J. P., EL-KHATEEB, M., DAGGAG, H., BAQAIN, Z. H. & HU, J. C. 2013. FAM20A mutations can cause enamel-renal syndrome (ERS). *PLoS Genetics,* 9**,** e1003302.

WANG, S. K., HU, Y., YANG, J., SMITH, C. E., RICHARDSON, A. S., YAMAKOSHI, Y., LEE, Y. L., SEYMEN, F., KORUYUCU, M. & GENCAY, K. 2016. Fam83h null mice support a neomorphic mechanism for human ADHCAI. *Molecular Genetics & Genomic Medicine,* 4**,** 46-67.

WARSHAWSKY, H. 1968. The fine structure of secretory ameloblasts in rat incisors. *The Anatomical Record,* 161**,** 211-29.

WARSHAWSKY, H., BAI, P. & NANCI, A. 1987. Analysis of crystallite shape in rat incisor enamel. *The Anatomical Record,* 218**,** 380-90.

WAUGH, D. S. 2011. An overview of enzymatic reagents for the removal of affinity tags. *Protein expression and purification,* 80**,** 283-293.
WAZEN, R. M., MOFFATT, P., ZALZAL, S. F., YAMADA, Y. & NANCI, A. 2009. A mouse model expressing a truncated form of ameloblastin exhibits dental and junctional epithelium defects. *Matrix Biology,* 28**,** 292-303.
WEERHEIJM, K. L., JALEVIK, B. & ALALUUSUA, S. 2001. Molar-incisor hypomineralisation. *Caries Research,* 35**,** 390-1.
WEK, R. C., JIANG, H. Y. & ANTHONY, T. G. 2006. Coping with stress: eIF2 kinases and translational control. *Biochemical Society Transactions,* 34**,** 7-11.
WEN, H., FINCHAM, A. & MORADIAN-OLDAK, J. 2001. Progressive accretion of amelogenin molecules during nanospheres assembly revealed by atomic force microscopy. *Matrix Biology,* 20**,** 387-395.
WHO 2017. Sugars and dental caries. World Health Organization.
WIEDEMANN-BIDLACK, F. B., BENIASH, E., YAMAKOSHI, Y., SIMMER, J. P. & MARGOLIS, H. C. 2007. pH triggered self-assembly of native and recombinant amelogenins under physiological pH and temperature in vitro. *Journal of structural biology,* 160**,** 57-69.
WILLIAMSON, M. P. 1994. The structure and function of proline-rich regions in proteins. *Biochemical Journal,* 297 (Pt 2)**,** 249-60.
WONG, S. S. 1991. *Chemistry of protein conjugation and cross-linking*, CRC press.
WRIGHT, J., HART, P., ALDRED, M., SEOW, K., CRAWFORD, P., HONG, S., GIBSON, C. & HART, T. 2003. Relationship of phenotype and genotype in X-linked amelogenesis imperfecta. *Connective Tissue Research,* 44**,** 72-78.
WRIGHT, J. T., CARRION, I. A. & MORRIS, C. 2015. The molecular basis of hereditary enamel defects in humans. *Journal of Dental Research,* 94**,** 52-61.
WRIGHT, J. T., HART, T. C., HART, P. S., SIMMONS, D., SUGGS, C., DALEY, B., SIMMER, J., HU, J., BARTLETT, J. D., LI, Y., YUAN, Z. A., SEOW, W. K. & GIBSON, C. W. 2009. Human and mouse enamel phenotypes resulting from mutation or altered expression of AMEL, ENAM, MMP20 and KLK4. *Cells Tissues Organs,* 189**,** 224-9.
WU, J. & KAUFMAN, R. J. 2006. From acute ER stress to physiological roles of the Unfolded Protein Response. *Cell Death & Differentiation,* 13**,** 374-84.
XIAO, Y. & ISAACS, S. N. 2012. Enzyme-linked immunosorbent assay (ELISA) and blocking with bovine serum albumin (BSA)—not all BSAs are alike. *Journal of immunological methods,* 384**,** 148-151.
XU, L., HARADA, H., YOKOHAMA-TAMAKI, T., MATSUMOTO, S., TANAKA, J. & TANIGUCHI, A. 2006. Reuptake of extracellular amelogenin by dental epithelial cells results in increased levels of amelogenin mRNA through enhanced mRNA stabilization. *Journal of Biological Chemistry,* 281**,** 2257-2262.
XU, Y., PISTON, D. W. & JOHNSON, C. H. 1999. A bioluminescence resonance energy transfer (BRET) system: application to interacting circadian clock proteins. *Proceedings of the National Academy of Sciences,* 96**,** 151-156.
YADEGARI, Z., BANDEHPOUR, M., KAZEMI, B. & SHARIFI-SARASIABI, K. 2015. Expression of Recombinant Human Amelogenin in Iranian Lizard

Leishmania and Its Biological Function Assay. *Iranian Journal of Public Health,* 44**,** 987-96.

YAHYAZADEHFAR, M., BAJAJ, D. & AROLA, D. D. 2013. Hidden contributions of the enamel rods on the fracture resistance of human teeth. *Acta Biomaterialia,* 9**,** 4806-4814.

YAM, G. H.-F., GAPLOVSKA-KYSELA, K., ZUBER, C. & ROTH, J. R. 2007. Sodium 4-phenylbutyrate acts as a chemical chaperone on misfolded myocilin to rescue cells from endoplasmic reticulum stress and apoptosis. *Investigative ophthalmology & visual science,* 48**,** 1683-1690.

YAMADA, Y. & KLEINMAN, H. K. 1992. Functional domains of cell adhesion molecules. *Current Opinion in Cell Biology,* 4**,** 819-823.

YAMAKOSHI, Y., RICHARDSON, A. S., NUNEZ, S. M., YAMAKOSHI, F., MILKOVICH, R. N., HU, J. C. C., BARTLETT, J. D. & SIMMER, J. P. 2011. Enamel proteins and proteases in Mmp20 and Klk4 null and double-null mice. *European Journal of Oral Sciences,* 119**,** 206-216.

YAMAKOSHI, Y., TANABE, T., FUKAE, M. & SHIMIZU, M. 1994. Porcine amelogenins. *Calcified tissue international,* 54**,** 69-75.

YAMAZAKI, H., BENIASH, E., YAMAKOSHI, Y., SIMMER, J. P. & MARGOLIS, H. C. 2017. Protein Phosphorylation and Mineral Binding Affect the Secondary Structure of the Leucine-Rich Amelogenin Peptide. *Frontiers in Physiology,* 8.

YANG, Y., WANG, H. & ERIE, D. A. 2003. Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy. *Methods,* 29**,** 175-187.

YE, J., RAWSON, R. B., KOMURO, R., CHEN, X., DAVE, U. P., PRYWES, R., BROWN, M. S. & GOLDSTEIN, J. L. 2000. ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. *Molecular Cell,* 6**,** 1355-64.

YOSHIDA, H. 2007. ER stress and diseases. *The FEBS Journal,* 274**,** 630-658.

YOSHIDA, H., HAZE, K., YANAGI, H., YURA, T. & MORI, K. 1998. Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose-regulated proteins. Involvement of basic leucine zipper transcription factors. *Journal of Biological Chemistry,* 273**,** 33741-9.

YOSHIDA, H., MATSUI, T., YAMAMOTO, A., OKADA, T. & MORI, K. 2001. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell,* 107**,** 881-891.

YOSHIDA, H., OKADA, T., HAZE, K., YANAGI, H., YURA, T., NEGISHI, M. & MORI, K. 2000. ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Molecular and Cellular Biology,* 20**,** 6755-67.

YOUNG, L., JERNIGAN, R. & COVELL, D. 1994. A role for surface hydrophobicity in protein-protein recognition. *Protein Science,* 3**,** 717-729.

YOUNGER, J. M., CHEN, L., REN, H.-Y., ROSSER, M. F., TURNBULL, E. L., FAN, C.-Y., PATTERSON, C. & CYR, D. M. 2006. Sequential quality-control checkpoints triage misfolded cystic fibrosis transmembrane conductance regulator. *Cell,* 126**,** 571-582.

ZALZAL, S. F., SMITH, C. E. & NANCI, A. 2008. Ameloblastin and amelogenin share a common secretory pathway and are co-secreted during enamel formation. *Matrix Biology,* 27**,** 352-9.

ZEICHNER-DAVID, M., CHEN, L. S., HSU, Z., REYNA, J., CATON, J. & BRINGAS, P. 2006. Amelogenin and ameloblastin show growth-factor like activity in periodontal ligament cells. *European Journal of Oral Sciences,* 114**,** 244-253.

ZHU, L., USKOKOVIĆ, V., LE, T., DENBESTEN, P., HUANG, Y., HABELITZ, S. & LI, W. 2011. Altered self-assembly and apatite binding of amelogenin induced by N-terminal proline mutation. *Archives of Oral Biology,* 56**,** 331-336.

Supplementary data

Details of analytical SDS PAGE and western blotting

As indicated in section 2.1.1.4 p. 100 - 103 (“SDS PAGE and western blotting”), the sample loading for each SDS PAGE analysis was optimised. This was to obtain strong enough band staining to display the proteins present in the extracts, and at the same time, avoid overloading so that the bands could be distinguished clearly.

The volumes and concentrations differed depending on:

- The visualisation technique employed; e.g.: silver staining is much more sensitive than Coomassie Blue (Morrissey, 1981)

- The sample itself, e.g. The starting material was usually more concentrated than the purification fractions collected; it was sometimes so concentrated that dilution was necessary to be able to distinguish its contents on analytical SDS PAGE.

This appendix details the sample preparation and loading for the analytical SDS PAGE and western blotting Figures comprised in the results (Chapter 3). It details:

- The composition of samples before electrophoresis
- The molecular weight markers used (and their dilution, if there is).
- The volumes of samples and markers loaded

It specifies also the staining techniques employed.

- For Coomassie Blue staining, the duration of staining with Coomassie Blue is provided. As for the reagent employed, details are in the ‘Materials and Methods’ chapter (Section 2.1.1.4) pp. 100 - 102
- For silver staining, the duration of staining and development steps are provided here. As for details about the kit and whole procedure employed, see Materials and Methods’ chapter (Section 2.1.1.4) pp. 100 - 102
- For western blotting, this section provides with the duration of incubation with the primary and secondary antibodies, the dilution factor and the buffers in which the antibodies are diluted. The primary and secondary antibodies do not change between all western blots. The blotting procedure and reagents are all detailed in the Materials and Methods’ chapter (Section 2.1.1.4) pp. 101, 103

1.1 Expression of recombinant WT and mutant p.Y64H amelogenins in E. coli (Section 3.1.1.1)

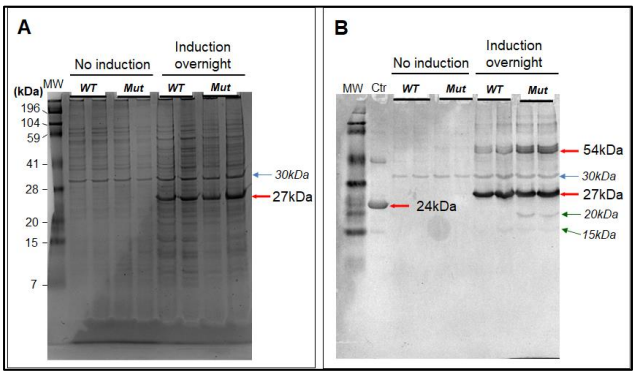


Figure 35 (Section 3.1.1.1 p. 137)

***SDS PAGE and western blotting of expressed* E. coli *proteins*.**

The sample preparation, loading and staining/blotting methods are detailed below.

	Analytical SDS PAGE (Figure 35A)	Western Blotting (Figure 35B)
Sample preparation	30µL sample +10 µL sample buffer 4X	
Volume sample loaded	10 µL test samples loaded 5 µL control loaded	
Marker (volume loaded)	Bio-Rad, cat # 161-0318 (3 µL)	
Staining	Coomassie Blue for 5-15 min. (until signal visible). Wash in distilled H_2O	1h30 with primary antibodies (1/10,000 in TTBS) 1h30 with secondary antibody (1/5000 in TTBS) 15 sec. development. Wash in distilled H_2O

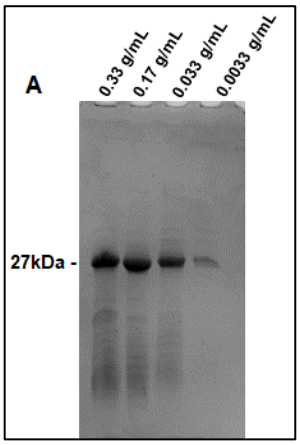


Figure 36A (Section 3.1.1.2.1 p. 139)

Optimisation of the weight of* E. coli *paste to volume of acetic acid in the extraction process.

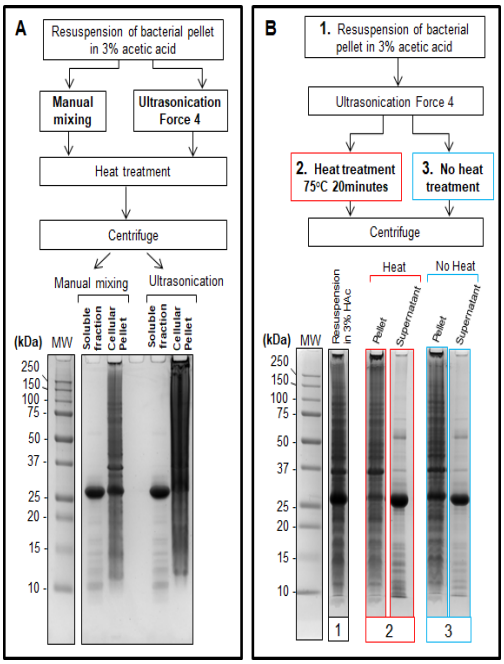
The sample preparation, loading and staining/blotting methods are detailed below.

	Analytical SDS PAGE (Figure 36A)
Sample preparation	30µL + 10 µL sample buffer 4X
Volume sample loaded	10 µL test samples loaded
Staining	Coomassie Blue for 5-15 min. (until signal visible). Wash in distilled H_2O

Figure 37 (Section 3.1.1.2.2 p. 141)

Optimisation of the mixing regimen and temperature on the yield of r-amelogenin extraction.

The sample preparation, loading and staining/blotting methods are detailed below.



Analytical SDS PAGE (Figure 37A)	
Sample preparation (volumes loaded)	*Soluble fractions*: Lyophilisation + Resuspension to identical volume in loading buffer 1X (10µL loaded) *Cellular pellet*: 15mg taken, mixed with 80µL sample buffer 1X (5 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	Coomassie Blue for 1h15 min. Washed in distilled H_2O

Analytical SDS PAGE (Figure 37B)	
Sample preparation (volumes loaded)	100 µL aliquots taken from solutions after steps 2 (heat treatment) and 3 (No heat treatment). Aliquots centrifuged at 3220g for 20 minutes. Supernatants and pellets lyophilised and resuspended to 100 µL sample buffer 1X. (10 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	Coomassie Blue overnight. Washed in distilled H_2O

1.2 Optimisation of 3% acetic acid extraction method (Section 3.1.2.1)

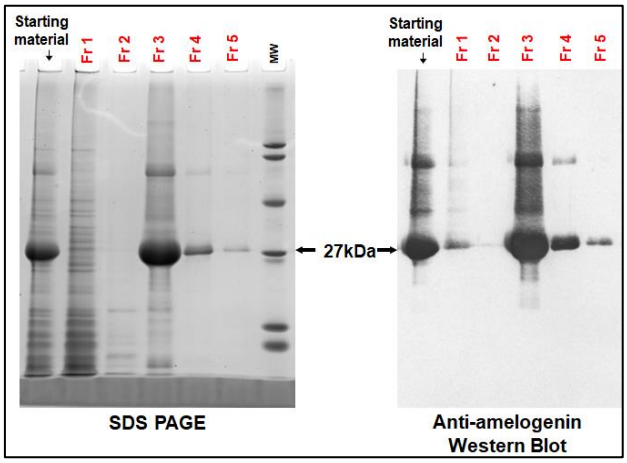


Figure 42B

(Section 3.1.2.1.1 p. 150)

Nickel column chromatography round 1.

The sample preparation, loading and staining/blotting methods are detailed below.

	Analytical SDS PAGE (Figure 42B)	Western blotting (Figure 42B)
Sample preparation	Samples lyophilised and resuspended to initial volume in sample buffer 1X	
Volume sample loaded	Starting material: 2.5 µL loaded Fr 1 – 5: 10 µL loaded	
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)	
Staining	Coomassie Blue overnight. Wash in distilled H_2O	1h30 with primary antibodies (1/2,000 in blocking solution) 1h30 with secondary antibody (1/3,000 in blocking solution) 5-7 min. development. Wash in distilled H_2O

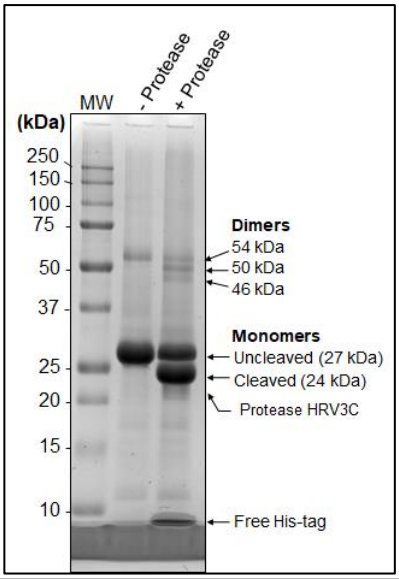


Figure 43 (Section 3.1.2.1.2 p. 153)

Optimisation of the mixing regimen and temperature on the yield of r-amelogenin extraction.

The sample preparation, loading and staining /blotting methods are detailed below.

Analytical SDS PAGE (Figure 43)	
Sample preparation	Samples lyophilised and resuspended to initial volume in sample buffer 1X (10 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	Coomassie Blue overnight Wash in distilled H_2O

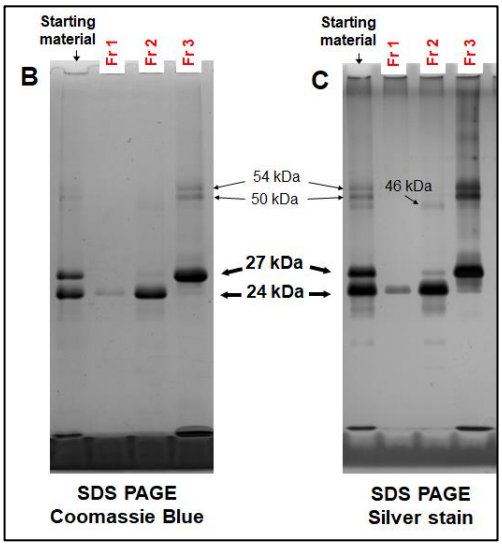


Figure 44

(Section 3.1.2.1.2 p. 156)

Nickel column chromatography round 2.

The sample preparation, loading and staining /blotting methods are detailed below.

	Analytical SDS PAGE (Figure 44B)	Analytical SDS PAGE (Figure 44C)
Sample preparation	Samples lyophilised and resuspended to initial volume in sample buffer 1X	Samples lyophilised and resuspended in sample buffer 1X, diluted 10x
Volume sample loaded	Starting material: 5 µL loaded Fr 1 – 3: 10 µL loaded	Starting material: 5 µL loaded Fr 1 – 3: 10 µL loaded
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)	Bio-Rad, cat # 161-0373, Diluted 20x (3 µL)
Staining	Coomassie Blue overnight. Wash in distilled H_2O	Silver staining (kit # 24612, ThermoFisher Scientific) 20 min. in staining solution 3 min.10 second. in development

1.3 Purification of r-amelogenin using nickel column chromatography (Section 3.1.2.2)

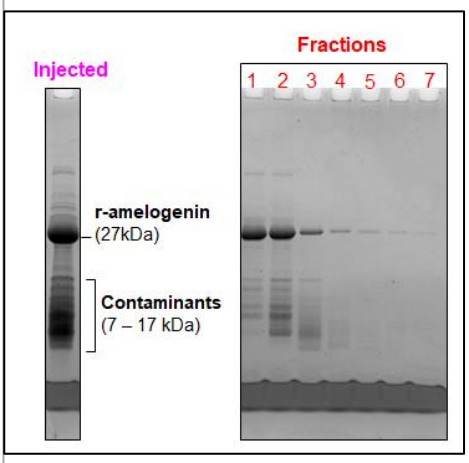


Figure 46 (Section 3.1.2.2.1 p. 163)

Separation of acetic acid extract components by size exclusion chromatography using Bio-gel-p30 matrix, 35 cm column length

The sample preparation, loading and staining /blotting methods are detailed below.

Analytical SDS PAGE (Figure 46)	
Sample preparation	Samples lyophilised and resuspended to initial volume in 1X sample buffer (10 µL loaded)
Staining	1h15 min. in Coomassie Blue. Wash in distilled H_2O

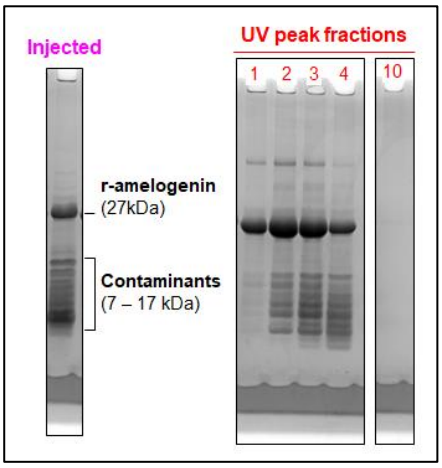


Figure 47 (Section 3.1.2.2.2 p. 165)

Separation of acetic acid extract components by size exclusion chromatography using Bio-gel-p10 matrix, 35 cm column length

The sample preparation, loading and staining /blotting methods are detailed below.

Analytical SDS PAGE (Figure 47)	
Sample preparation	Samples lyophilised and resuspended to initial volume in 1X sample buffer (10 µL loaded)
Staining	1h15 min. in Coomassie Blue. Wash in distilled H_2O

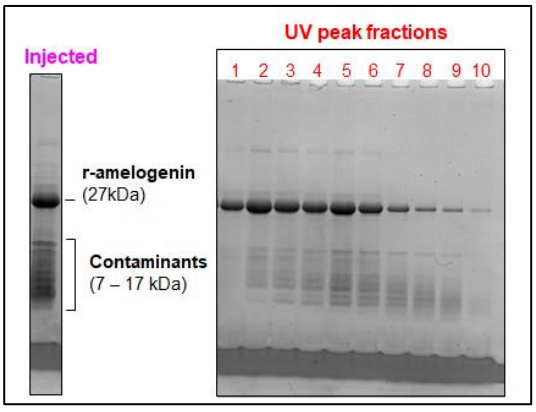


Figure 48 (Section 3.1.2.2.3 p. 166)

Separation of acetic acid extract components by size exclusion chromatography using Bio-gel-p10 matrix, 95 cm column length.

The sample preparation, loading and staining /blotting methods are detailed below.

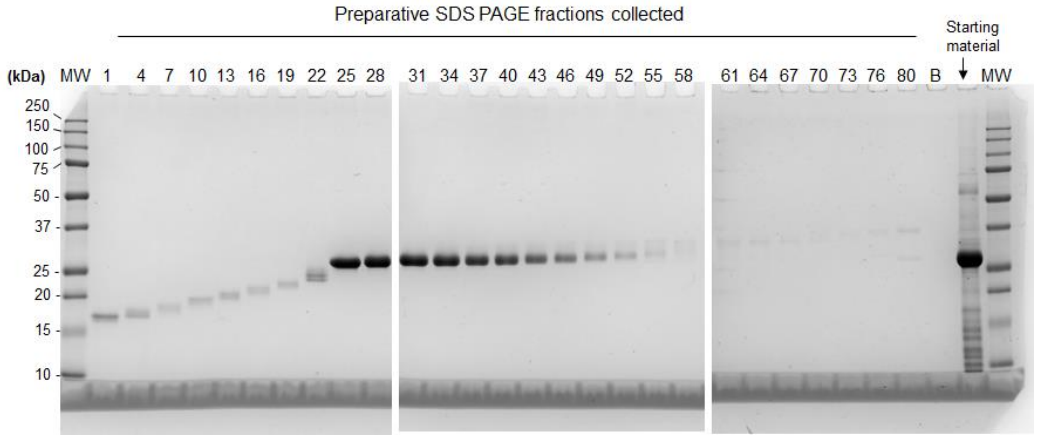
Analytical SDS PAGE (Figure 48)	
Sample preparation	Samples lyophilised and resuspended to initial volume in 1X sample buffer (10 µL loaded)
Staining	1h15 min. in Coomassie Blue. Then wash in distilled H_2O

1.4 Purification of r-amelogenin using preparative SDS PAGE (Section 3.1.2.3)

Figure 50 (Section 3.1.2.3.1 p. 170)

Analytical SDS PAGE with Coomassie Blue staining to identify the fractions of interest (those containing r-amelogenin).

The sample preparation, loading and staining /blotting methods are detailed below.

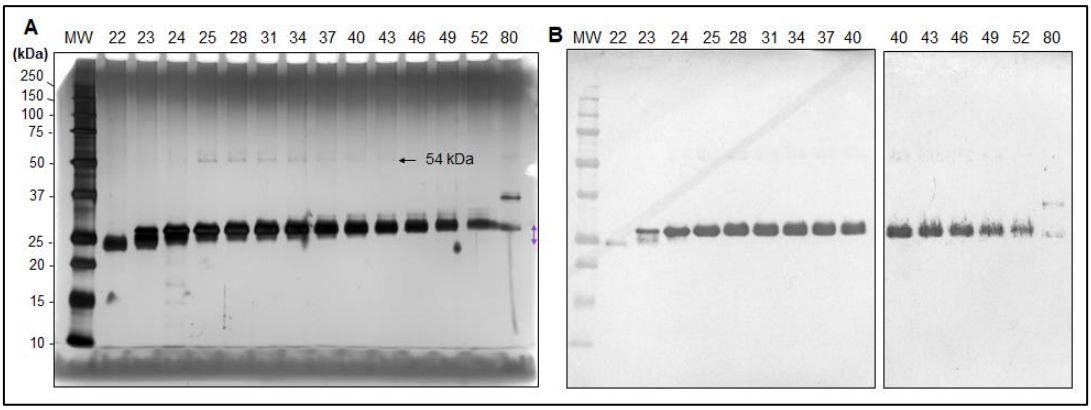


Analytical SDS PAGE (Figure 50)	
Sample preparation	Samples lyophilised and resuspended to initial volume in sample buffer 1X
Volume sample loaded	Starting material: 5 µL loaded Fractions + blank (B): 10 µL loaded
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	Coomassie Blue overnight Wash in distilled H_2O

Figure 51 (Section 3.1.2.3.2 p. 171)

Analytical SDS PAGE of the fractions of interest (obtained by preparative SDS PAGE) with silver staining and western blotting.

The sample preparation, loading and staining /blotting methods are detailed below.

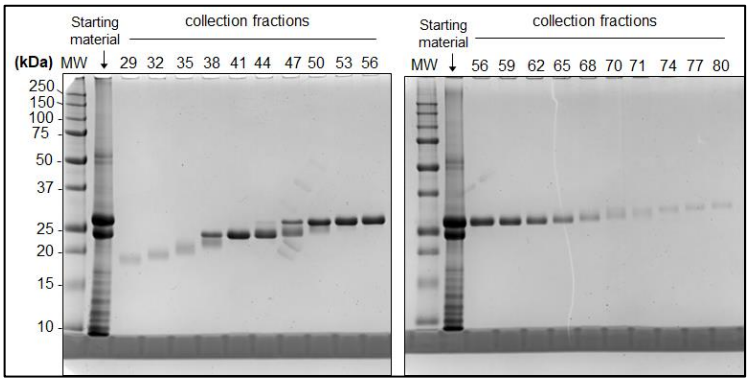


	Analytical SDS PAGE (Figure 51A)	Western blotting (Figure 51B)
Sample preparation	Samples lyophilised and resuspended in sample buffer 1X, diluted 10x (5 µL loaded)	Samples lyophilised and resuspended in sample buffer 1X, (5 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373, Diluted 20x in sample buffer 1X (2.5 µL loaded)	Bio-Rad, cat # 161-0373, (3 µL loaded)
Staining	Silver staining (kit # 24612, ThermoFisher Scientific) 15 min. in staining solution 5 min.10 second. in development	1h30min. with primary antibodies (1/1,000 in blocking solution) 1h30min. with secondary antibodies (1/3,000 in blocking solution) 2 min 10 second. development Wash in distilled H_2O

Figure 54A (Section 3.1.2.3.2 p. 177)

Analytical SDS PAGE showing the isolation of 'cleaved' r-amelogenin from 'uncleaved r-amelogenin using preparative SDS PAGE. (A) Coomassie Blue staining.

The sample preparation, loading and staining /blotting methods are detailed below.

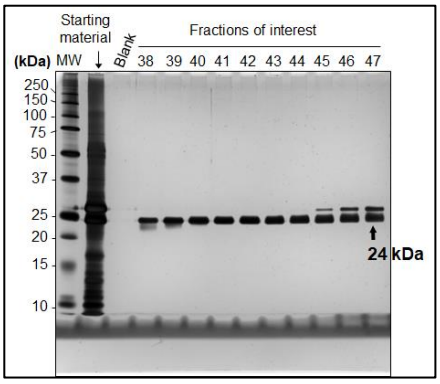


Analytical SDS PAGE (Figure 54A)	
Sample preparation	Samples lyophilised and resuspended to initial volume in sample buffer 1X
Volume sample loaded	Starting material: 2.5 µL loaded Fractions: 10 µL loaded
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL loaded)
Staining	Coomassie Blue overnight Wash in distilled H_2O

Figure 54B (Section 3.1.2.3.2 p. 177)

Analytical SDS PAGE showing the isolation of 'cleaved' r-amelogenin from 'uncleaved r-amelogenin using preparative SDS PAGE. (B) Silver staining.

The sample preparation, loading and staining /blotting methods are detailed below.



Analytical SDS PAGE (Figure 54B)	
Sample preparation	Samples lyophilised and resuspended in sample buffer 1X with 10x dilution
Volume sample loaded	Starting material: 2.5 µL loaded Fractions: 5 µL loaded
Marker (volume loaded)	Bio-Rad, cat # 161-0373 Diluted 20x in loading buffer 1X (2.5 µL loaded)
Staining	Silver staining (kit# 24612, ThermoFisher Scientific) 30 min. in staining solution 5 min.05 second. development

1.5 Fluorescence-based binding assay (Section 3.2.1)

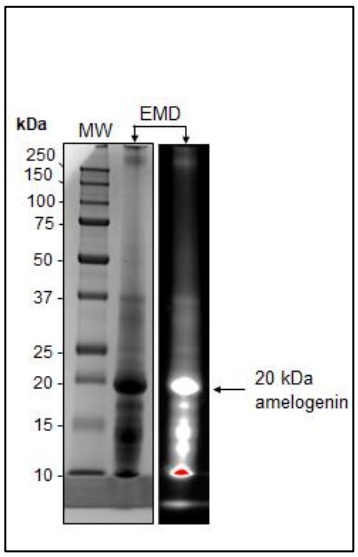


Figure 55 (Section 3.2.1.1 p. 180)

EMD proteins sucessfully labelled with FITC

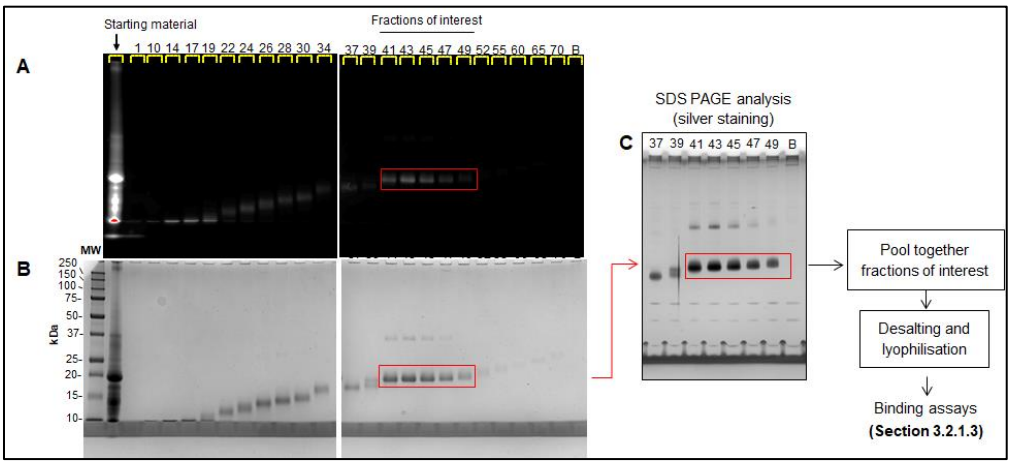
The sample preparation, loading and staining /blotting methods are detailed below.

Analytical SDS PAGE (Figure 55)	
Sample preparation	Samples lyophilised and resuspended to initial volume in sample buffer 1X (2.5 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	1) In distilled H_2O. Picture taken: choosing « fluorescein (Blot) » application 2) In Coomassie Blue overnight . Wash in distilled H_2O

Figure 57 (Section 3.2.1.2 p. 182)

Fractions obained when FITC-labeled EMD proteins were subjected to preparative SDS PAGE.

The sample preparation, loading and staining /blotting methods are detailed below.



Analytical SDS PAGE (Figure 57 A and B)	
Sample preparation	Samples lyophilised and resuspended to initial volume in sample buffer 1X
Volume sample loaded	Starting material: 2.5 µL loaded Fractions + blank (B): 10 µL loaded
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	1) In distilled H_2O. Picture taken: choosing « fluorescein (Blot) » application 2) In Coomassie Blue 1h 15min. Wash in distilled H_2O

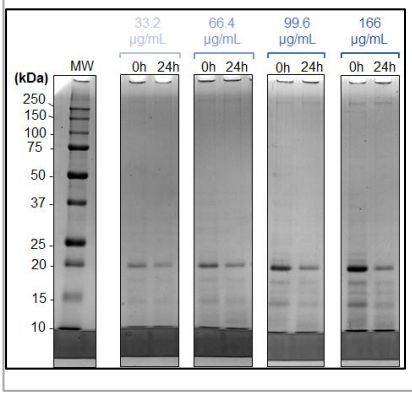
Analytical SDS PAGE (Figure 57C)	
Sample preparation (Volume loaded)	Samples lyophilised and resuspended to initial volume in sample buffer 1X Diluted 10x (5 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373 Diluted 20x in loading buffer 1X (3 µL loaded)
Staining	Silver staining (kit# 24612, ThermoFisher Scientific) 20 min. in staining solution 2 min 30 sec. development

1.6 Monitoring the kinetics of amelogenin-amelogenin interactions as a function of UV absorbance (Section 3.2.2)

Figure 61B (Section 3.2.2.1 p. 192)

A range of initial EMD concentrations (33.2 – 166 µg/mL) was used to investigate EMD adsorption onto the microwells surfaces.

The sample preparation, loading and staining /blotting methods are detailed below.



Analytical SDS PAGE (Figure 61B)	
Sample preparation (volume loaded)	Samples lyophilised and resuspended to initial volume in sample buffer 1X (10 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	In Coomassie Blue overnight . Wash in distilled H_2O

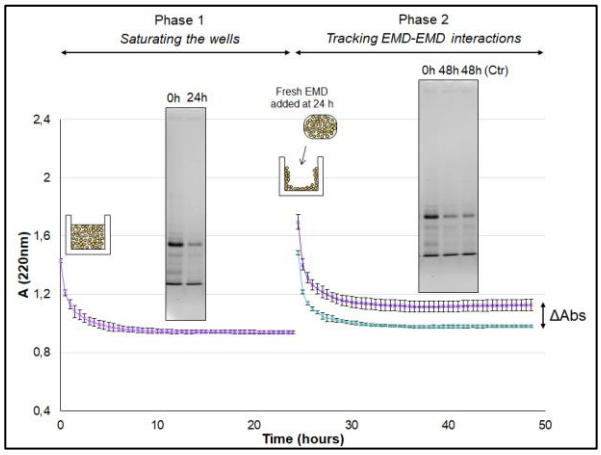


Figure 64 (Section 3.2.2.2 p. 198)

Analysis of EMD-EMD binding kinetics.

The sample preparation, loading and staining /blotting methods are detailed below.

Analytical SDS PAGE (Figure 64)	
Sample preparation (volume loaded)	Samples lyophilised and resuspended to initial volume in sample buffer 1X (10 µL loaded)
Staining	In Coomassie Blue overnight . Wash in distilled H_2O

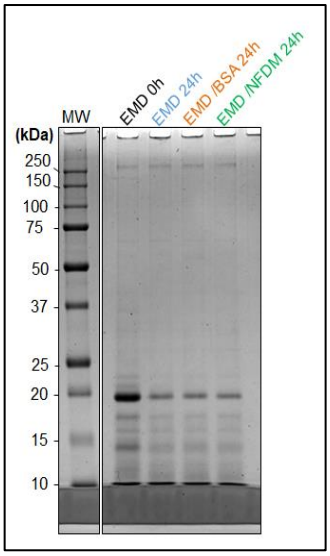


Figure 68 (Section 3.2.2.2 p. 205)

Ability of goldstandard blockers BSA and NFDM to block EMD from binding to microwell surfaces.

The sample preparation, loading and staining /blotting methods are detailed below.

Analytical SDS PAGE (Figure 68)	
Sample preparation (volume loaded)	Samples lyophilised and resuspended to initial volume in sample buffer 1X (10 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	In Coomassie Blue overnight . Wash in distilled H_2O

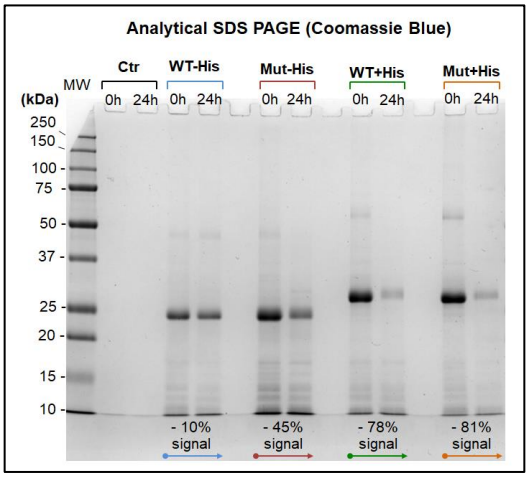


Figure 69B
(Section 3.2.2.5 p. 209)

Comparison of r-amelogenin depletion from solution after 24h incubation in microwells for WT^{-His}, Mut^{-His}, WT^{+His}, Mut^{+His}.

The sample preparation, loading and staining /blotting methods are detailed below.

Analytical SDS PAGE (Figure 69B)	
Sample preparation (volume loaded)	Samples lyophilised and resuspended to initial volume in sample buffer 1X (10 µL loaded)
Marker (volume loaded)	Bio-Rad, cat # 161-0373 (3 µL)
Staining	In Coomassie Blue overnight . Wash in distilled H_2O

REFERENCES

MORRISSEY, J. H. (1981) Silver stain for proteins in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity. *Analytical biochemistry,* 117**,** 307-310.