Mechanisms and Functions of Collagen Glycosylations in Bone

Marnisa Sricholpech

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Dentistry (Oral Biology)

Chapel Hill

2010

Approved by

Professor	Mitsuo Yamauchi
Professor	Eric Everett
Professor	Heather Yeowell
Professor	Janet Rubin
Associate Professor	Yoshiyuki Mochida

Abstract

Marnisa Sricholpech

Mechanisms and Functions of Collagen Glycosylations in Bone

(Under the direction of Professor Mitsuo Yamauchi)

O-linked glycosylation of hydroxylysine (Hyl) is one of the unique post-translational modifications found in collagens and collagen-like proteins. In type I collagen, some of the helical Hyl residues are galactosylated, forming galactosylhydroxylysine (G-Hyl) which can be further glucosylated into glucosylgalactosylhydroxylysine (GG-Hyl). The critical importance of these glycosylated Hyl residues was implicated since alterations in their levels were associated with several human connective disorders. To date, the multifunctional enzyme, lysyl hydroxylase 3 (LH3), has been shown to be the major glucosyltransferase enzyme, while its galactosyltransferase function is still debatable. For bone type I collagen, little is known about the regulatory mechanisms and the significance of Hyl glycosylation. Therefore, in this study, we have aimed to elucidate the formation mechanism and functions of the glycosylated Hyl in type I collagen by utilizing mouse osteoblast (MC3T3-E1 (MC) cells) culture system. Short hairpin RNA technology was employed to stably suppress the expression of LH3 gene (*Plod3*) and generate single cell-derived clones (Sh clones). Characterization of type I collagen, synthesized by the Sh clones, showed significant level decrease of GG-Hyl with concomitant increase of G-Hyl while total Hyl remained unchanged, thus indicating the major function of LH3 in G-Hyl glucosylation (Study I). By

mass spectrometry, specific molecular loci and forms of glycosylation have been identified at residues α1-87, α1-174 and α2-173. In addition, the effect of lowered LH3-mediated glucosylation was observed in the formation and maturation of intermolecular cross-links, collagen matrix organization and mineralization (Study II). Most recently, novel collagen galactosyltransferase enzymes, glycosyltransferase 25 domain 1 and 2 (GLT25D1 and D2), have been discovered and characterized. We have shown in study I that Glt25d1 is the only isoform expressed in MC cells. By suppressing Glt25d1, the type I collagen synthesized showed significantly lower levels of both G-Hyl and GG-Hyl (Study III). In conclusion, the results from all these studies clearly indicate that for bone type I collagen, Hyl galactosylation is modulated by Glt25d1 and subsequent glucosylation by LH3. Moreover, the glucose units in the GG-Hyl residues appeared to play essential roles in the formation of normal collagen template for the mineralization process.

Acknowledgements

This dissertation study was carried out at the NC Oral Health Institute, School of Dentistry, University of North Carolina at Chapel Hill and would not have been complete without the guidance and collaborative support from several people.

First and foremost, I would like to express my gratitude to my advisor Prof. Mitsuo Yamauchi for giving me the greatest opportunity to be in the program and develop basic research skills under his guidance. His passion in the field of Collagen Biochemistry had inspired me to pursue a dissertation project related to that. Even though, at times, it has been very tough with a lot of trials and errors, but now, I have realized that this is what "re-search" is all about. Moreover, I am most thankful for all the time he has spent mentoring me scientifically and also guiding me on how to live a happy life. His suggestions, patience and kind consideration would forever be in my thoughts.

I wish to thank all the members of the collagen biochemistry lab, past and present, for all the friendship, help and support throughout the years. I specifically would like to thank Dr. Duenpim Parisuthiman (Art), since she was the person who supervised me in my first lab rotation and helped me settling down when I first arrived in Chapel Hill. Next, I would not have understood about molecular cloning without the step-by-step instruction of Dr. Yoshiyuki Mochida, whose well-rounded talent was greatly missed when he had to move to Boston University. But even so, he has always been willing to help trouble-shooting the problems and giving useful comments and suggestions. So for that, I will always be grateful.

I would also like to thank Dr.Phimon Atsawasuwan, Dr.Patricia Miguez, Dr.Hideaki Nagaoka, Dr.Megumi Yokoyama for all their help, support and encouragements.

I am very grateful to Dr.Irina Perdivara and Dr.Kenneth Tomer from the Laboratory of Structural Biology, National Institute of Environmental Health Sciences for the wonderful collaboration. A significant part of this dissertation is greatly accomplished because of the invaluable contribution of Dr.Perdivara in the mass spectrometry analysis.

This is also a great opportunity to express my gratitudes to all the professors, my preliminary exam and dissertation committee members (Dr. Patrick Flood, Dr. Timothy Wright, Dr. Ching-Chang Ko, Dr. Heather Yeowell, Dr. Eric Everett, Dr. Janet Rubin, Dr. Yoshiyuki Mochida and Dr. Mitsuo Yamauchi) for their structural comments, suggestions and supports throughout the course of my studies. Moreover, I am very thankful to be in the company of the wonderful, supportive and fun group of oral biology students and for Ms. Cindy Blake, our program manager, who is always there for us.

Another group of people that I have to thank is the Thai community here in Chapel Hill. They have certainly made my life here, aside from the studies, more fun and relaxing with all the trips and activities that we have been involved in. I am pleased to thank Dr. and Mrs. Maixner for always welcoming me into their home during Thanksgivings and New Year's Eve. Thanks to all the members of the Thai Student Associations, past and present, for the friendship and fun times.

For my studies here, I have had the honour of being a recipient of the scholarship from the Royal Thai Government, for which I would always be grateful and hope that the experience that I have gained during my time here could, in some way, be of value to the dental and scientific education in Thailand. Aside from that, I would like to express my

gratitude to Dr.Theeralaksna Suddhasthira, the current Dean of Faculty of Dentistry, Mahidol University for her help, suggestions and encouragements all these years. To all my teachers in Thailand who as ever given me the knowledge and skills to be the person I am today, I greatly thank them. To my colleagues and friends in the Department of Oral Surgery and Oral Medicine, Faculty of Dentistry, Srinakharintwirot University, I am thankful for their consideration and supports.

Finally, my utmost thankfulness goes to my family, my mother, sister and brother, for all their love, understanding, patience, caring and support all along. For my father, I am sure he has been watching over me and I hope that my achievements will make him proud.

Tables of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

ACP	aldol condensation product
AFM	atomic force microscopy
Ald	aldehydes
AP	alkaline phosphatase
Asn	asparagine
Asp	aspartic acid
ATCC	American type culture collection
Bip	binding proteins
BMP	bone morphogenetic protein
bp	basepair(s)
C-	carboxy-
CBB	Coomassie Brilliant Blue
cDNA	complimentary deoxyribonucleic acid
COL1A2	type I collagen alpha 2
Cu	copper
DCPD	dicalcium phosphate dehydrate
DDR2	discoidin domain receptor 2
DDW	distilled deionized water
DEBS	dominant epidermolysis bullosa simplex
deH-	dehydro-
DHLNL	dihydroxylysinonorleucine

DHNL	dihydroxynorleucine
d-Pyr	deoxypyridinoline
DTT	dithiothreitol
ECM	extracellular matrix
EDS	Ehlers Danlos syndrome
EIC	extracted ion chromatogram
ER	endoplasmic reticulum
ESI	electrospray ionization
ETD	electron transfer dissociation
EV	empty vector
FACIT	fibril-associated collagen with interrupted triple helices
FBS	fetal bovine serum
G-	galactosyl
G_Sh clones	MC cell derived clones stably suppressing Glt25d1
GAG	glycosaminoglycan
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GG-	glucosylgalactosyl
GG-Hyl	glucosylgalactosylhydroxylysine
GGT	galactosylhydroxylysyl glucosyltransferase
G-Hyl	galactosylhydroxylysine
Glt25d1	mouse glycosyltransferase 25 domain 1
Glt25d2	mouse glycosyltransferase 25 domain 2
Glu	glutamine

Gly	glycine
GT	hydroxylysyl galactosyltransferase
HEK	human embryonic kidney
HHL	histidinohydroxylysinonorleucine
HHMD	histidinohydroxymerodesmosine
His	histidine
HLKNL	hydroxylysino-keto-norleucine
HLNL	hydroxylysinonorleucine
HNL	hydroxynorleucine
HPLC	high performance liquid chromatography
HSP	heat shock protein
Hyl	hydroxylysine
Hyl^{ald}	hydroxylysine aldehyde
Hyp	hydroxyproline
kDa	kilodalton
KO	knockout
LH	lysyl hydroxylase
LKNL	lysine-keto-norleucine
LNL	lysinonorleucine
LOX	lysyl oxidase
LOXL	lysyl oxidase-like
Lys	lysine
Lys^{ald}	lysine aldehyde

M	molar
m/z	mass to charge ratio
MBL	mannose-binding lectin
MC	MC3T3-E1 osteoblastic cell line from mouse calvaria
mRNA	messenger ribonucleic acid
MS	mass spectrometry
N-	amino-
NaB^3H_4	tritiated sodium borohydride
NH_3	ammonia
OI	osteogenesis imperfecta
OPG	osteoprotegerin
P3H	prolyl-3-hydroxylase
P4H	prolyl-4-hydroxylase
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCP	procollagen C-proteinase
PCR	polymerase chain reaction
PDI	protein disulfide isomerase
PH	prolyl hydroxylase
PLOD	procollagen-lysine 1, 2-oxoglutarate 5-dioxygenase
PNP	procollagen N-proteinase
PPI	peptidyl-prolyl *cis-trans* isomerase
Prl	pyrrole

Pro	proline
PVDF	polyvinylidene fluoride
Pyr	pyridinoline
res	residue
Ser	serine
Sh clones	MC3T3-E1 cells derived clones stably suppressing LH3
SLRPs	small leucine rich proteoglycans
ßAPN	ß-aminopropionitrile
v/v	volume/volume
w/w	weight/weight
Wk	week
α	alpha
ß	beta
δ	delta
γ	gamma
ε	epsilon
f	femto
µ	micro
m	milli
n	nano
A	alanine
C	cysteine
D	aspartic acid

E	glutamic acid
G	glycine
H	histidine
I	isoleucine
K	lysine
L	leucine
M	methionine
P	proline
R	arginine
S	serine
Y	tyrosine

Chapter 1

Review of Literature

1.1 Collagen superfamily

Collagens are a family of extracellular proteins which mainly serve as the structural basis of all tissues. In the vertebrates, there are now at least 29 collagen types, encoded by more than 40 different genes, which can be categorized into families based on their supramolecular assemblies and other specific features. The major categories include fibrillar collagens, fibril-associated collagens with interrupted triple helices (FACITs), network-forming collagens, transmembranous collagens and others (1-2). To date, increasing numbers of proteins, which contain triple-helical collagenous domains (e.g. C1q, adiponectin, collectins, ficolins, surfactant protein A and D, acetylcholinesterase, macrophage scavenger receptors, ectodysplasin and etc.) (3-5), have been discovered and many of those are involved in innate immunity.

Fibrillar collagens (type I, II, III, V, XI, XXIV and XXVII) constitutes the group of collagens which form fibrils with unique repeated banding pattern, called D-periodicity which arise from the ordered staggering of the collagen molecules. Each molecule is formed from three polypeptide chains named α chains, either in a homotrimeric or a heterotrimeric fashion, depending on the collagen type and some existing variants. Another common

characteristic of fibrillar collagens is the long triple helical region with unique Gly-X-Y triplet repeats, flanked by the short non-helical part called telopeptides. Among the fibrillar collagens, type I collagen is the most abundant collagen found in various connective tissues such as skin, bone, tendon, dentin and etc and has been shown to portray critical roles for maintaining the integrity and elasticity of the tissues.(3)

1.2 Type I collagen biosynthesis

Type I collagen, one of the fibrillar collagens, is formed from the syntheses of procollagen α1 and α2 chains, encoded by COL1A1 and COL1A2 genes respectively, in the endoplasmic reticulum (ER). At which time they undergo extensive co- and post-translational modifications including the unique hydroxylation of proline (Pro) and lysine (Lys) residues, subsequent O-linked glycosylation of specific hydroxylysine (Hyl) residues and asparagine-linked glycosylations. Proline residues in the Y position of the Gly-X-Y triplet are mostly hydroxylated by prolyl-4-hydroxylase (P4H), while some of those at the X position are hydroxylated by prolyl-3 hydroxylase (P3H). Hydroxyproline (Hyp) has been shown to be the key to the triple-helix stabilization. Lysine hydroxylation and its subsequent glycosylation will be reviewed in detail in section 1.6. The modified procollagen chains, two α1 and one α2, further associate by forming disulfide bonds at the C-propeptide region, and then fold into triple helical conformation towards the N-terminal region. Each of the polypeptide chains contains unique repetitive amino acid sequences, glycine-proline-X (Gly-Pro-X) or glycine-X-hydroxyproline (Gly-X-Hyp), which allow the folding of the 3 chains in to triple helical conformation where the Gly residues are located inside the molecule and the other amino acids facing outwards probably involving in the intermolecular interactions. In

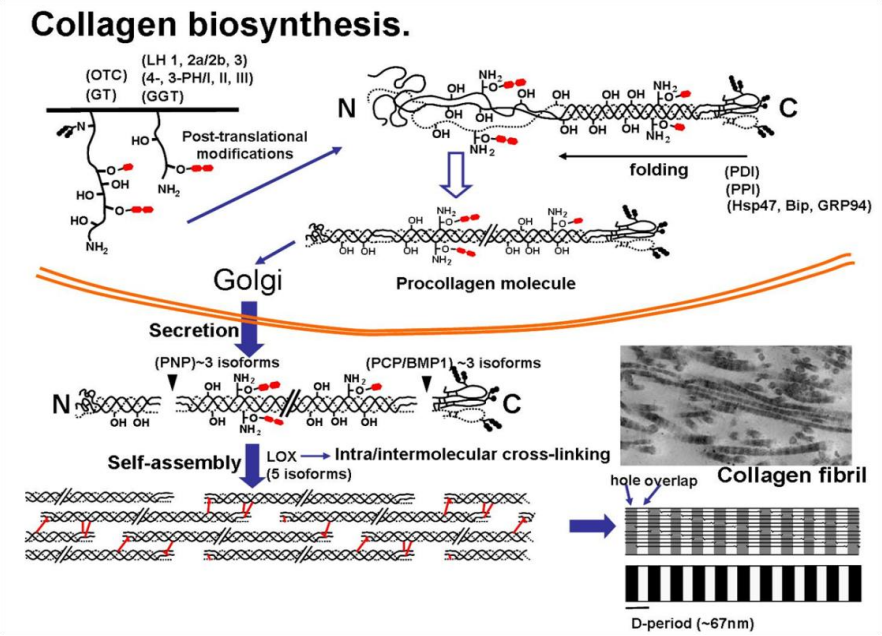

Figure 1.1) Schematic for the biosynthesis of type I collagen. The intracellular events include extensive post-translational modifications such as hydroxylation, glycosylation (both O- and N-linked), association of pro α chains and folding into triple helical molecule from the C- to N-terminus. The extracellular events involve the removal of both N- and C-propeptides, self assembly of collagen molecules into a fibril, enzymatic oxidative deamination of the telopeptidyl Lys and Hyl residues by LOX and subsequent intra- and intermolecular cross-linking. The collagen molecules are packed in parallel and are longitudinally staggered with respect to one another by some multiple of axial repeat distance, D (~67nm). This packing arrangement creates repeated regions of high and low packing density , i.e. overlap and hole regions, respectively, showing a characteristic banding pattern of collagen fibril seen at utltrastructural level. LH: lysyl hydroxylase, PH: prolyl hydroxylase, GGT: galactosylhydroxylysyl glucosyltransferase, GT: hydroxylysyl galactosyltransferase, OTC: oligsaccharyl transferase complex, PDI: protein disulfide isomerase, PPI: peptidyl-prolyl *cis-trans* isomerase, Hsp: heat shock protein, Bip: binding proteins, GRP: glucose-regulated protein, PNP: procollagen-N-proteinase, PCP: procollagen-C-proteinase, BMP: bone morphogenetic protein, LOX: lysyl oxidase, Solid line: pro α1 chain, Dotted line: pro α2 chain. Diagram reproduced with permission from Yamauchi M.

this process, several proteins act as chaperones in trimerization and folding of the α chains either by selectively binding to the unfolded procollagen α chains to prevent premature triple helix formation (e.g. P4H, protein disulphide isomerase (PDI), Bip/Grp78, Grp94, and immunophillins), or bind to the completely folded collagen molecule to stabilize the triple helix and possibly prevent premature aggregation of the procollagen molecules in the ER.

The procollagen molecules are then transported through the Golgi apparatus and secreted into the extracellular matrix, where the N- and C-propeptides are cleaved by their respective procollagen N-proteinase and procollagen C-proteinase, after which, the collagen molecules would spontaneously self-assemble in ordered stagger arrays. Each molecule would overlap each other by 234 residues, forming the 67 nm D-period repetitive regions of collagen which consist of the unique hole and overlap zones in the collagen fibrils (Figure 1.1) (reviewed in (3,6).

1.3 Type I collagen fibrillogenesis

Formation of type I collagen fibrils in physiological condition occurs as a spontaneous process. The processed collagen molecules self assemble and align by the interactions of the charge and hydrophobic regions in the helical region (7). Besides, the factors that could regulate collagen fibrillogenesis include the following:

- The levels of post-translational modifications e.g. Lys hydroxylation and glycosylation. Collagen with higher levels of modification is associated with the formation of fibrils with smaller diameter. (8-11).
- Collagen-binding small leucine-rich proteoglycans (SLRPs), containing the unique leucine-rich repeats, flanked by cysteine-rich clusters and attachments of glycosaminoglycan (GAG) chains, are shown to control collagen fibril formation (12-14). DCN is localized in close association with collagen fibers and is a well-characterized regulator of collagen fibrillogenesis (15). DCN knockout (KO) mice developed lax and fragile skin that exhibited loosely packed type I collagen fibrils, with larger and irregular diameter probably due to the uncontrolled lateral fusion of

fibrils (16). By manipulating the expression level of DCN in the mouse osteoblast cell culture system, we also have observed similar abnormalities of the collagen fibrils (17).

- Type III and V collagen have been shown to be the regulators of type I collagen fibrils lateral growth. Heterotypic fibrils formed from those three different collagen types can be found in skin. However, their presence in bone is considered low (18).
- Interestingly, some extracellular matrix proteins have been found to promote collagen fibril formation in vitro, It was speculated that the protein-collagen interaction may stabilize the initial aggregation of the collagen molecules for further fibril growth (reviewed in (3)).

1.4 Collagen cross-links in mineralized tissues

The collagen molecules, spontaneously aggregated into fibrils, are further stabilized by the final step of post-translational modification, the formation of intra and intermolecular cross-links. They are initiated by the generation of aldehydes at specific Lys or Hyl residues in the telopeptide regions of the α chains by the oxidative deamination function of the lysyl oxidase family. It includes lysyl oxidase (LOX) and lysyl oxidase-like (LOXL), LOXL2, LOXL3, and LOXL4 (19). LOX and LOXL has been shown to be active for fibrillar collagen while LOXL2 has been associated with basement membrane type IV collagen. As for the other isoforms, their substrate specificity have not been clearly defined (reviewed in (3)). It was reported that LOX is highly functional for growing, native collagen fibrils and was suggested that the intermolecular interaction between collagen molecules was important for the enzyme activity (20). Studies have shown that the binding sites for LOX in type I

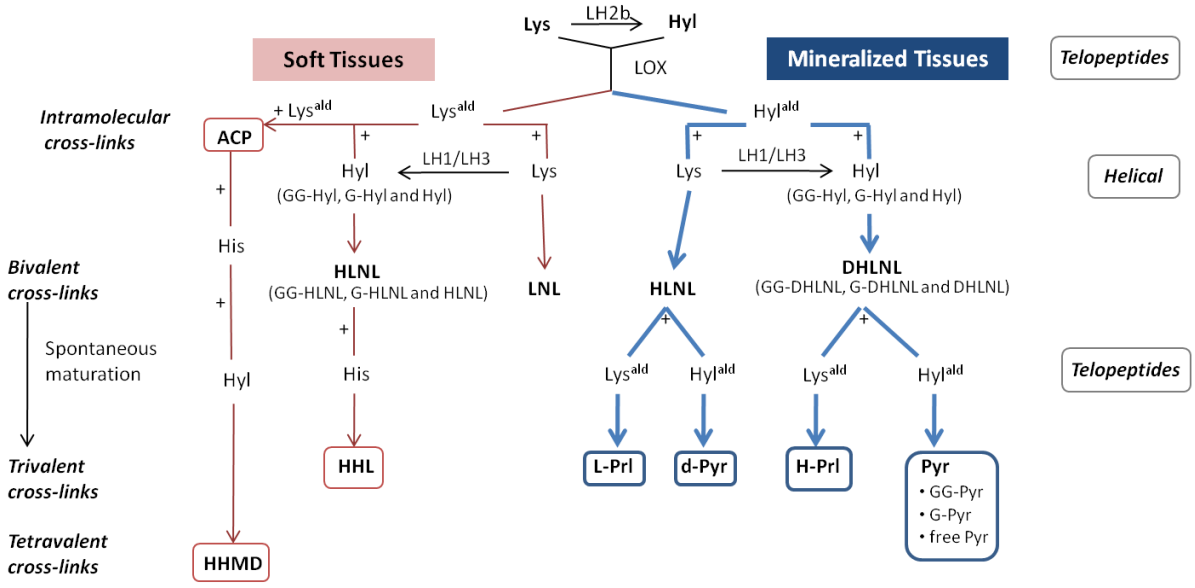

Figure 1.2) Collagen cross-linking pathways LH: Lysyl hydroxylase, LOX: Lysyl oxidase, ACP: Aldol condensation product, Lys^{ald}: Lys aldehyde, Hyl^{ald}: Hyl aldehyde, Bivalent cross-links are shown in their reduced forms: HLNL:hydroxylysinonorleucine, DHLNL:dihydroxylysinonorleucine, HHMD: histidinohydroxymerodesmosine, HHL: histidinohydroxylysinonorleucine, Pyr: pyridinoline, Prl: pyrrole, d-Pyr: deoxypyridinoline, GG-: glucosylgalactosyl-, G-: galactosyl-, L-: lysyl-, H-: hydroxylysyl-

collagen are in the triple helical region (21), potentially in the area with highly conserved sequences (Hyl-Gly-His-Arg) where it can catalyze the formation of aldehydes in the telopeptides of the adjacent collagen molecule (reviewed in (22)). The oxidative deamination activity of LOX enzymes can be inhibited by the well known lathyritic agent, β-aminopropionitrile (BAPN)(23). The resulting Lys aldehydes (Lys^{ald}) or Hyl aldehydes (Hyl^{ald}) are determinants for the tissue-specific cross-linking pathway by their involvement in a series of spontaneous intra- or intermolecular reactions and thus provides the matrices with tensile strength and elasticity which are essential for the functional integrity of the tissue.

1.4.1 Immature bivalent reducible cross-links

In fibrillar collagens, the telopeptides each contain one site of cross-link i.e. residue 9 of the N-terminus (α1-9^{N}, α2-9^{N}) and residue 16 of the C-terminus (α1-16^{C}). In the helical region, the cross-linking sites are found at residues 87 and 930. The Lys^{ald} and Hyl^{ald} at those telopeptidyl sites then spontaneously react with the helical Lys or Hyl residues and form two unique, tissue-specific, cross-linking pathways (Figure 1.2)(reviewed in (6,24)).

Hydroxylysine aldehyde-derived pathway (Keto-amines). The cross-links derived from this pathway are mostly found in mineralized tissues such as bone and dentin. The Lys residues in the telopeptides are mainly hydroxylated, converted into Hyl^{ald} and react with the ε-amino group of Hyl or Lys residues in the helical region of the adjacent collagen molecule. The Schiff bases formed (dehydro-dihydroxylysinonorleucine (deH-DHLNL) and dehydro-hydroxylysinonorleucine (deh-HLNL)) then undergo Amadori rearrangements to form hydroxylysino-keto-norleucine (HLKNL, Hyl^{ald} X Hyl) or lysine-keto-norleucine (LKNL, Hyl^{ald} X Lys) respectively. These keto-amine forms are more stable and may contribute, in part, to the insolubility of mineralized tissue collagen. However, for these immature cross-links to withstand the acid hydrolysis, they have to be stabilized by the reduction process with tritiated sodium borohydride ($NaB^{3}H_{4}$), and analyzed in their reduced forms i.e. dihydroxylysinonorleucine (DHLNL) and hydroxylysinonorleucine (HLNL) (reviewed in (6,24)) .

Lysine aldehyde-derived pathway (Aldimines). This leads to the formation of cross-links in non-mineralized tissue such as skin and tendon. The Lys^{ald} in the telopeptides reacts with Hyl or Lys in the helical region of the juxtaposed molecule, forming dehydro-hydroxylysinonorleucine (deH-HLNL, Lys^{ald} X Hyl) or dehydro-lysinonorleucine (deH-LNL,

Lys^{ald} X Lys). In addition, a unique reducible tetravalent cross-link, dehydro-histidino-hydroxymerodesmosine (deH-HHMD) can be formed between the aldol condensation product of two Lys^{ald}, histidine (His) and a helical Hyl (Lys^{ald} X Lys^{ald} X His X Hyl). Similar to the keto-amine cross-links, the aldimines are also analyzed in their reduced forms i.e. hydroxylysinonorleucine (HLNL), lysinonorleucine (LNL) and histidino-hydroxymerodesmosine (HHMD). It is notable that the Hyl^{ald} and Lys^{ald} derived HLNL are structural isomers, therefore they both co-elute and are indistinguishable by the analysis methods commonly utilized (reviewed in (6,24)).

1.4.2 Mature trivalent cross-links

In aging collagen, the levels of immature reducible cross-links were found to decline over time along with solubility of collagen while the strength of the tissue were shown to increase (25-27). These findings had led to the speculation of the presence of the maturational product of the immature cross-links. To date, in mineralized tissues, two major forms of the trivalent cross-links have been identified and characterized. In bone, unlike other tissues, it was shown that the level of immature cross-links remain relatively high, throughout life (28). The possible reasons could be, in part, that bone is under constant remodeling therefore new collagen is always formed. In addition, it is suggested that the mineralization process may inhibit the maturation of bivalent into trivalent cross-links (29)

Pyridinium cross-links (Pyridinolines). These acid stable, ultraviolet sensitive fluorescent cross-links are predominant in collagen of mineralized tissues (e.g. bone, cartilage) in which Lys residues in the telopeptides are highly hydroxylated. Two major forms have been identified which include pyridinoline (Pyr, or known as hydroxylysyl

pyridinoline (HP)), and deoxy-pyridinoline (d-Pyr, or known as lysyl pyridinoline (LP)). It was pointed out that the term deoxy-pyridinoline may be misleading, since the formation occurs with the Lys rather than Hyl and may not involve the loss of water in the condensation reaction, as the name suggested (24). Pyridinoline, first identified by Fujimoto et al in bovine Achilles tendon, was shown to link between two Hyl^{ald} and one Hyl (Hyl^{ald} X Hyl^{ald} X Hyl) (30). Deoxy-pyridinoline ties between two Hyl^{ald} and one Lys (Hyl^{ald} X Hyl^{ald} X Lys) and was thought to originate from the reaction of LKNL with Hyl^{ald} (31). It is found mostly in mineralized tissues such as bone and dentin (32), and suggestive of having some role in the mineralization process (33).

Several mechanisms have been proposed for the formation of Pyr (Hyl^{ald} X Hyl^{ald} X Hyl). Eyre and Oguchi suggested that it links three collagen molecules and forms from the condensation reaction between two keto-amines, 2 X HLKNL (Hyl^{ald} X Hyl) with the release of one Hyl residue(34). Alternatively, a condensation between one keto-amine and one free Hyl^{ald} from the same collagen molecule has been proposed (35) which is suggestive that it ties only between two collagen molecules. Another possible pathway is proposed through the condensation reaction between deH-DHLNL and its ketoamine form, HLKNL (36)

Pyrroles. This unstable compound, which reacted to p-dimethylaminobenzaldehyde and called Ehrlich chromogen, was identified by Kuyper et al from bovine tendon collagen (37) and later isolated from bovine bone collagen (38). This trivalent cross-link was found exclusively at the N-telopeptide region and cross-linked between Hyl^{ald} X Lys^{ald} X Lys or Hyl (37-39). The presence and abundance of pyrrole cross-links have been correlated to the strength of tissues since decreased levels are found with increasing age and in osteoporotic bone (reviewed in (39)).

1.5 Type I collagen-based mineralization

The mineralized tissues (e.g. bone, dentin, enamel and cementum), are formed from a complex series of events regulating the interaction between the two major factors i.e. the inorganic mineral phase and the organic matrix. From the dry weight of the mineralized tissue, the mineral phase constitutes about two-thirds of the mass while the other one-third is the organic phase, of which 85-90% is collagen; 10-15% is cells and non-collagenous proteins. The mineral phase provides stiffness while the collagen fibers provide the ductility and ability to absorb energy to bone tissue. Unlike other mineralized tissues, the major organic component in enamel is not collagen but amelogenin and enamelin (40). Type I collagen matrix serves as a structural framework to support the nucleation and growth of mineralized nodules. In specific tissue or areas where mineralization occurs, the highly saturated level of the mineral ions (e.g. calcium and phosphate ions) has to be maintained by specific regulatory mechanisms. Mineral deposition occurs both within and between the fibrils, initially forming intermediate precursors (e.g. amorphous calcium phosphate (ACP), or, dicalcium phosphate dihydrate (DCPD)) and gradually mature into the stable form, hydroxyapatite ($(Ca_{10}(PO_4)_6(OH)_2$)/HAP) (41). Several studies have tried to elucidate the regulatory mechanisms of the initiation and progression of collagen-based mineralization, however, the data obtained from those studies cannot give enough clarifications. By using the calcifying avian tendon, a model for mineral formation in vertebrate calcifying tissues, and an atomic force microscopy (AFM), Siperko and Landis in 2001, have shown that collagen-based mineralization is an ordered and tightly regulated process. As depicted in Figure 1.3, after the spontaneous assembly of collagen molecules, mineral deposition induces the formation of crystal platelets in the gaps of the hole and overlap zones, which continuously

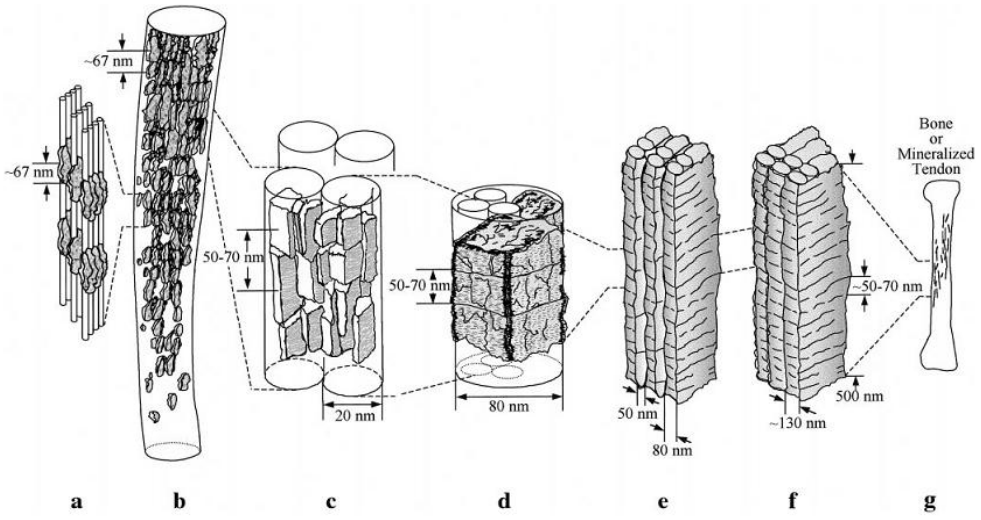

Figure 1.3) Schematic diagram of the putative organization and interaction of collagen and mineral at different structural levels of hierarchy for a calcified vertebrate tissue. (a) Based on the model of collagen assembly proposed by Hodge and Petruska (1963) and the manner of its association with mineral described by Weiner and Traub (1986), McEwen et al. (1992), and Landis et al. (1993, 1996a), crystal platelets nucleate in collagen channels or gaps created by periodic (67 nm) hole and overlap zones. (b) Crystal platelets grow in length along their crystallographic c-axes and in width through the channel or gap spaces (Landis et al., 1993). (c) As collagen macromolecules grow as well to microfibrils and fibrils (20 nm in diameter), crystals fuse into larger and thicker plates in which their periodic deposition (50–70 nm) and parallel nature are still maintained (Landis et al., 1996a). At this stage, interfibrillar plates may be developing. (d) The crystal plates grow larger in all dimensions at the level of collagen fibers (80 nm and greater in diameter). (e) Fibers next associate to create a series of parallel plate aggregates that may vary in length, width, and thickness. Some of the aggregates may represent the interfibrillar plates formed earlier. A frequent thickness observed for the aggregates is 80 nm and a frequent size along their length is 500 nm. Individual plate aggregates are initially separated by 50 nm (e), but this space gradually disappears as mineral deposition proceeds in the tissue and aggregates would grow to thicknesses of 130 nm (f). Edges of plate aggregates still maintain 50- to 70-nm periodicity, indicative of the basic collagen structure underlying mineral formation. (g) Ultimately the plate aggregates become lamellar in shape and constitute a portion of bone or mineralized tendon. Independent of the mineralization associated with the hole and overlap zones, there is surface mineralization (Nylen et al., 1960; Landis et al., 1996b) of the collagen structures in b–f. This aspect of mineral formation is not depicted in b–d. (e) In the schematic (e) is equivalent to Fig. 3b on 90° rotation. Reprinted from J Stuct Biol, 2001, Vol 135 (3), Siperko L.M. & Landis W.J. Aspects of mineral structure in normally calcifying avian tendon, p.313-20, Copyright (2001), with permission from Elsevier.

grow in size through the gap and along the axis of the fibrils in all dimensions to the level of collagen fiber (a-d). Association of collagen fibers, with the ongoing process of mineral

deposition, generates parallel plate aggregates that finally represent the lamellar pattern in bone or mineralized tissue (e-g) (42). Lees in 2003 had proposed that mineralization of collagen rely largely on the molecular packing and chemical structure of collagen molecules. The author has shown that the deposited minerals replace the same amount of water removed from the collagen fibrils (43).

For this reason, a proper formation of the collagenous template is indeed crucial for the mineralization process and alteration in the collagen fibril formation by any factors (e.g. intrinsic or extrinsic factors) may affect the mineral deposition. Defective collagen matrix from structural alteration of the collagen molecules results in abnormal mineralization (18,44), as shown in the studies of osteogenesis imperfecta (OI) (45-46), osteoporosis (47). It has also been observed in several studies performed in our lab that, smaller collagen fibrils formed by mouse osteoblastic cell line, MC3T3-E1, overexpressing decorin (DCN) or lysyl hydroxylase 2b (LH2b) are not efficient templates for mineral deposition (48-49). LH2b is known for hydroxylating the lysine (Lys) residues in the telopeptide regions of collagen while DCN, a member of the SLRPs family, is a negative regulator of collagen fibrillogenesis and mineralization (48,50-51).

1.6 Hydroxylysine glycosylation in collagen

This type of post-translational modification is specific to collagens and proteins with collagen-like triplet repeats, Gly-X-Y, and resulting in two forms of glycosylated Hyl which are galactosylhydroxylysine (G-Hyl) and glucosylgalactosylhydroxylysine (GG-Hyl). The formation involves three sequential enzyme-mediated steps of Lys hydroxylation, Hyl galactosylation and finally, G-Hyl glucosylation, as shown in figure 1.4.

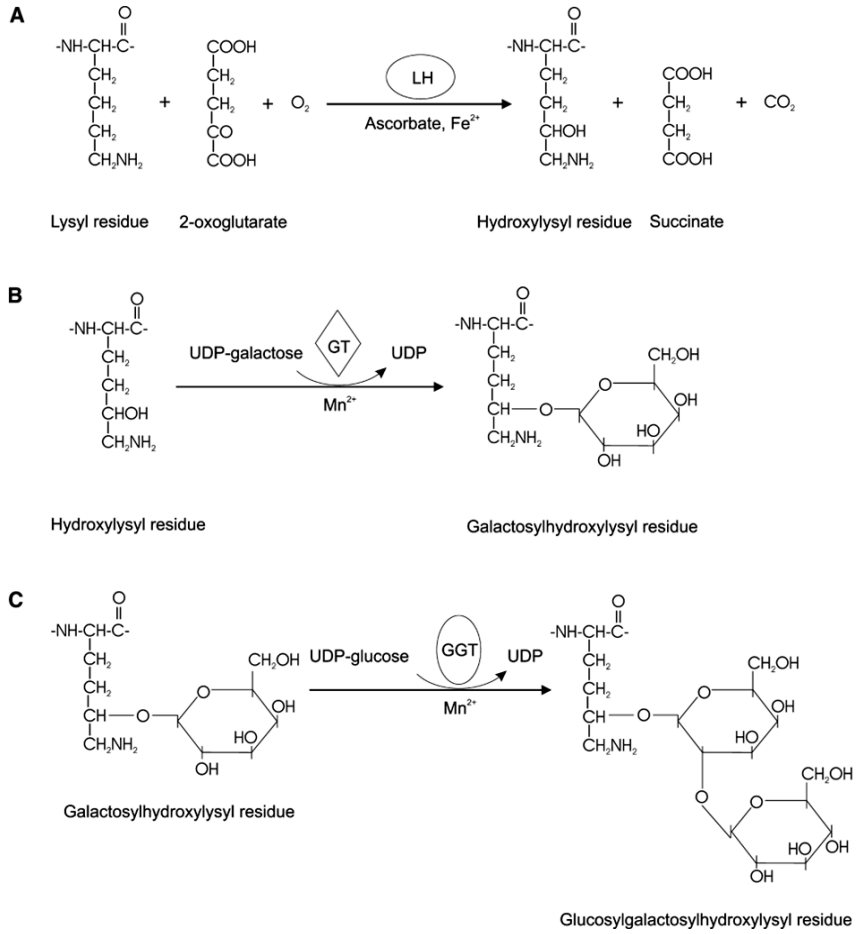

Figure 1.4) Enzymatic reactions catalyzing Lys hydroxylation and subsequent glycosylation. Reaction catalyzed by (A) lysyl hydroxylase (LH), (B) collagen galactosyltransferase (GT), and (C) collagen glucosyltransferase (GGT). Reprinted from J Cell Physiol, 212(2), Myllyla et al, Expanding the lysyl hydroxylase toolbox: New insights into the localization and activities of lysyl hydroxylase 3 (LH3), p.323-9, Copyright (2007), with permission from John Wiley & Sons, Inc.

1.6.1 Lysine hydroxylation

Enzymes and reactions. Lysyl hydroxylase (LH) (EC 1.14.11.4) is one of the 2-oxoglutarate dioxygenases which catalyzes the hydroxylation reaction of Lys residues in the procollagen α chains and proteins with collagenous sequences, co- and post-translationally. The reaction requires ferrous iron, oxygen, 2-oxoglutarate and ascorbate as cofactors and in turn releases succinate and carbon dioxide (CO_2) along with the formation of Hyl (Figure 1.4A) (52). In collagens, Hyl residues are present exclusively in the Y position of X-Y-Gly

triplets while in the non-helical region of α chains, Gly is replaced by serine in the N-telopeptide (X-Hyl-Ser) and by alanine in the C-telopeptide (X-Hyl-Ala) of the α chains. These findings were suggestive of the presence of different LH isoforms, which were later identified and characterized. Some of the Hyl residues, exclusively in the helical region, serve as attachment sites for carbohydrate units, galactose (G-) or glucose-galactose (GG-) forming G-Hyl and GG-Hyl respectively. All of the Hyl species were shown to participate in the formation of intermolecular cross-links. The level of Lys hydroxylation varies among different collagen types, same collagen types in different tissues, developmental stages and pathological conditions (reviewed in (3,6,52).

Lysyl hydroxylase isoforms. The differences in the sequence-specific sites of Lys hydroxylation and the varied levels of Hyl found in different tissues and collagen types had prompted the search for the other LH isoforms and to date 3 of them were identified, LH1, LH2 and LH3 encoded by procollagen-lysine 1, 2-oxoglutarate 5-dioxygenase (PLOD; PLOD1, PLOD2 and PLOD3 respectively). Unlike the other two isoforms, the gene encoding LH2 is alternatively spliced leading to two different variants designated as LH2a and LH2b (or the short and long form, respectively) (53). LH2b contains 63 bp of an extra exon named 13A (53), from which the amino acid sequence were shown to be conserved between human, mouse and rat (54). Among the two variants, LH2b was shown to be the major isoform in most of the tissues (53), including bone (55). The three LH isoforms were cloned and characterized in human (56-57), mouse (58) , rat (54) and zebrafish (59-60) whereas one orthologue is found in *C. elegans* (61). From phylogenetic analysis, among the three

isoforms, LH3 is the oldest isoform showing the most homology to the ancestral LH gene in *C. elegans*, while LH1 and LH2 are derived from more recent gene duplications (58).

Substrate specificity of lysyl hydroxylase isoforms. Several studies have tried to determine the sequence or collagen type specificity for the activities of the three LH isoforms by means of different substrate types (62-63), correlation of the mRNA levels with collagens (64), measuring the levels of end products (55,65), evaluating the different responses to LHs inhibitor (66), collagen analysis from patients with known disorders related to LHs (67), or LHs deficient mice (68-70). Collectively, however, it was shown that there are no strict collagen types or sequence requirements for each LH isoforms. It was suggested that LH2b mainly hydroxylates Lys residues in the telopeptide regions (54-55,62,71-72) while LH1 preferentially catalyzes the hydroxylation of Lys residues in the triple helical region of fibrillar collagen, especially the residues involved in the intermolecular cross-links (62,65-66,69,73). As for LH3, its contribution in hydroxylating the Lys residues in the helical region of fibrillar collagen is still unclear, but it showed its critical roles for type IV and type VI collagen biosynthesis (62,69-70). In the condition that the function of one isoform was impaired, compensatory mechanism from the other 2 isoforms were speculated, and may depend in most part on the tissue-specific expression of each isoform (68).

Subcellular localization of lysyl hydroxylase isoforms. Lysyl hydroxylase 1 has been localized to the endoplasmic reticulum (ER) with a unique mechanism from the 40 amino acid stretch at the C-terminus of the LH1 sequence, unlike the other proteins containing the known ER-retention motifs, KDEL or double lysine. This unique sequence is present with

high degree of homology in all the three LH isoforms indicating that all the isoforms employ the same ER-retention mechanism (74-75).

Unexpectedly, LH3 can be localized both inside and outside of the cell, since it can be detected in the cell lysate, at the cell surface and in the cultured medium. The extracellular LH3 is of higher molecular size than the intracellular LH3 from the asparagine-linked (N-linked) glycosylation during the secretion process. The potential roles of the secreted LH3 have been shown by the ability of recombinant LH3 to modify extracellular proteins (e.g. serum proteins, secreted proteins or proteins on the cell surface) in their native state, suggesting a possible catalytic activity for native collagen (76) which contradicts the earlier reports that triple helical conformation inhibits further modifications (77-78). However, the specific substrates (e.g. collagens or proteins with collagenous sequences) for the extracellular function of LH3 have never been reported.

Moreover, to date, LH3 is the only enzyme modulating the transfer of glucose units to G-Hyl residue, while there are some conflicting data regarding the galactosylation activity of LH3 in transferring galactose units to free Hyl (70,79-81). The glycosyltransferase activity of LH3 will be reviewed in more detail in section 1.6.2.

Lysyl hydroxylase isoforms and human disorders. Mutations in *PLOD1*, which led to a deficiency in the LH1 enzyme, were shown to be the cause of Ehlers-Danlos syndrome EDS type VIA, a disorder characterized by neonatal muscle hypotonia, progressive kyphoscoliosis, abnormal scarring and easy bruising of the skin, increased risk of arterial rupture and ocular fragility (82-87). From these patients, there was a significant decrease in the level of Hyl in the skin collagen and the cross-linking pattern was also altered (73,88-89).

In addition, mutations in *PLOD1* were also found in patients affected with Nevo syndrome, a rare, autosomal recessive disorder with clinical features similar to those of EDS VIA (90). *Plod1* null mice, recently generated, showed muscular hypotonia, increased risk of aortic rupture, and alteration in collagen fibrils. A decrease in Lys hydroxylation and a change in the cross-linking pattern were found to be varied among different tissues of the knockout mice, showing the most significant changes in the skin rather than the other tissues, similar to the phenotypes seen in EDS VIA patients (68). Interestingly, in another subtype of EDS type VIB, *PLOD1* was not mutated and LH activity and cross-linking profiles were normal, however decreases in the mRNA expression of LH2 and LH3 have been observed (91-92). It has been reported, most recently that mutations in the *CHST14*, encoding dermatan-4-sulfotransferase 1 (D4ST-1) involved in the biosynthesis of the dermatan sulfate glycosaminoglycan (GAG) chains, may be one cause of EDS type VIB (93-94). In addition, another novel form of EDS showing similar clinical features with the kyphoscoliotic form was designated as spondylocheiro dysplastic form of EDS (SCD-EDS). Interestingly, in this form of EDS, Lys and Pro residues in the patients' collagen were underhydroxylated despite the normal LH and P4H activities measured from the cell homogenates. It was reported that the disorder has been attributed to a mutation in the Zn^{2+} transporter gene, *SLC39A13*. The mutation most likely resulted in the build up of Zn^{2+} in the ER, which competitively inhibited the function of Fe^{2+}, the essential cofactor for LH and P4H activities (95).

Mutation in *PLOD2*, resulting in underhydroxylation of Lys residues in the telopeptide region of collagen, have been identified in patients with Bruck syndrome type 2, an autosomal recessive inherited disorder, characterized by fragile bones, congenital joint contractures, scoliosis and osteoporosis (72,96-97). However, Bruck syndrome type 1 was

identified in one family showing the link to chromosome 17p12, while no mutation was found in *PLOD2*, suggesting heterogeneity of the syndrome (67). Most recently, Kelley et al have shown that mutations in *FKPB10* (also known as *FKPB65*), encoding a chaperone protein the FK506 binding protein 10, located in chromosome 17p12, are the causes of Bruck syndrome type 1 (98). To date, the generation of *Plod2* manipulated mice model, has not been reported.

The biological significance of *PLOD3* has been illustrated from the LH3-knockout mice which were embryonically lethal at E9.5 from abnormal type IV and type VI collagen biosynthesis and distribution which resulted in the fragmentation of the basement membrane (69-70,99). Recently, mutations in *PLOD3* have been associated with human disorders affecting several connective tissues (100-101). (detailed description in section 1.6.2)

1.6.2 Hydroxylysine glycosylation

Enzymes and reactions. Glycosylated hydroxylysines are generated by the group of enzymes known as collagen glycosyltransferases, which transfer the carbohydrate units (D-galactose and D-glucose) O-glycosidically to the hydroxyl group of specific Hyl residues in the α chains. They consist of hydroxylysyl galactosyltransferase (GT) (EC 2.4.1.50) and galactosylhydroxylysyl glucosyltransferase (GGT) (EC 2.4.1.66) which form galactosylhydroxylysine (G-Hyl) and glucosylgalactosylhydroxylysine (GG-Hyl) respectively. Structural analysis of GG-Hyl, by Spiro in 1967, showed that galactose is attached to the hydroxyl group of Hyl by a β-glycosidic bond, while glucose is linked by an α-glycosidic bond to C-2 of the galactose unit forming the complex structure of 2-O-α-D-glucopyranosyl-O-β-D-galactopyranosylhydroxylysine (Figure 1.5) (102). Both reactions

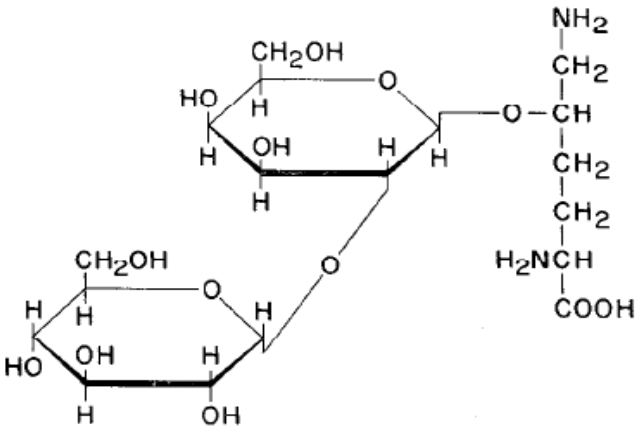

Figure 1.5) The structure of di-glycosylated hydroxylysine. The di-glycosylated hydroxylysine was characterized from glomerular basement membrane showing the structure of 2-*O*-α-D-glucopyranosyl-*O*-β-D-galactopyranosylhydroxylysine. This research was originally published in Jounal of Biological Chemistry. Spiro R.G. The structure of the disaccharide unit of the renal glomerular basement membrane. *J Biol Chem*. 1967; 242: 4813-23. © The American Society for Biochemistry and Molecular Biology.

require the presence of free ε-amino group of the Hyl residue, the UDP-galactose or UDP-glucose as carbohydrate donors, and the most preferred bivalent cation cofactor, Mn^{2+} (103). The complete triple helix structure of collagen molecule inhibits further glycosylation (77-78). These enzymes have been extensively studied and characterized from tissue extracts of different species (104-107), while the highest activities were obtained from whole chick embryo extract. About a decade ago, these enzymatic activities were found in the multifunctional enzyme, LH3 (79-80,108).

The role of LH3 as collagen glycosyltransferase. Lysyl hydroxylase 3 and *C. elegans* LH, the only ortholog in nematode, were both shown to be multifunctional enzymes containing the LH, GT and GGT activities in vitro (108). The GT and GGT activities of LH3 were identified by the ability of the recombinant LH3 generated from SF9 cells and the product of cell-free translation of LH3 cDNA in transferring the sugar moieties to Hyl

residues in calf skin gelatin substrate (79-80). The GGT activity was also confirmed by the partial inhibition of the activity from the binding of the polyclonal antibody of LH3, as well as the antibody raised against chicken galactosyl glucosyltransferase enzyme (79). In LH3, the domain and specific sequences for the GT and GGT activity (aspartic acid (D) repeats or DXD-like motifs) are located in the N-terminus while those for the LH activity are in the C-terminus. The LH and the glycosyltransferase enzymatic activities are shown to be independently regulated (108). The biological significance of the GGT activity of LH3 has been demonstrated from the studies of LH3 deficient mice which had an embryonic lethal phenotype at E9.5 due to the impaired type IV collagen biosynthesis and secretion resulting in fragmentation of the basement membrane. The severity of the phenotype showed a significant correlation to the level of GGT activity (69-70) which was further shown to be essential for the oligomerization and secretion of type VI collagen (99). As for the GT activity of LH3, when compared to the GGT, it has been shown to be significantly lower (79-80) or undetectable (70,81). The biological significance was always in question until recently, a novel family of collagen galactosyltransferase enzyme encoded by the *GLT25D1* and *GLT25D2* has been discovered and characterized in regards to the tissue distribution and enzymatic activity but the substrate specificity was not clearly defined (81).

As mentioned in section 1.6.1, LH3 is also localized and functional outside the cells, since it can be detected in the culture medium, serum and at the cell surface (76). However, the specific substrates (e.g. collagens or proteins with collagenous sequences) for the extracellular LH3 are still not known. By overexpressing the glycosyltransferase-deficient fragment of LH3 or silencing the gene in HT1080 cell, exclusively producing type IV collagen, abnormal cell morphology, growth and eventually cell death were observed.

Interestingly, when the purified glycosyltransferase-deficient fragment was added to the HT1080 cells exogenously, similar effects on the morphology and vitality of the cells were seen faster. However, the effects on the cells were reversible upon removal of the mutated fragments. The effects in the other cell types studied (adult lung fibroblasts, adult skin fibroblasts and osteosarcoma cells), were very mild, if any, indicating cell-type specificity, probably in relation to the receptor protein composition on the cell membrane (109). Moreover, Risteli et al in 2009 reported that lowered level of LH3 resulted in decreased deposition, impaired organization of collagens, and cell morphology changes from the alteration in the organization of cytoskeletal proteins (101).

Identification of novel collagen galactosyltransferases. Most recently, Schegg et al in 2009 have cloned and characterized novel collagen galactosyltransferase enzymes, named glycosyltransferase 25 domain 1 and 2, encoded by the *GLT25D1* and *GLT25D2* respectively, showing that they are capable of transferring galactose, but not glucose units, to collagens and mannose-binding lectin (MBL) (81), a collagen-like protein containing glycosylated hydroxylysines (110). In the same study, LH3 was shown to be the sole GGT enzyme while the GT activity was not detected. Among the two isoforms, GLT25D1 exhibited broad expression in several fetal and adult human tissues, and was suggested to be the main collagen galactosyltransferase enzyme, while GLT25D2 was only detected in a few cell types (81). Unlike LH3, GLT25D1 and GLT25D2 were shown to be strictly maintained in the ER by the ER retention motif, RDEL (81,111). Immunofluorescence analysis in liver cells showed that GLT25D1 co-localizes with MBL and LH3 in the ER (111).

1.6.3 Glycosylated hydroxylysines in collagen fibrillogenesis

To date, the biological roles of glycosylated Hyl residues have not been clearly defined, however it is evidently seen in collagens that, optimal levels of Hyl glycosylations are essential for the formation of normal extracellular matrices hence healthy connective tissues. With their unique characteristics of having bulky sugar units projecting outwards and orienting parallel to the backbone of the collagen molecules (9), they are believed to be one of the factors controlling collagen fibril diameter. Previous studies have shown that collagens with higher level of Hyl glycosylation formed fibrils with smaller diameter *in vitro* (8-11,112), possibly from the extensive amount of sugar projections preventing the regular hexagonal packing of the collagen molecules. Besides, the hydrophilic nature of the sugar units may attract more water to the surface of the collagen molecules, possibly affecting the interactions between collagen molecules or between collagen and its binding proteins.

1.6.4 Glycosylated hydroxylysines in collagen cross-linking

Several studies, analyzing type I collagen from various tissues and species, had identified Hyl residue 87 in both α1 and α2 chains to be mainly di-glycosylated (113-117). In bone type I collagen, the immature bivalent cross-links derived from Hyl residue 87 and Hyl^{ald} in the C-terminal telopeptide were found to be mainly di-glycosylated (reviewed in Table 1.1) (118-120). Interestingly, pyridinoline (Pyr), the trivalent cross-link maturing from the condensation reaction between two DHLNL complexes, predominantly found in human bone was found to be mono-glycosylated (G-Pyr) or non-glycosylated (38,121-122). Analysis of the posttranslational modifications from various functional sites in rat femur has shown an inverse correlation between the level of glycosylated Hyl residues and pyridinium

Table 1.1 Review of glycosylated cross-links identification. (38,119-123)

References	Species/tissue	Types of cross-links	Site / % Glycosylation	Glycosylation types
Robins & Bailey 1974 [118]	bovine bone 1yr	DHLNL	not specified	GG-DHLNL > G-DHLNL
	articular cartilage	DHLNL		GG-DHLNL > G-DHLNL
	rat skin	HLNL		GG-HLNL > G-HLNL + Hexosyl-Hyl
	fetal calf skin	DHLNL		GG-DHLNL (major form)
Henkel et al 1976 [123]	rabbit bone 1yr	DHLNL	α1-16C + α1-87 (80%)	GG-DHLNL : G-DHLNL 1:1
	tendon	DHLNL, HLNL, HHMD	α1-16C + α1-87 (minimal)	not specified
	skin	HLNL, HHMD	α1-16C + α1-87 (100%)	GG-HLNL : G-HLNL 4.3:1
Robins 1983 [121]	sheep diaphyseal	Pyr	not specified	G-Pyr
				GG-Pyr (very low yield)
Hanson & Eyre 1996 [38]	human bone	Pyr	2xα1-16C + α1-87	G-Pyr
Gineyts et al 2001 [120]	human bone	Pyr	not specified	free Pyr > G-Pyr (no GG-Pyr)
	cartilage			free Pyr > minimal GG (No G)
	synovium			GG-Pyr > free Pyr > G-Pyr
Eriksen et al 2004 [119]	human bone	DHLNL	α1-16C + α1-87 (70%)	GG-DHLNL/G-DHLNL 2.0
			α1-16C + α2-87 (61%)	GG-DHLNL/G-DHLNL 1.3
		HLNL	α1-16C + α1-87 (7%)	not specified
			α1-16C + α2-87 (20%)	not specified

Abbreviations: HLNL:hydroxylysino-norleucine, DHLNL:dihydroxylysinonorleucine, HHMD: histidinohydroxymerodesmosine, Pyr: pyridinoline, GG-: glucosylgalactosyl-, G-: galactosyl-, Hyl: hydroxylysine, α1-16C: residue 16 at C-telopeptide of alpha 1 chain, α1-87: residue 87 at helical region of alpha1 chain, α2-87: residue 87 at helical region of alpha 2 chain.

cross-links, especially GG-Hyl and Pyr (124). These observations may implicate a potential negative regulatory role of the glucose units in the GG-Hyl residue for the maturation of collagen cross-links, which occurs spontaneously between two juxtaposed bivalent cross-links. The possible explanations would be, firstly, that the bulkiness of the galactose-glucose attachments may sterically hinder the reaction forming the trivalent cross-links. However, GG-Pyr has been detected in several other tissues (121) therefore, the steric hindrance caused by the glucose unit seems less likely. Secondly, since the presence of GG-Pyr shows high tissue specificity, differences in the micro-environment between tissues may have a selective role in the cross-links maturation process.

1.6.5 Glycosylated hydroxylysines in extracellular matrix formation

Spiro in 1969 reported higher level of GG-Hyl residues in insoluble collagen compared to the soluble one purified from the same tissue (e.g. rat skin or rat tail tendon), suggesting the possibilities that glucosylation may continue after cross-links have been formed or that the collagen molecules with more GG-Hyl residues would be deposited and retained in the matrix (125). Later, it was shown that the transfer of the glucose units to G-Hyl by the glucosyltransferase enzyme only occur before the complete folding of the α chains into triple helical conformation (77), therefore the continued glucosylation of collagen molecules would be less likely. Moreover, it has been reported that collagens with higher levels of glycosylation were shown to have increased resistance to mammalian collagenase digestion (9). In skin fibroblast culture from the patients with dominant epidermolysis bullosa simplex (DEBS), a disorder with defective GGT activity from decreased level of LH3 protein, collagen deposition into the extracellular matrix was reported to be lower when compared to the control (101). Collectively, the data from these previous reports suggested that collagen molecules with higher levels of glycosylation may be deposited into the matrix and stabilized by the formation of cross-links while collagen with lower levels of glycosylation may be less cross-linked and more susceptible to collagenase digestion.

1.6.6 Glycosylated hydroxylysines as markers of collagen metabolism

In the process of collagen metabolism, these glycosylated Hyl residues are not re-utilized by the cell for collagen biosynthesis; therefore, they are secreted into the serum and subsequently excreted with urine (126). Comparing to Hyp, the levels of Hyl-glycosides excreted in urine were not affected by normal diets and shown to quantitatively reflect the

degree of collagen turnover and somehow indicative of the tissue origin. Higher urinary excretion of total glycosylated Hyl has been shown in healthy children and some heritable connective tissue disorders with higher degree of collagen turnover. Among different connective tissue collagens, G-Hyl is the predominant form in bone type I collagen resulting in the uniquely low GG-Hyl to G-Hyl ratio, which was also observed when measured in urine. Therefore, these glycosylated Hyl residues and their ratios were used as biological markers for collagen turnover (127), especially in bone, a marker of growth in children (128), and as a tool to monitor treatment response in patients (129-132). The level of free G-Hyl in serum was also shown to be a useful marker for bone resorption (133). In addition to the direct measurement of the level of glycosylated Hyl residues, the GGT activity in serum has been indicated as a sensitive marker for collagen metabolism in various rheumatic diseases (134) neuromuscular, hepatic and dermatological diseases (135-137), and after physical exercise (138-139).

1.6.7 Glycosylated hydroxylysines and age-related variations

Age-related variation in glycosylation has been observed from the analysis of human and rat skin collagen which demonstrated a rapid decrease of the GG-Hyl/free Hyl ratio and a somewhat constant G-Hyl/free Hyl ratio during maturation, therefore suggesting a functional role for GG-Hyl while G-Hyl serves as an intermediate (140). Cetta et al in 1982 analyzed rabbit Achilles tendon of different ages and showed fast decreases in the level of GG-Hyl and GG-Hyl/G-Hyl ratio during fetal life and in the first month after birth, which corresponds to the rapid increase in the collagen fiber diameter. The markedly decrease level of GG-Hyl, as compared to G-Hyl, points to the possibility of a conversion of GG-Hyl into G-Hyl (141). A

glucohydrolase enzyme, with specific activity for GG-Hyl residues has been purified from chick embryo (142), rat spleen (143), and detected in cell lysate of human skin fibroblast (144). However, characterization of the GG-Hyl α-D glucosidase enzyme from rat kidney has demonstrated that the enzyme has optimal activity on free GG-Hyl and to a lesser degree on GG-Hyl containing peptides and proteins respectively. This could suggest the potential role of the enzyme at the later stage of the collagen metabolism process rather than at the collagen molecules in the extracellular matrix (145).

1.6.8 Glycosylated hydroxylysines and protein interactions

Several studies had identified, in type I collagen from different species, the glycosylation sites (reviewed in Table 1.2) which are noticeably present more to the N-terminal region of the collagen molecule and are located in close proximity to or in the gap zone, according to the functional and ligand binding regions mapped in one D-period repeat containing the complete collagen sequence from elements of 5 monomers and include an overlap and gap zone (146). The gap zone was shown to be the site for the initiation and progression of mineralization (Figure 1.2) (42). The presence of the glycosylated Hyl residues exclusively to this region, may suggest a regulatory role of the sugar attachments for the mineralization process either directly by facilitating the deposition of minerals or indirectly by interacting with other non-collagenous proteins, both of which have never been investigated.

Moreover, those mapped regions were shown to be the binding sites of various extracellular proteins e.g. integrins, proteoglycans, phosphophoryn and discoidin domain receptor 2 (DDR2) (146). The functional association between these proteins and Hyl

Table 1.2 Review of the glycosylated hydroxylysine sites identification. (103,113-116,147)

Reference	Species/tissue	Site/peptide sequence	Types
Butler & Cunningham 1966 [112]	Guinea pig skin collagen (soluble)	α1- 87/GMKGHR	GG-Hyl
Cunningham & Ford 1968 [113]	Guinea pig skin collagen (insoluble)	α1- 87/GMKGHR	GG-Hyl
Morgan et al 1970 [114]	Guinea pig skin collagen Carp swim bladder Human skin collagen	α1-87/GMKGHR α1-87/GMKGHR, α2-87/GIKGHR α1-87/GMKGHR, α2-87/GIKGHR GPK	GG-Hyl G-Hyl
Volpin & Veis 1973 [145]	Bovine skin (Insoluble)	α1-CB5 (87-123) α1-CB6 (820-1011) α2-CB4(7-326) α2-CB(3-5)(356-1011)	GG-Hyl G-Hyl GG-Hyl GG-Hyl
	Bovine dentin (Insoluble)	α1-CB5 (87-123) α1-CB3 (403-551) α1-CB6 (820-1011) α2-CB4(7-326) α2-CB(3-5)(356-1011)	GG-Hyl G-Hyl G-Hyl G-Hyl/GG-Hyl G-Hyl/GG-Hyl
Aguilar et al 1973 [115]	Rat skin collagen (soluble & insoluble)	α-1 : 87/GMKGHR (major) α-1 : GMK, GPK (minor) α-2 : 174/GPKGEL α-2 : 87/GFKGIR	GG-Hyl G-Hyl G-Hyl G-Hyl/GG-Hyl
Fietzek & Kuhn 1976 reviewed in Kivirikko 1979 [102]	Calf skin collagen	α-1 : 87/GMKGHR α-1 : 684/GPK α-2 : 174/GPKGEL, 87/GFKGIR α-2 : 219/GAK	G-Hyl/GG-Hyl G-Hyl/GG-Hyl G-Hyl/GG-Hyl G-Hyl or GG-Hyl *

*: type not specified, CB: cyanogens bromide peptide, G: glycine, M: methionine, K: lysine, H: histidine, R: arginine, I: isoleucine, P: proline, E: glutamic acid, L: leucine, F: phenylalanine, A: alanine G-Hyl: galactosylhydroxylysine, GG-Hyl: glucosylgalactosylhydroxylysine,

glycosylation has never been elucidated except for DDR2, a tyrosine kinase receptor specific for collagen, which its activation was reported to be significantly reduced after deglycosylation of the collagen by sodium-m-periodate (148). Moreover, it was demonstrated that the degree of collagen glycosylation affects the arthritogenicity to type II collagen. The incidence and severity of the arthritis showed strong correlation to the level of G-Hyl residues (149).

1.7 Collagen glycosylation-related human disorders

1.7.1 Levels of collagen glycosylation

The levels of collagen glycosylation have been shown to vary among different collagen types (5,150), different functional regions in the same tissue (124,151), different tissues, developmental stages (103,152-153) and diseased conditions (100-101,154-160). Among different collagens, the abundance of free Hyl and glycosylated Hyl residues in the tissue varies with the distinct expression and the organization degree of collagen types. The highest content has been reported in the less organized collagen type IV-rich basement membrane, in a disaccharide form (GG-Hyl). Lower levels of Hyl glycosides are found in collagen types that form organized fibrils such as type I, II and III (150), among which, type I collagen has the lowest level of glycosylation. The abundance of the two forms of modified Hyl residues (GG-Hyl, G-Hyl and their ratio) also shows varying levels in the same collagen type from different tissues e.g. bone, skin and tendon (127,161). Besides, species specific difference in the GG-Hyl to G-Hyl ratio has been observed between bone type I collagen from human and rat (162). The degree of Hyl glycosylation has been shown to have a significant contribution towards the quality of extracellular matrices synthesis, organization and maturation, as seen from several connective tissue disorders in which different levels of collagen glycosylations were observed.

1.7.2 Human disorders with collagen underglycosylation

Underglycosylation of collagen, as a result of the deficiency in the GGT enzymatic activity in the skin and serum, was detected in a group of patients from one family with dominant epidermolysis bullosa simplex (DEBS), characterized by the occurrence of serous,

nonscarring blistering of the skin after minor trauma (154). Most recently, the cause for the defective GGT activity in this specific subtype of DEBS has been reported, showing four polymorphic nucleotides in the promoter and 5' untranslated region of the PLOD3 gene but no mutation in the coding region, resulting in about 67% LH3 mRNA expression (101). In another recent finding, two heterozygous mutations in the PLOD3 coding region, which resulted in a significant decrease in GGT activity, have been associated with a rare syndrome of congenital malformations affecting several tissues and organs and shared common features with several other known collagen disorders (100). Biochemical analysis of collagens from mice models for Kashin-Beck disease, a rare degenerative, osteoarticular disorder associated with severe skeletal deformation and dwarfism, showed elevated Lys hydroxylation in bone type I collagen, and to a lesser degree in type II collagen which also showed a slight increase in the level of glycosylated Hyl residues (163). Interestingly, significant decrease in the levels of GG-Hyl but not G-Hyl residues were reported in bone type I collagen of mice receiving long term supplementation of fulvic acid, shown to be one of the main causes for Kashin-Beck disease, (164).

1.7.3 Human disorders with collagen overglycosylation

Overglycosylation of bone type I collagen, on the other hand, has been associated with poor bone quality as seen in some osteogenic disorders i.e. osteogenesis imperfecta (OI) (156,160), post-menopausal osteoporosis (155,158,165-166) osteosarcoma, and osteofibrous dysplasia (159). In addition, higher level of glycosylated Hyl and GG-Hyl to G-Hyl ratio were observed in the granulation tissue while lower levels were found in normal skin and scar tissue respectively, suggesting the role of Hyl glycosylation in the wound healing

process (167). Brinckmann et al in 1999 has shown an increased level of Hyl glycosylation as a result of overhydroxylation of skin type I collagen from the patients suffering chronic venous insufficiency which resulted in lipodermatosclerosis, a condition characterized by excessive accumulation of extracellular matrices and leading to hardening of the skin (157).

Chapter 2

Specific Aims

In type I collagen, the presence of glycosylated hydroxylysines, i.e. galactosylhydroxylysine (G-Hyl) and glucosylgalactosyl hydroxylysine (GG-Hyl) are long known, and their level alteration has been associated with several human connective tissue disorders. However, the enzymatic mechanism for the formation of G-Hyl and GG-Hyl in type I collagen has not been clearly defined. Moreover, the biological roles of these Hyl species in type I collagen still remained elusive. Lysyl hydroxylase 3 (LH3), the multifunctional enzyme containing both Lys hydroxylation and Hyl glycosylation activities, is designated as the major collagen glucosyltransferase enzyme with some conflicting data with its collagen galactosylation activity. From the previous publications on LH3, data regarding its role for type I collagen has been very limited. Preliminary studies from our laboratory have shown that LH3 is expressed in mouse osteoblast cell line, MC3T3-E1 (MC), mainly producing type I collagen. Therefore, the studies to be shown here in this dissertation were aimed to uncover the role of LH3 for type I collagen by utilizing MC3T3-E1 cell line as a model. The specific aims were as follow:

1. To elucidate the enzymatic function of LH3 for bone type I collagen by generating MC cell-derived clones stably suppressing LH3 gene (Sh clones) with short hairpin RNA technology.

2. To investigate the effect of LH3 suppression on the level of Lys hydroxylation and Hyl glycosylations in type I collagen, collagen fibrillogenesis, collagen cross-linking, collagen matrix formation and mineralization.
3. To identify specific molecular loci and types of Hyl glycosylation in MC and Sh collagen with mass spectrometry.

During the course of this study, novel collagen galactosyltransferase enzymes, glycosyltransferase 25 domain 1 and 2 (GLT25D1 and D2) were identified and characterized. Among the two isoforms, we have determined that only Glt25d1 is expressed in MC cells (Chapter 3: Study I). Therefore, an additional aim was:

4. To elucidate the function of Glt25d1 for bone type I collagen by generating MC cell-derived clones stably suppressing Glt25d1 (G_Sh clones) with short hairpin RNA technology and characterize their collagens.

Chapter 3

Study I

LYSYL HYDROXYLASE 3 GLUCOSYLATES GALACTOSYLHYDROXYLYSINE RESIDUES IN TYPE I COLLAGEN IN OSTEOBLAST CULTURE

Marnisa Sricholpech[1], Irina Perdivara[2], Hideaki Nagaoka[1], Megumi Yokoyama[1], Kenneth B. Tomer[2] and Mitsuo Yamauchi[1]

[1] NC Oral Health Institute, School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, and [2] Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA.

Abstract

Lysyl hydroxylase 3 (LH3), encoded by *Plod3*, is the multifunctional collagen modifying enzyme possessing the lysyl hydroxylase (LH), hydroxylysyl galactosyltransferase (GT) and galactosylhydroxylysyl glucosyltransferase (GGT) activities. Though an alteration in type I collagen glycosylation has been implicated in several osteogenic disorders, the role of LH3 in bone physiology has never been investigated. In order to elucidate the function of LH3 in bone type I collagen modifications, we have employed a short hairpin RNA technology in mouse osteoblastic cell line, MC3T3-E1 (MC), generated single cell-derived clones stably suppressing LH3 (Sh clones) and characterized the phenotype. *Plod3* expression and the LH3 protein levels in the Sh clones were significantly suppressed when compared to the controls, MC and the clone transfected with an empty vector (EV). In comparison to controls, type I collagen synthesized by Sh clones (Sh collagen) showed a significant decrease in the extent of glucosylgalactosylhydroxylysine with a concomitant increase of galactosylhydroxylysine while the total number of hydroxylysine residues was essentially unchanged. In an *in vitro* fibrillogenesis assay, Sh collagen showed accelerated fibrillogenesis compared to the controls. In addition, when recombinant LH3-V5/His protein was generated in 293 cells and subjected to GGT/GT activity assay, it showed GGT but not GT activity against denatured type I collagen. The results from this study clearly indicate that the major function of LH3 in osteoblasts is to glucosylate galactosylhydroxylysine residues in type I collagen and that lowering of LH3 level significantly affects type I collagen fibrillogenesis.

Introduction

Collagens are a large family of extracellular matrix proteins comprising at least 29 different genetic types (*1-2*). Among those types, fibrillar type I collagen is the most abundant protein and it is the major structural component in most connective tissues including bone. One of the critical steps in collagen biosynthesis, which contributes to the functional integrity of the tissues, is the post-translational modifications including the hydroxylation of specific proline (Pro) and lysine (Lys) residues, glycosylation of specific hydroxylysine (Hyl) residues, and the formation of covalent intermolecular cross-links. Though several functions have been proposed for collagen glycosylation such as control of collagen fibrillogenesis (8-11,112) , cross-linking (38,118,121-124,168), remodeling (126-133) collagen-cell interaction (148-149) (*reviewed in section 1.6.3-1.6.8*), the function is still not well defined due in part to the lack of clear understanding of the mechanism of this modification.

In fibrillar collagens, glycosylation occurs at specific Hyl residues by hydroxylysyl galactosyltransferase (GT) (EC 2.4.1.50) and galactosylhydroxylysyl glucosyltransferase (GGT) (EC 2.4.1.66) resulting in the formation of galactosylhydroxylysine (G-Hyl) and glucosylgalactosylhydroxylysine (GG-Hyl) respectively. Recently, these enzymatic activities were found in the multifunctional enzyme, lysyl hydroxylase 3 (LH3), one of the collagen modifying enzymes in the lysyl hydroxylase (LH) family, encoded by *PLOD3* (79-80,108).

The GT and GGT activities of LH3 were identified by measuring the transfer of radiolabeled sugar moieties to Hyl residues in calf skin gelatin substrate (79-80). The biological significance of the GGT activity of LH3 in the biosynthesis of types IV and VI collagen has been demonstrated from the studies of LH3 deficient mice (69-70,99). In

addition, defective GGT activity, resulting from four polymorphic nucleotides in the non-coding region of *PLOD3*, was detected in a family with dominant epidermolysis bullosa simplex (DEBS) (101,154). In another recent finding, two heterozygous mutations in the *PLOD3* coding region, have been associated with a rare syndrome of congenital malformations affecting several connective tissues and organs (100). As for the GT activity of LH3, the data has been somewhat conflicting and the biological significance is still not clearly understood. This activity could be performed by recently reported novel enzymes, glycosyltransferase 25 domain 1 and 2 encoded by *GLT25D1* and *GLT25D2*, respectively (81) *(reviewed in section 1.6.2).*

It has been reported that altered collagen glycosylation is associated with bone disorders (155-156,158-160,165-166,169) *(reviewed in section 1.7).* However, the role of LH3 in type I collagen modification and bone biology have not been investigated.

In order to elucidate the function of LH3 for bone type I collagen, we have employed an *in vitro* loss-of-function approach in mouse osteoblast-like cell line, MC3T3-E1 (MC) by generating single cell-derived clones stably suppressing LH3. Type I collagen produced by the clones were analyzed for the extent of modifications (e.g. GG-Hyl, G-Hyl and free Hyl) and fibrillogenesis. In addition, recombinant LH3-V5/His fusion protein was generated and its enzymatic functions were investigated by means of a newly developed high performance liquid chromatography (HPLC)-based glycosyltransferase activity assay.

Experimental Procedures

Cell Lines and Culture Conditions

MC3T3-E1 subclone #4 (MC), a well characterized nontransformed mouse osteoblast-like cell line (170) was purchased from American Type Culture Collection (CRL-2593). The cells were grown in α-MEM (Invitrogen, Carlsbad, CA, USA) containing 10% FBS (Invitrogen) and supplemented with 100 U/ml penicillin G sodium, 100 μg/ml streptomycin sulfate. 293 HEK (Human Embryonic Kidney) cells were purchased from Clontech (Mountain View, CA, USA). The cells were maintained in Dulbecco's Modified Eagle Media (DMEM; Invitrogen) with high glucose (4.5g/l) and same supplements as described above. Both cell lines were cultured in a 5% CO_2 atmosphere at 37ºC and the medium was changed twice a week.

Expression of Plod Family, Glt25d1 and Glt25d2 in MC cells

MC cells were plated at a density of $2x10^5$ cells/35mm dish and after 48 hours, total RNA was extracted with TRIzol reagent (Invitrogen), and two μg of total RNA was converted into cDNA by using the Omniscript Reverse Transcriptase kit (Qiagen, Valencia, CA, USA). Quantitative real-time PCR was performed using the sequence specific primer for LH1 (*Plod1*; ABI assay number: Mm00599925_m1), LH2 (*Plod2*; Mm00478767_m1), LH3 (*Plod3*; Mm00478798_m1), Glt25d1 (*Glt25d1*; Mm00600638_m1), and Glt25d2 (*Glt25d2*; Mm00625070_m1). The reactions were prepared and analyzed in triplicates by the ABI Prism 7000 Sequence detection system (Applied Biosystems, Foster city, CA, USA). The mRNA expression level of LH3 relative to glyceraldehyde 3-phosphate dehydrogenase (GAPDH; ABI assay number 4308313) was analyzed by the $2^{-\Delta\Delta CT}$ method (171).

Generation of short hairpin expressing plasmid targeting Plod3

The target sequences for suppression of *Plod3* were designed by siDesign Center website (Dharmacon RNAi technologies). The three highest ranked targets were selected (si#1 AGAAGGAAATGGAGAAATA [start position 956], si#2 CCACAAGGGTGTAGATTAT [2715], si#3 GAACAAAACAGGAAGGTAA [1897]) and converted into short hairpin oligonucleotide sequences specific for pSilencer2.1 vector by the siRNA converter (Ambion, www.ambion.com). This method added BamHI and HindIII restriction site overhangs at the 5' end of the sense and antisense oligonucleotide respectively, to facilitate efficient directional cloning into the pSilencer vectors. The two complimentary 55-mer oligonucleotides for each target were synthesized (MWG Eurofins, Huntsville, AL, USA), and annealed in the annealing buffer (10mM Tris pH 7.5-8.0, 50mMNaCl, 1mM EDTA) according to the condition described in the instruction manual for pSilencer2.1 vector. The annealed oligonucleotides were then ligated into the pSilencer2.1-U6/neo vector (Ambion, Austin, TX, USA, a generous gift from Dr. Franck Polleux, Neuroscience Center, University of North Carolina at Chapel Hill), linearized with BamHI and HindIII restriction enzymes digestion, and transformed into bacteria competent cells. The clones with 100% nucleotide sequence accuracy (Eton Bioscience, Durham, NC, USA) for each target (Sh1, Sh2, and Sh3) were selected and tested for their suppression efficiency by transient transfection into MC cells using FuGENE6 transfection reagent (Roche, Indianapolis, IN, USA). The non-transfected MC cells and those transfected with the original pSilencer2.1-U6/neo plasmid (EV; encoding a hairpin siRNA sequence not found in any genome database) were used as controls. After 48 hours, total RNA was extracted, converted into cDNA and analyzed for the expression of *Plod3* by quantitative real-time PCR as

described above. The suppression efficiency among the different ShRNA target sites was then compared in percentage relative to the controls.

Generation and characterization of MC single cell-derived clones stably suppressing LH3

MC cells were transfected with the selected pSilencer2.1-U6/neo-Plod3 plasmid by using FuGENE6 transfection reagent (Roche). After 48 hours, the transfected cells were trypsinized, plated at a low density and maintained under constant selection with 400 μg/ml G418 sulfate (G418; Mediatech, Inc., Manassas, VA, USA). Several single cell-derived Sh clones were isolated by cloning rings and maintained in α-MEM, 10% FBS, 100 units/ml penicillin, 100 μg/ml streptomycin along with 400 μg/ml G418. The single cell-derived EV clones were isolated and maintained in the same manner as described above. The selected Sh clones were then characterized by comparing the level of *Plod3* expression to those of the EV clone and the non-transfected MC cells, by using quantitative real-time PCR. Three to six Sh clones with different levels of *Plod3* suppression were then characterized for its effects on the mRNA expression of LH1 (*Plod1*; ABI assay number: Mm00599925_m1), LH2 (*Plod2*; Mm00478767_m1), Glt25d1 (*Glt25d1*; Mm00600638_m1), and type I collagen α2 chain (*Col1A2*; Mm00483888_m1). The effect on cell proliferation was assessed by using CellTier 96 MTS assay (Promega, Madison,WI, USA) as previously reported (172).

Generation of anti-mouse LH3 antibody

Polyclonal rabbit anti-mouse LH3 antibody was commercially generated (Bethyl Laboratories, Inc., Montgomery, TX, USA) by immunization with the HPLC purified, synthetic peptide, corresponding to the amino acids 357 to 371 (FSAVKLVGPEEALSA)

conjugated to KLH. The agarose-bound peptide was used for the affinity purification of the antibody from the hyperimmune rabbit sera obtained from the immunized animal. The sensitivity and specificity of the antibody was determined by Western Blot analysis of purified recombinant LH3-V5/His tagged protein, as described below.

Immunoprecipitation and Western Blot analysis

To determine the LH3 protein level, the Sh clones and the controls were plated onto 35mm dishes at a density of $3x10^5$ cells/dish. After culturing for 7days, the cultured medium was collected and the cells were washed with PBS, lysed with lysis buffer containing 150 mM NaCl, 20 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1% TritonX-100, 1% deoxycholate, 1.5% aprotinin, and 1 mM phenymethylsulfonylfluoride, centrifuged and the supernatant collected. The cultured medium was incubated overnight with anti-LH3 antibody (see above), followed by incubation with protein G-sepharose 4B conjugate beads (Invitrogen) for 30 minutes, washed with lysis buffer, solubilized in SDS sample buffer (100 mM Tris-HCl, pH 8.8, 0.01% bromphenol blue, 36% glycerol, and 4% SDS) in the presence of 10 mM dithiothreitol, and separated by 4–12% NuPAGE Bis-Tris gel (Invitrogen). As for the cell lysates, the total protein concentration was measured by the DC protein assay kit (Bio-Rad Laboratories, Hercules, CA, USA) according to the manufacturer's protocol, and 20 μg of total protein was directly applied to 4-12% NuPAGE Bis-Tris gel. The separated proteins were transferred to a polyvinylidene fluoride (PVDF) membrane (Immobilon-P; Millipore Corp., Bedford, MA, USA) and subjected to Western blot analysis with anti-LH3 antibody followed by incubation with anti-rabbit IgG conjugated to alkaline phosphatase (AP) secondary antibody (Thermo Scientific, Rockford, IL, USA). The immunoreactivities were

detected with AP conjugate substrate kit (BioRad Laboratories). Equal protein loading of cell lysate was confirmed by Western blotting with anti-β-actin antibody.

Purification of type I collagen

The same number of cells of MC, EV and Sh clones were plated on to 10-cm dishes and maintained with twice-a-week change of medium (α-MEM , 10% FBS, 100units/ml penicillin, 100μg/ml streptomycin). When the cells reached confluence, 50μg/ml of ascorbic acid was added and cultured for at least 2 weeks, the cell/matrix layer was then collected and subjected to type I collagen extraction as previously described (173). Briefly, the cell/matrices were washed with PBS and cold distilled water, lyophilized and subjected to limited pepsin digestion (20% w/w) in 0.5N cold acetic acid with constant stirring at 4˚C for 24 hours. After ultracentrifugation at 35,000 rpm for 2 hours, the supernatant was collected and type I collagen was precipitated with 0.7M NaCl. The precipitate was then reconstituted with 0.5N acetic acid, dialyzed against cold distilled water for 48 hours at 4˚C and lyophilized. Collagen solution was prepared by reconstituting the lyophilized collagen in 0.01N acetic acid to the concentration of 0.5 μg/μl. The purity of type I collagen obtained from the control cells and clones was assessed by 6%Tris-Glycine gel electrophoresis (Invitrogen) stained with Coomassie Brilliant Blue (CBB) and by amino acid analysis (see below).

In vitro fibrillogenesis assay

The procedure previously described (12) was slightly modified for the use of a 96-well microtiter plate. Briefly, lyophilized type I collagen purified from each group was

redissolved in 0.01 M acetic acid at a concentration of 0.2µg/µl. The reactions were prepared in triplicates, in a 96-well microtiter plate pre-cooled on ice, containing the reaction buffer (140mM NaCl, 30mM sodium phosphate (PBS) at pH7.3) (174) and purified type I collagen in the ratio of 1:1 (v/v), and placed in the Powerwave X 340 spectrophotometer (BioTek, Winooski, VT, USA) equilibrated at 37˚C. Formation of collagen fibrils was monitored by measuring the absorbance at 400nm (turbidity), at 1000 seconds intervals for 15 readouts. The turbidity-time curve was plotted and the differences in the rate of collagen fibril formation and the potential fibril diameter among the clones and the controls were compared.

*Purification of GG-Hyl and G-Hyl standard*s

Natural marine sponge was used to isolate GG-Hyl residues (175-176). Fifty milligrams of sponge was pulverized in a Spex freezer mill (SPEX CertiPrep, Metuchen, NJ, USA), subjected to base hydrolysis with 2N NaOH in sealed polypropylene tubes flushed with N_2 gas, and incubated at 110°C for 22 hours. The hydrolysate was acidified to a ~pH 3 with 2N HCl, and filtered. It was then applied to a standardized molecular sieve column filled with Bio-Gel® P-2 resin (Bio-Rad Laboratories) equilibrated in 0.1N acetic acid buffer at a flow rate of 0.7ml/min and monitored by absorbance at 214 nm (177). Several collagen cross-link standards were used as molecular weight standards to isolate the GG-Hyl (MW: 487 Da) -enriched fractions. The GG-Hyl residues eluted slightly earlier than the Pyr standard (MW: 428 Da). In order to generate the monoglycosylated form, G-Hyl, the purified GG-Hyl was subjected to mild acid hydrolysis with 0.8N HCl at 110°C for 1 hour. The resulting G-Hyl was purified using the P-2 column as described above. The identities of the purified GG-Hyl and G-Hyl were confirmed by mass spectrometry.

Mass Spectrometry analysis of glycosylated hydroxylysine standards

The fraction containing GG-Hyl, G-Hyl and free Hyl from the mild acid hydrolysis was analyzed by nano-ESI-MS on a Waters QTof Global mass spectrometer (Milford, MA). A 1:100 dilution (v/v) in a solvent mixture containing 50 % methanol in deionized water acidified with 0.1 % formic acid was used for flow injection analysis. The mass spectra were acquired over the mass range *m/z* 150-1000. The source parameters employed were as follows: capillary voltage – 3.5 kV, cone voltage - 80 V, collision energy 10 V. The MS/MS spectra were acquired using targeted analysis of the ions of interest. The collision energy was optimized for each ion, ranging from 10 to 15 V for the singly protonated ion of G-Hyl, and from 15-25 V for the singly protonated ion of GG-Hyl.

Quantification of GG-Hyl, G-Hyl and Free Hyl

Lyophilized collagen samples were hydrolyzed with 6N HCl (Sequenal Grade; Thermo Scientific) *in vacuo* after flushing with N_2 at 110°C for 22 hours. The hydrolysates were subjected to amino acid analysis on a Varian 9050 HPLC system (9050/9012; Varian Associates Inc., Palo Alto, CA, USA), equipped with a strong cation exchange column (AMINOSep #AA-911, Transgenomic, Omaha, NE, USA) (173). The level of total Hyl (GG-, G- and free Hyl) in a collagen molecule was calculated based on the value of 300 residues of Hyp per collagen molecule.

To determine the level of glycosylated Hyl, lyophilized collagen samples were also subjected to base hydrolysis and the hydrolysates were prepared in the same manner as described above. Then the GG-, G- and free Hyl in the hydrolysates were separated and quantified with a specific gradient program (Table 3.1) developed on the same HPLC system

mentioned above. Separation of the different Hyl species from one another and from other amino acids was achieved in 75 minutes. The color factor of Hyl was used to calculate the quantity of GG-Hyl and G-Hyl (141,178-179). Since Hyp is partially degraded during base hydrolysis, the amount of total Hyl/collagen was determined by acid hydrolysis (see above). Based on the ratio of the three Hyl forms (GG-, G- and free Hyl) and the total Hyl values obtained from acid hydrolysis, each form of Hyl was quantified as residues/collagen molecule.

Molecular cloning of mouse LH3 cDNA

Total RNA was isolated from MC cells using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. The first-strand cDNA was synthesized from 2 μg of total RNA by means of the Omniscript Reverse Transcriptase Kit (Qiagen). Specific primers for the coding sequence of mouse LH3 (GenBank accession NM011962) were designed as follow: forward primer including restriction enzyme sequence for EcoRI, 5'_GCGAATTCCGGGCCTGGGTGTAGCCT_3'; and reverse primer including the sequence for XhoI, 5'_GCCTCGAGGGGGTCAACAAAGGACACCATG_3'. PCR amplification was performed by the ProofStart DNA polymerase (Qiagen) with an annealing temperature of 60.0°C for 40 cycles. After digesting the PCR products with EcoRI and XhoI restriction enzymes, the inserts were ligated into the pcDNA3.1/V5/His_verA vector (Invitrogen) generating the pcDNA3.1/LH3-V5/His constructs. The molecular size of the ligated insert of 2389 bp was analyzed by EcoRI and XhoI restriction enzymes digestion followed by 2% agarose gel electrophoresis. Plasmids containing digested inserts were

sequenced at the UNC-CH DNA Sequencing Facility and showed 100% homology to the mouse LH3 in the database.

Purification of LH3-V5/His fusion protein

293 HEK cells were transfected with pcDNA3.1/LH3-V5/His plasmid using FuGENE 6 transfection reagent (Roche) according to the manufacturer's instructions, and the cells were cultured in the presence of 400 μg/ml of G418 (Mediatech, Inc.) to isolate stably transfected single cell-derived clones. Based on Western Blot analysis with anti-V5-AP antibody, the clone that synthesized the highest level of LH3-V5/His protein was selected and cultured to collect the cultured medium. LH3-V5/His protein was purified using a nickel-nitrilotriacetic acid agarose resin (Qiagen), and eluted with elution buffer containing 0.2M NaCl, 0.1M glycine, 0.1% glycerol, 50mM urea, 20mM Tris, and 0.05M histidine, pH 7.5. The purified proteins were pooled, dialyzed, concentrated with centrifugal filter with MWCO 30,000 (Centricon; Millipore Corp.), and stored at -20 °C until use. The protein concentration was measured using a DC protein assay kit (Bio-Rad Laboratories) and the purity of the purified LH3-V5/His protein was determined by 4-12% NuPAGE Bis-Tris gel and CBB staining. In order to confirm its identity, the gel-resolved proteins were transferred to a PVDF membrane (Immobilon-P; Millipore Corp.) and subjected to Western blot analysis with anti-V5-AP antibody (Invitrogen) or anti-LH3 antibody (Bethyl Laboratories, Inc.) with or without the epitope peptide (CFSAVKLVGPEEALSA). The immunoreactivities were detected in the same manner as described above.

HPLC-based collagen glycosyltransferase assays

Pepsinized bovine skin collagen solution (*PureCol*® 3mg/ml) was purchased from Advanced BioMatrix (Advanced BioMatrix, Tucson, AZ, USA). Denatured collagen substrate was prepared by incubating the *PureCol*® solution in distilled water at 60°C for 30 minutes and rapidly cooled on ice before use. The *PureCol*® solution without heat denaturation was used as a native collagen substrate. The assay for GT or GGT reaction was performed in pre-cooled 0.65ml microcentifuge tubes containing 120μg native or denatured collagen substrate, 60μM UDP-Glc or UDP-Gal (Calbiochem / EMD Chemicals, Gibbstown, NJ, USA), 12μl of reaction buffer (16.7mM MnCl2, 1mM DTT, 2mg/ml BSA, 0.1M sodium bicarbonate buffer pH7.8), recombinant LH3-V5/His protein and distilled water in a total volume of 100μl. The tubes were incubated at 37°C for 45 minutes and the reactions were terminated by placing the tubes on ice. The reaction products were dried by speed vacuum centrifugation (Savant Instrument Inc., Hicksville, NY, USA), hydrolyzed with base and analyzed for the level of GG-Hyl, G-Hyl and free Hyl as described above.

Statistical analyses

Statistical analyses were performed using Jmp®8.0 software (SAS Institute Inc., Cary, NC, USA). Statistical differences were determined by Kruskal-Wallis one-way analysis of variance, means comparison to the controls by Dunnett's Method and means pair comparison by Student's t test. The data were presented as means ± SD and a p value less than 0.05 was considered significant.

Results

Expression of Plod family, Glt25d1 and Glt25d2 in MC cells

At 48hrs of MC cells culture, the expressions of *Plod1-3* and *Glt25d1-2* were analyzed by quantitative real-time PCR. By normalizing to *GAPDH*, the results showed that *Plod3* was expressed at the highest level among the *Plod* family. In addition, *Glt25d1* was highly expressed in MC cells while *Glt25d2* was undetectable (Figure 3.1). The expression of *Plod3* and *Glt25d1* remained high and that of *Glt25d2* was undetectable up to 4 weeks of culture (data not shown). Among these genes, the highest expression was constantly observed for *Plod3* ($p < 0.05$).

Generation of clones stably suppressing Plod3

By transiently transfecting MC cells with pSilencer2.1-U6/Neo plasmids encoding hairpin oligonucleotides targeting three different sites in *Plod3*, the expression levels of *Plod3* relative to *GAPDH* from all three targets were significantly suppressed. In comparison

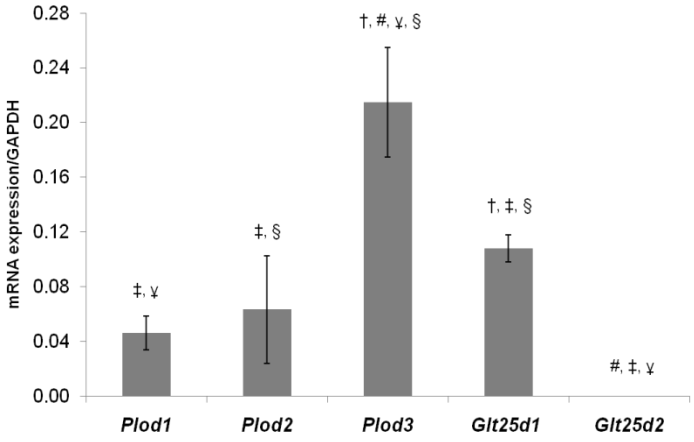

Figure 3.1) Gene expression of *Plod* and *Glt25* families in MC cells. After 48 hours of culture of MC cells, the mRNA expression of *Plod1*, *Plod2*, *Plod3*, *Glt25d1* and *Glt25d2* relative to internal control, *GAPDH*, was analyzed by quantitative real time PCR. Note that the expression of *Glt25d1*, not *Glt25d2*, is expressed in MC cells. Error bars indicate mean ± SD of three independent experiments. Significantly different from †: *Plod1*, #: *Plod2*, ‡: *Plod3*, ¥: *Glt25d1*, §: *Glt25d2*. ($p < 0.05$)

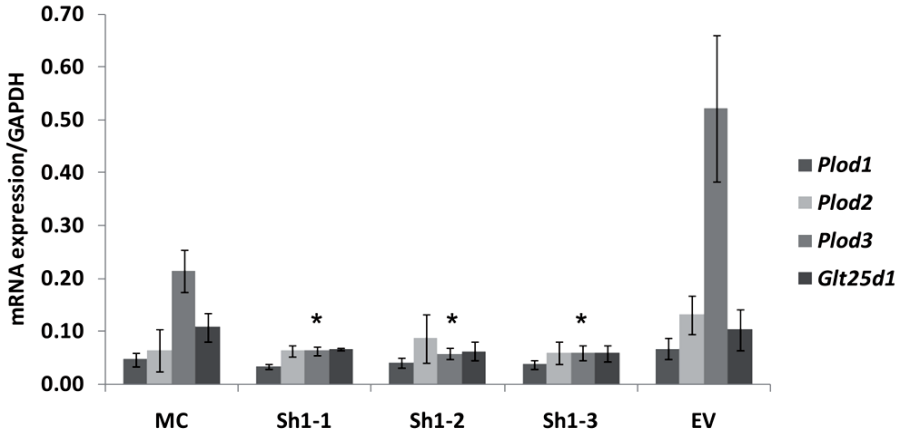

Figure3.2) Gene expression of *Plod* family and *Glt25d1* in single cell-derived Short hairpin (Sh) clones. After 48 hours of culture, real time PCR was performed to determine the suppression levels of *Plod3* in the single cell-derived Sh clones and its potential effects on *Plod1*, *Plod2* and *Glt25d1* expression. The mRNA expression was normalized to the internal control, *GAPDH*. Only the expressions of *Plod3* in the Sh clones were significantly different when compared to the controls, MC and EV. Error bars indicate mean ± SD of three independent experiments. *, significantly different (p<0.0001) from MC and EV.

to the non-transfected MC cells (100%), the expression levels of *Plod3* in Sh1, 2 and 3 were 20%, 22% and 46%, respectively. The *Plod3* expression in the transfected control (EV) was ~80% of that of MC cells. The suppression levels of *Plod3* among the different targeting sequences were essentially the same when normalized to *Col1A2* (data not shown). The targeting sequence number one (Sh1), showing the most suppression, was selected to generate single cell-derived clones stably suppressing LH3 by constant selection with G418. Several clones were isolated and characterized.

Characterization of Sh1-derived clones stably suppressing LH3

Three representative clones (Sh1-1, 1-2 and 1-3) stably suppressing *Plod3*, that exhibited similar cell proliferation patterns to that of controls, were selected for further characterization. Quantitative real-time PCR analyses, from three independent experiments,

showed statistically significant decreases in the expression of *Plod3* in all Sh clones (p<0.0001), compared to the EV and MC controls (Figure 3.2). The higher expression of the *Plod3* in the EV clone compared to MC parental cell was constantly observed during the isolation and characterization of the EV clones, likely due to the effect of transfection. The effect of *Plod3* suppression on the expression of the other *Plod* family members and *Glt25d1* was also investigated. When normalized to *GAPDH*, the expressions of *Plod2* and *Glt25d1* were not significantly changed in all the three Sh clones when compared to the controls. *Plod1* expression was also mostly unaffected except for Sh1-1 in that the expression of *Plod1* was slightly lower when compared to EV (p<0.05) but not significantly different from MC.

The LH3 protein levels from the clones were assessed by direct Western blot analysis for the cell lysate and immunoprecipitation/Western blot analysis for the cultured medium (Figure 3.3). In Figure 3.3A, an immunoreactive band was observed at the expected molecular weight of LH3 at ~ 87kDa in the cell lysates of the controls and the Sh clones; however the intensity was decreased in all Sh clones. Comparable loading amounts of the cell lysates was confirmed by Western blotting detection of β-actin. Figure 3.3B shows the

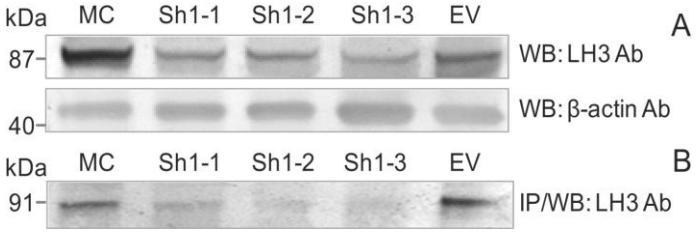

Figure 3.3) The levels of LH3 protein produced by the Sh1 derived clones and controls
A) Western blot analyses of cell lysates with anti-LH3 antibody and anti-β-actin antibody. Twenty μg of total protein from each sample was loaded. B) Immunoprecipitation-Western blot analyses of the cultured media with anti-LH3 antibody. The intensity of the immunoreactive bands from Sh1-1,-2, and-3 were significantly lower than that of the controls, MC and EV, in both cell lysate and cultured media. WB: Western blotting, IP: Immunoprecipitation, Ab: Antibody.

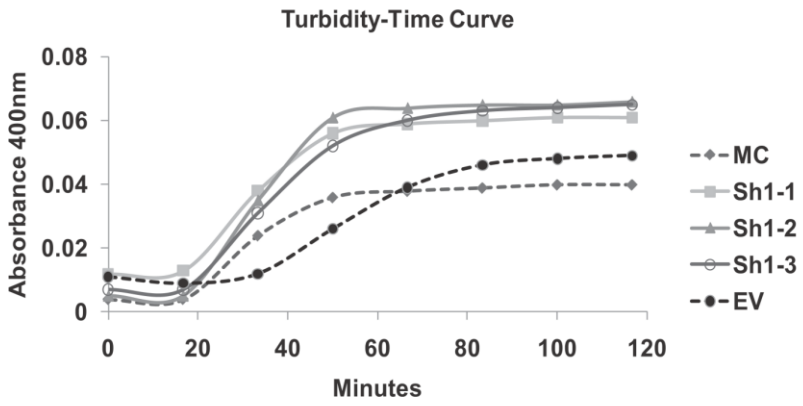

Figure 3.4) *In vitro* fibrillogenesis assay
After 2 weeks of culture, type I collagen without telopeptides was purified from the cell/matrix layers of MC cells, Sh (1-1, 1-2 and 1-3) and EV clones. Twenty µg of purified collagens (0.2µg/µl) were incubated in a 96-well microtiter plate at 37°C and the 400nm absorbance (turbidity) was measured (see Experimental procedures for details). The kinetics of the collagen fibrils formation are shown as a turbidity-time plot.

result of immunoprecipitation followed by Western blot analysis with anti-LH3 antibody from the cultured media collected from the clones and controls. Compared to the controls, all three Sh clones showed lower intensity of the immunoreactive bands observed at ~91 kDa.The apparent molecular weight of these bands was slightly higher than those observed in the cell lysate (~87 kDa). This observation is consistent with a previous report showing that the extracellular LH3 protein is modified by asparagine-linked glycosylation during the secretion process (76). Together, it is evident that suppression of *Plod3* by short hairpin RNA expression plasmid resulted in decreases in the levels of LH3 protein, both intra- and extracellularly.

Fibril formation kinetics of purified type I collagen

Type I collagen without telopeptides was purified from 2 weeks cell/matrix layer of the MC, EV and Sh1-1, -2 and -3. The purity of the collagens was determined by 6% Tris-

Glycine gel electrophoresis, where only the collagen associated chains (α chains with some β and γ components) were identified (data not shown). In addition, amino acid analyses showed approximately 100 residues Hyp and ~340 residues of glycine/1000 amino acids in MC, ShEV and all the Sh clones indicating a high level of collagen purity. The purified collagen from the clones and controls were subjected to an *in vitro* collagen fibrillogenesis assay to compare the fibril formation kinetics among the groups. Figure 3.4 showed the turbidity-time curve plotted from the 400nm absorbance measured over time. Comparing to the controls, MC and EV, the collagen from all three Sh clones showed shorter lag time and higher turbidity plateau indicating faster rate of fibril formation and potentially larger collagen fibrils diameter, respectively (180). The result shown is the representative of several independent assays.

Purification and NanoESI-Qtof analysis of GG-Hyl and G-Hyl standards

The GG-Hyl and G-Hyl residues were purified from the base hydrolysate of natural marine sponge by molecular sieve chromatography. The identity of the purified GG-Hyl residues was indirectly confirmed through the conversion of GG-Hyl into G-Hyl and free Hyl by mild acid hydrolysis. The upper panel in Figure 3.5A shows the chromatogram of the potential GG-Hyl residue separated on a strong cation-exchange column detected by the ninhydrin-derivatization method. The program was specifically developed to separate GG-, G- and free Hyl from one another and from other known amino acids. The lower panel shows the same compound after the mild acid hydrolysis demonstrating a decrease in the amount of the potential GG-Hyl and the appearance of two additional ninhydrin-positive peaks eluting at later time points. The double peak eluting at ~70 minutes is consistent with the elution of

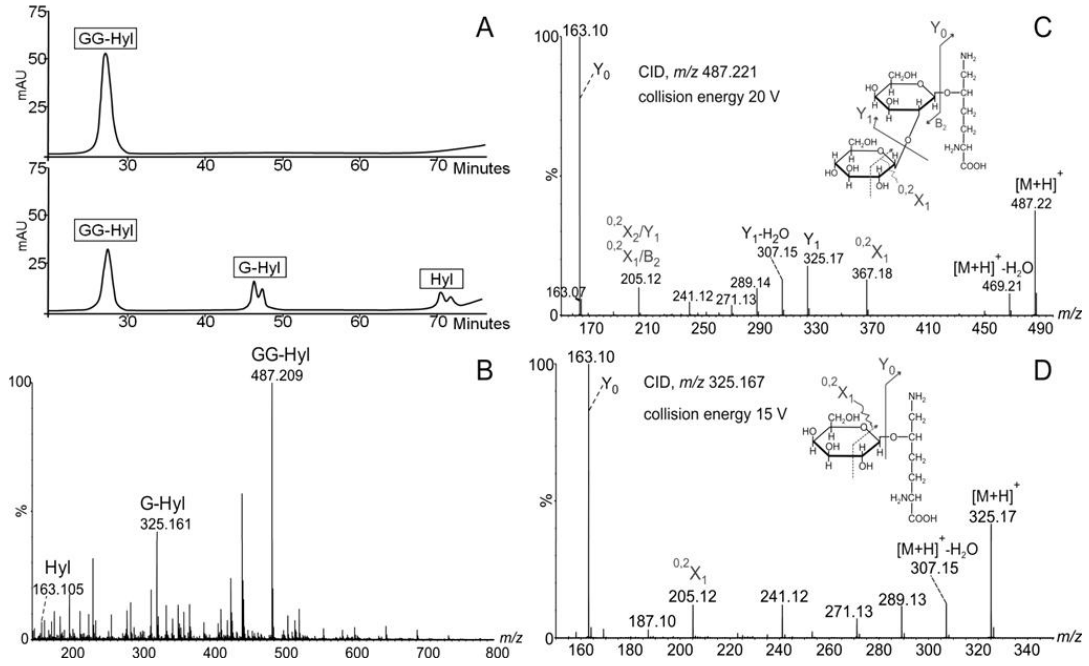

Figure 3.5) Purification and identification of glycosylated Hyl standards.
A) Upper panel: ion-exchange chromatogram of potential GG-Hyl residues isolated from a P-2 column; Lower panel: ion-exchange chromatogram of the same compound after mild acid hydrolysis showing the conversion into free Hyl and an intermediate product, potentially G-Hyl. B) Mass spectrum of the mixture of Hyl, G-Hyl and GG-Hyl from the amino acid analysis, showing the molecular ions corresponding to each species. C-D) MS/MS spectra of the ions of *m/z* 487.221 and 325.167, confirming the structures of GG-Hyl and G-Hyl respectively. Hyl: hydroxylysine, G-Hyl: galactosyl Hyl, GG-Hyl: glucosylgalactosyl Hyl.

the Hyl standard. The double peak eluted at ~47 minutes is most likely the intermediate product in the conversion of GG-Hyl into free Hyl, i.e. G-Hyl residue, based on the physico-chemical characteristics of the molecule and from previous reports (176,178,181-182).

To confirm the identity, the mild acid hydrolysate containing a mixture of the potential three Hyl species was analyzed on a QTof Global mass spectrometer by nanoESI. The most abundant ion in the MS spectrum shown in Figure 3.5B, observed as *m/z* 487.209 (1+), was assigned to hydroxylysine modified with dihexose, GG-Hyl. The ion corresponding to the G-Hyl species was observed as *m/z* 325.161. Unmodified Hyl was observed as the singly protonated ion of *m/z* 163.105. The structural characterization of each

compound was performed by MS/MS, shown in Figures 3.5C and 3.5D. The fragment ions observed in the MS/MS spectrum of the ion of *m/z* 487.209, obtained with a collision energy of 20 V, are characteristic for the CID spectra of glycoconjugates, confirming the structure GG-Hyl (Figure 3.5C). The fragment ions of *m/z* 325.165 and *m/z* 163.103, assigned to Y_1 and Y_0 (183) are derived from the subsequent neutral loss of each hexose from the non-reducing end, whereas the ion of *m/z* 367.178, assigned to $^{0,2}X_1$ resulted from a cross-ring cleavage in the first monosaccharide ring. The ion of *m/z* 205.116 is most likely formed in a secondary fragmentation event, by cross-ring cleavage within the Y_1 fragment. This hypothesis is supported by the fact that, by increasing the collision energy, this ion is observed with higher relative abundance. The fragment ions observed in the CID spectrum of the ion of *m/z* 325.167, shown in Figure 3.5D, were obtained using collision energy of 15 V. The most abundant ion, *m/z* 163.101, Y_0, is derived from the fragmentation of G-Hyl from the non-reducing end.

Quantification of GG-Hyl, G-Hyl and Free Hyl

To separate and quantify the level of GG-Hyl, G-Hyl and free Hyl, a new program was developed on the amino acid analysis system in our laboratory by modifying the buffer gradient previously reported (173), shown in Table 3.1. The separation of the glycosylated Hyl and the other amino acids standard are shown in Figure 3.6. This specific program can be used to analyze base hydrolysates of collagens, as the carbohydrate units are retained under these conditions. Figure 3.7 shows the levels of the different Hyl species in type I collagen purified from control cells and clones, analyzed in triplicates. The results revealed that there was a significant decrease in the level of GG-Hyl with a concomitant level increase of G-Hyl

Table 3.1 HPLC gradient system for the separation of GG-Hyl, G-Hyl and free Hyl

Minutes	*Buffer A*	*Buffer B*	*Oven Temp.*	*UV Monitor*
Initial	100	0		
1.5	100	0		
3	60	40	60°C	
27	59	41	67 min	
33	43	57		
45	43	57		570nm
46	12	88		
70	12	88	90°C	
85	10	90	43 min	
110	0	100		
120	0	100		

Buffer A: 125 mM tartaric acid disodium, 172 mM maleic acid, 2.5% isopropyl alcohol, pH 2.78
Buffer B: 0.3M NaoH, 107.7 mM maleic acid, 48.5 mM boric acid, pH 9.91

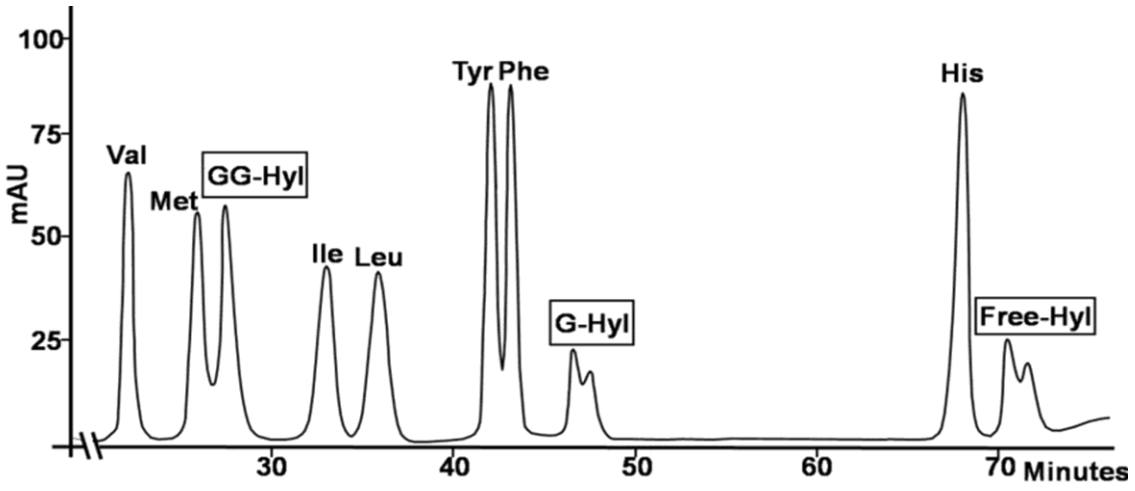

Figure 3.6) HPLC elution profile of GG-, G- and free Hyl relative to other amino acids.
Three-letter abbreviation is used to indicate amino acids.

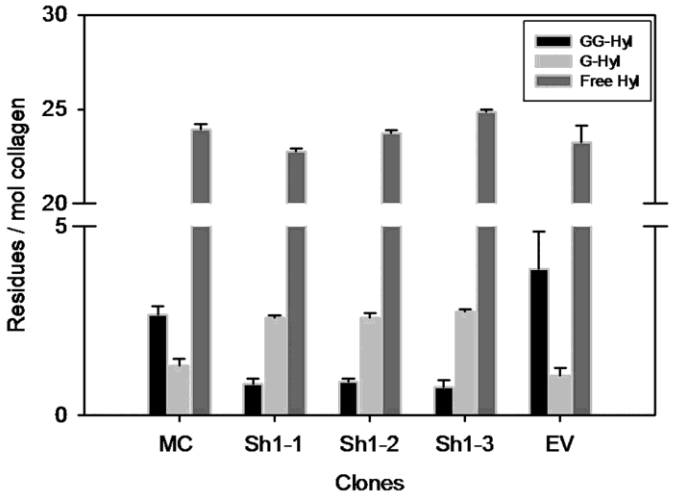

Figure 3.7) The glycosylation levels of Hyl residues in the purified type I collagen synthesized by Sh1 clones and controls.
The levels of GG-Hyl, G-Hyl and free Hyl are shown as residues/mol collagen. Error bar indicate mean ± SD of triplicate analyses of the hydrolysates.

in collagen from Sh clones, while the levels of free Hyl remained comparable between the Sh clones and the controls, MC and EV. This modification profile was also confirmed in type I collagen purified from the other two independent sets of cultures. This clearly indicates that the major role of LH3 for type I collagen in osteoblasts is to transfer glucose units to G-Hyl residues.

Purification of recombinant LH3-V5/His fusion protein

To confirm the glycosyltransferase activity of LH3, recombinant LH3-V5/His fusion protein was purified from the cultured medium of 293 HEK cells. An aliquot from the collected fractions were analyzed by 4-12% NuPAGE Bis-Tris gel stained with CBB and by Western blot analyses (Figure 3.8). By CBB staining, a major single band was observed at ~96kDa, consistent with the expected molecular weight of LH3-V5/His fusion protein (91 kDa of LH3 + 5 kDa tag) (*lane 1*). The identity of the purified protein was further confirmed by Western blot with anti-V5 (*lane 2*) and anti-LH3 antibodies (*lane 3*). When peptide blocking was performed, the intensity of the immunoreactive bands to anti-LH3 antibody

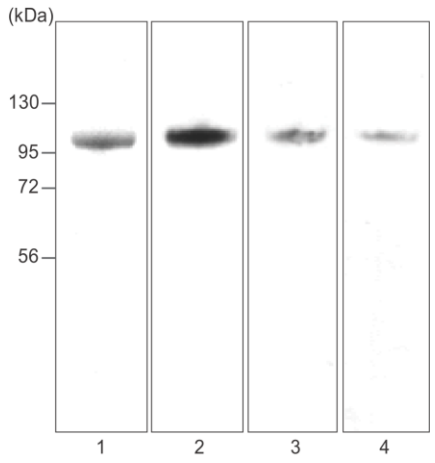

Figure 3.8) Characterization of recombinant LH3-V5/His tagged protein.
Purified LH3-V5/His protein was subjected to 4-12% NuPAGE Bis-Tris gel and stained with CBB (*lane 1*), WB analysis with anti-V5 antibody (*lane 2*), anti-LH3 antibody (*lane 3*) and anti-LH3 antibody with peptide block (*lane 4*). The major band at the level of ~96 kDa showed immunoreactivities to both anti-V5 and anti-LH3 antibody.

was decreased indicating the competitive binding of the peptides to anti-LH3 antibody (*lane 4*). The fractions were then pooled, dialyzed and concentrated.

Glycosyltransferase activity of LH3-V5/His fusion protein

The results of the GT and GGT activity assays of LH3-V5/His fusion protein with bovine skin collagen substrate are shown in Figure 3.9. When LH3-V5/His fusion protein was incubated with denatured collagen substrate, the level of GG-Hyl was significantly increased with a concomitant decrease in the level of G-Hyl (Figure 3.9A) leading to a reverse in the GG-Hyl/G-Hyl ratio (Figure 3.9B). However, the transfer of galactose units to free Hyl was not observed as the GGHyl/GHyl ratio did not change (Figure 3.9B). When native bovine skin collagen was used as a substrate, the GGT activity was not observed (Figure 3.9A-B). To confirm the specificity of the enzymatic functions, the dose-effect of LH3-V5/His protein was examined (Figure 3.9C). With denatured collagen substrate, the change in the GG-Hyl/G-Hyl ratio and the dose-effect were observed only for the GGT activity but not the GT activity. The GGT activity was readily detected with the amount of LH3-V5/His protein as low as 0.1% of the collagen substrate (w/w) indicating strong potency

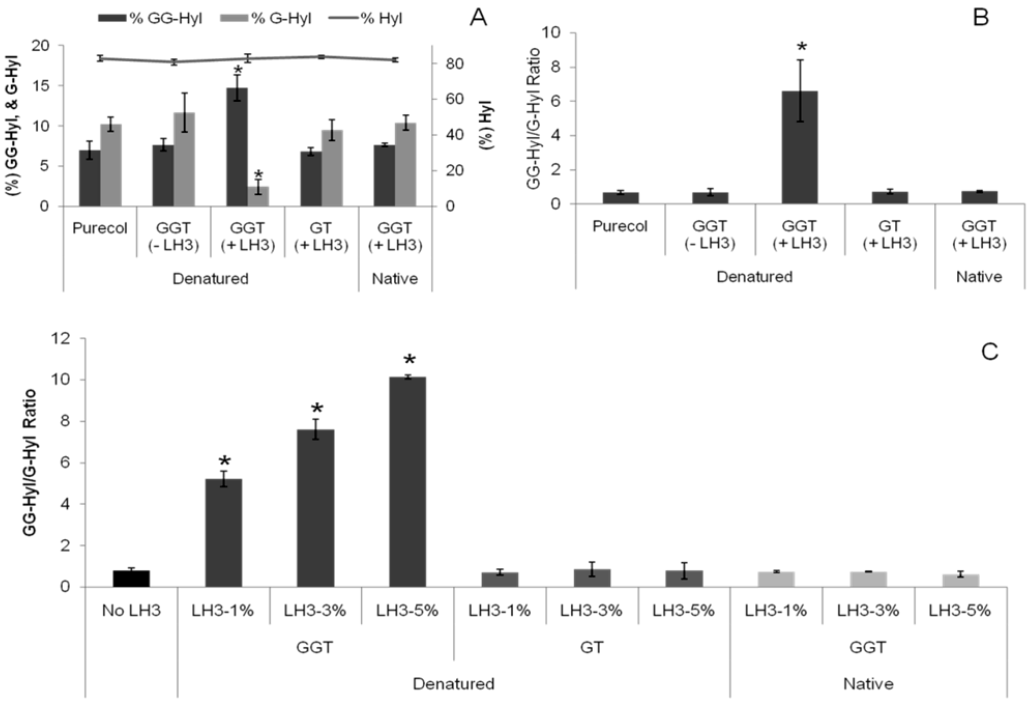

Figure 3.9) HPLC-based glycosyltransferase activity assays of recombinant LH3-V5/His against denatured and native type I collagen. The transfer of galactose or glucose units (GT or GGT respectively) by the purified LH3-V5/His protein was performed with bovine skin collagen (*PureCol*®) substrate. The base hydrolysates of the reaction products were analyzed for the level of G-Hyl or GG-Hyl by HPLC (see Figure 3.6). A) Percentage of GG-Hyl, G-Hyl and free Hyl, B) GG-Hyl/G-Hyl ratios of the reaction products, C) GG-Hyl/G-Hyl ratios of the GGT and GT activity assays using increasing levels of LH3-V5/His proteins with denatured or native collagen substrate. GT: hydroxylysyl galactosyltransferase, GGT: galactosylhydroxylysyl glucosyltransferase, LH3: LH3-V5/His protein. Error bars indicate mean ± SD from triplicate assays. *, significantly different from PureCol (in A & B) and from No LH3 (in C), $p < 0.05$.

of the recombinant LH3-V5/His fusion protein as an enzyme (data not shown). When native collagen was used as a substrate, the GG-Hyl/G-Hyl ratio was not changed with any levels of LH3-V5/His protein examined (Figure 3.9C). Similar results were observed when rat tail tendon collagen was used as a substrate for the GT and GGT activity assays (data not shown). The results shown here indicate that the purified LH3-V5/His protein is enzymatically active with a specific function of glucosylation of G-Hyl residues, and suggest that the specific substrates for LH3 are the collagen α-chains rather than the native collagen

molecules, which is consistent with previous reports showing that the triple helical conformation of collagens inhibited further post-translational modifications (77-78).

Discussion

During the past decade, several reports have shown the unique features of LH3, i.e. its multifunctionality (LH, GT and GGT activities) (79-80,108,184) and its intra- as well as extracellular localization (76). The critical importance of this enzyme for the connective tissue integrity was clearly demonstrated through the gene knockout mouse study and others (69,99-101). The enzymatic functions of LH3 were shown through *in vitro* activity assays using lysates from SF9 cells transfected with LH3 expression plasmid, cell-free translation or purified recombinant LH3 protein as the enzyme source (79-80,108,184). The GGT activity of LH3 and its biological significance have been well demonstrated, however, data regarding the GT activity of LH3 has been somewhat conflicting (70,79-80). Recently, Schegg *et al* have reported the cloning of novel collagen galactosyltransferase enzymes, GLT25D1 and GLT25D2, and showed that these enzymes were capable of transferring galactose units to free Hyl residues in collagen while LH3 was not (81). However, direct measurements of the amount of G-Hyl or GG-Hyl residues were rarely performed in the assays used in the past. In this study, we have successfully purified the G-Hyl and GG-Hyl standards from natural marine sponge and developed a specific HPLC program which allowed us to effectively separate and quantify the levels of GG-Hyl, G-Hyl and free Hyl residues in the base hydrolysate of the collagen samples.

Type I collagen is the major collagen type in most connective tissues including skin, bone and tendon. However, little is known about the role of LH3 in this major fibrillar collagen at present. In this study, by employing the short hairpin RNA technique in mouse osteoblasts, we have successfully generated single cell-derived clones stably suppressing

LH3, and characterized its effect on type I collagen post-translational modifications. Our data demonstrated that the major function of LH3 is to glucosylate G-Hyl residues in type I collagen. The simultaneous increase in the level of G-Hyl residues observed from the LH3 Sh clones is most likely the result of diminished level of glucose transfer by LH3 suppression. The gene expression of GLT25D1, another potential galactosyltransferase, was unaffected by LH3 suppression. In addition, the levels of free and total Hyl residues in the purified type I collagen were essentially unchanged among the Sh clones and controls suggesting that the LH activity of LH3, if any, is minimal for bone type I collagen. In the past years, several groups have tried to determine the sequence or collagen type specificity for the activities of LH isoforms (55,62-70). The results from these studies indicated that there are no strict collagen types or sequence requirements for the individual LH isoforms. It was suggested that LH2b, the major isoform of LH2 in bone, mainly hydroxylates Lys residues in the telopeptide regions, thus, determining the collagen cross-linking pathway (54-55,62,71-72). In contrast, LH1 preferentially catalyzes the hydroxylation of Lys residues in the triple helical region of fibrillar collagen including those residues pairing with the telopeptidyl Hyl/Lys aldehydes to form the intermolecular cross-links (62,65-66,69,73). As for LH3, its contribution to Lys hydroxylation in the helical region of fibrillar collagen is still unclear (62,69-70). When the function of one isoform is impaired, compensatory mechanisms from the other two isoforms may occur, and may depend largely on the tissue-specific expression of each isoform (68).

In this study, the enzymatic function of LH3 as the main GGT but not GT was also confirmed by using the enzymatically active recombinant mouse LH3-V5/His fusion protein generated by 293HEK cells culture. By using HPLC, we could directly measure the levels of

GG-Hyl and G-Hyl in the reaction products of the GGT and GT activity assays. The specificity of the glucosylation activity of LH3-V5/His protein against denatured type I collagen was demonstrated by the dose dependent increase of GG-Hyl when increasing amounts of LH3-V5/His protein were used. Furthermore, no transfer of galactose was observed even with increasing amounts of LH3-V5 protein. This result is consistent with studies showing that the GT activity of LH3 protein is very low (79-80) or even undetectable (70,81) in contrast to the GGT activity. As for the expressions of *Glt25d1* and *2*, only the former isoform was detected in MC cells. From our results and those recently reported by Schegg *et al* (81), it is likely that galactosylation of Hyl in bone type I collagen is catalyzed by Glt25d1, followed by subsequent glucosylation by LH3.

The ability of recombinant LH3 to modify extracellular proteins (collagen or other proteins), as shown by GGT activity assay, indicated that LH3, unlike LH1 and LH2, is localized and functional in the extracellular space as well (76). However, the specific substrates for the extracellular function of LH3, such as collagens or proteins with collagenous sequences, have never been identified but assumed to be collagen. Interestingly, this finding contradicts with the earlier reports that triple helical conformation inhibits further modifications (77-78). Thus, it was proposed that in physiological temperature, there is a possibility of micro-unfolding of the triple helix which allows further modifications (76). In our study, when native bovine skin collagen was used as a substrate for the GGT activity assay, an increase in the level of GG-Hyl was not observed, as seen with denatured collagen substrate. This finding could suggest that the specific substrate for the enzymatic function of extracellular LH3 may not be type I collagen or that the secreted LH3 may play a different role.

From the current study, it is apparent that a decrease in glucosylation of G-Hyl in type I collagen significantly affects the collagen fibril formation kinetics observed in the *in vitro* fibrillogenesis assay. The collagen molecules with less glucosylation could self-aggregate faster and form thicker fibrils due in part to the lesser steric hindrance for the molecular packing. This result is consistent with other reports showing that the degree of collagen modification is inversely correlated to the rate and diameter size of collagen fibrils formed *in vitro* (8-9,185).

It has been shown that one of the major cross-linking sites in the helical region of type I collagen (residue 87 of α1 and α2 chains) are mainly di-glycosylated (113-116,118,123,168) (reviewed in Table 1.2). Thus, further studies are warranted to investigate the potential effects of the pattern (di- vs, mono-glycosylation) and extent of glycosylation on cross-link formation and maturation.

Taken together, this study shows that, by manipulating the LH3 gene expression in mouse osteoblasts, the major function of LH3 is to transfer glucose units to G-Hyl residues in the α chains of bone type I collagen. This result has also been corroborated with the HPLC-based glycosyltransferase activity assays, using recombinant LH3-V5/His fusion protein. The significant decrease in the transfer of glucose units, through LH3 suppression, results in the change in the *in vitro* collagen fibril formation kinetics. It can be concluded that the levels and types of Hyl glycosylation, and possibly, its loci in type I collagen are indeed critical for proper formation of collagen fibrils.

Chapter 4

Study II

LYSYL HYDROXYLASE 3-MEDIATED GLUCOSYLATION AT SPECIFIC MOLECULAR LOCI REGULATES THE FORMATION OF COLLAGEN CROSS-LINKS AND THE EXTRACELLULAR MATRIX MINERALIZATION

Marnisa Sricholpech[1], Irina Perdivara[2], Megumi Yokoyama[1], Hideaki Nagaoka[1], Kenneth B. Tomer[2] and Mitsuo Yamauchi[1]

[1] NC Oral Health Institute, School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, and [2] Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA.

Abstract

Recently, by employing the short hairpin RNA technology, we have successfully generated single cell-derived clones stably suppressing lysyl hydroxylase 3 (LH3) (Sh clones) in mouse osteoblast cell line, MC3T3-E1 (MC), and demonstrated the function of this enzyme as glucosyltransferase in type I collagen. To further elucidate the biological significance of this specific post-translational modification in type I collagen, we characterized the collagen produced by Sh clones and controls. Mass spectrometry analysis identified the specific glycosylation sites at α1-87 (but not α2-87), α1-174 and α2-173, and forms, galactosylhydroxylysine (G-Hyl) and glucosylgalactosylhydroxylysine (GG-Hyl) in type I collagen of MC. In these sites, GG-Hyl was significantly diminished with concomitant increase in the G-Hyl in Sh collagen compared to the control. From the cultured cell/matrix layers, the formation of covalent intermolecular cross-links in Sh clones was significantly lower when compared to the controls. When the cell/matrix layers from the clones were subjected to in vitro incubation up to 4 weeks, the decreases in immature collagen cross-links in Sh clones were slower than those of controls but the increases in mature cross-links were relatively comparable to controls. Collagen fibers from the Sh clones were thinner than those of the controls, as seen from picrosirius red staining. By in vitro mineralization assay, the onset of mineralized nodules formation in the Sh clones was significantly delayed when compared to MC and EV. These results indicate that the LH3-mediated glucosylation occurring at the specific molecular loci in the type I collagen molecule, plays critical roles in controlling collagen cross-linking, matrix organization and mineralization.

Introduction

O-linked glycosylation of Hyl residues has long been known to be unique post-translational modifications for collagens and proteins with collagenous sequences (5,150). They are modulated by collagen glycosyltransferases, the enzymes that transfer the carbohydrate units (D-galactose and D-glucose) *O*-glycosidically to the hydroxyl group of specific hydroxylysine (Hyl) residues. This group of enzymes comprise of hydroxylysyl galactosyltransferase (GT) (EC 2.4.1.50) and galactosylhydroxylysyl glucosyltransferase (GGT) (EC 2.4.1.66) which form galactosylhydroxylysine (G-Hyl) and glucosylgalactosylhydroxylysine (GG-Hyl) respectively (186). To date, the GT activity is shown to be modulated by the multifunctional enzyme lysyl hydroxylase 3 (LH3)(79-80), despite some conflicting data, and the most recently identified glycosyltransferase family 25 domain 1 and 2 (GLT25D1 and D2)(81). The GGT activity, on the other hand, is mainly catalyzed by LH3 (69-70,79,108,184) *(reviewed in section 1.6.2)*.

In normal conditions, studies have shown that the level of glycosylation varies among different collagen types (5,150), different functional regions in the same tissue (124,151), different tissues, and developmental stages (103,127,152-153,161). Previous studies have reported the identification of specific glycosylation sites in type I collagen from various species (113-117) *(reviewed in Table 1.2)*. To date, the biological functions of glycosylated Hyl residues have not been clearly defined, however studies have demonstrated their associations in collagen fibrillogenesis (8-10,187), susceptibility of collagen to enzymatic digestion (9), collagen cross-links formation (38,118,123,168), and collagen-matrix proteins interactions (148-149) *(reviewed in section 1.6.3-1.6.8)*. Moreover, it has been shown to have a significant contribution towards the quality of extracellular matrices synthesis, organization

and maturation, as seen from several connective tissue disorders in which different levels of collagen glycosylations were observed (100-101,154-159,163) *(reviewed in section 1.7)*.

In mineralized tissue (e.g. bone or dentin), type I collagen serves as a scaffold for mineral deposition by which the extent, types of post-translational modifications, and the molecular packing of collagen molecules were shown to have significant effects. Analyses of rat bone have demonstrated the differences between cortical and trabecular bone in terms of the types and degree of post-translational modifications and mineralization (124,151,188). Previous report from our laboratory, by Pornprasertsuk et al, has shown that by overexpressing LH2b (the telopeptidyl LH) in MC cells, there is a significant increase in the formation of Hyl^{ald}-derived cross-links, formation of collagen fibrils with smaller diameter and a delayed mineralization (49). On the other hand, underexpression of decorin (DCN), the negative regulator of collagen fibrillogenesis, in MC cells resulted in the formation of collagen fibrils with irregular shapes and larger diameter along with an enhanced rate of mineral deposition (17,48) *(reviewed in section 1.5)*.

From Study I (Chapter 3), we have shown that the major function of LH3 protein for bone type I collagen is to transfer glucose units to G-Hyl residues and that alteration in the level of glucosylation significantly affected the kinetics of in vitro collagen fibril formation. To further elucidate the biological significance of the LH3-mediated glucosylation, type I collagen from MC and one of the previously established Sh clones will be subjected to mass spectrometry analysis for the identification of specific glycosylation sites and forms (GG-Hyl and G-Hyl). Moreover, the effects of lowered level of type I collagen glucosylation, from LH3 suppression, on cross-links formation and maturation, extracellular matrix organization, maturation and mineralization, would be investigated.

Experimental Procedures

Cell culture, generation of Sh clones and purification of type I collagen

The pSilencer2.1-U6/neo-Plod3 construct encoding the short hairpin sequence targeting *Plod3* was generated, and the MC3T3-E1cell-derived clones stably suppressing *Plod3* (Sh clones) were obtained as previously described (Chapter 3). MC cells and the previously characterized empty vector transfected clone (EV clone: transfected with the original pSilencer2.1-U6/neo encoding a hairpin siRNA sequence not found in any genome database) were used as controls. The suppressions of *Plod3* were verified by quantitative real time PCR and the levels of LH3 protein were assessed by Western blotting or immunoprecipitation followed by Western blotting with anti-LH3 antibody. Type I collagen purified from the Sh clones and controls were characterized for their collagen glycosylation profiles and fibril formation kinetics (Chapter 3). The three Sh clones (Sh1-1, Sh1-2, and Sh1-3) and their purified type I collagens, reported previously will be further characterized in this study.

Identification of the specific Hyl glycosylation sites and forms by mass spectrometry

Proteolysis and SDS-PAGE separation. Purified type I collagen (5 µg) was dissolved in sample loading buffer and denatured at 95˚C for 10 minutes. The α1 and α2 chains were separated by SDS-PAGE on 4-12% Bis-Tris pre-cast gels (Invitrogen, Carlsbad, CA) for 1 hour at 200 V, 100 mA and 10 W. Following staining with Coomasie Simply Blue, the protein bands were excised and subjected to automated in-gel digestion with trypsin over night at 37˚C, using a Progest robot digester (Genomic Solutions, Harbor, MI). Alternatively,

purified type I collagen (10 µg) was dissolved in 25 mM ammonium bicarbonate, pH 7.6, and digested in presence of α-chymotrypsin (enzyme:substrate ratio of 1: 20), over night at 25˚C. The enzymatic mixture was separated by SDS-PAGE as described, followed by excision of the bands and automated in-gel digestion with trypsin. The samples were stored at -80° C until further use.

UPLC-nanoESI-QTof MS. The collagen proteolytic mixtures were analyzed by LC/MS on a Waters-Micromass Q-Tof Premier mass spectrometer equipped with a nanoAcquity UPLC system (Waters, Milford, MA). Analyses were performed on a 3 µm, 100 µm × 100 mm, Atlantis dC18 column (Waters, Milford, MA), using a flow rate of 300 nL/min. A C18 trapping column (180 µm × 20 mm) with 5 µm particle size (Waters, Milford, MA) was positioned in-line of the analytical column and upstream of a micro-tee union used both as a vent for trapping and as a liquid junction. Trapping was performed for 3 min at 5 µL/min flow rate, using the initial solvent composition. A 4 µL aliquot of the digest sample was injected onto the column. Peptides were eluted by using a linear gradient from 98% solvent A (0.1% formic acid in water (v/v)) and 2% solvent B (0.1% formic acid in acetonitrile (v/v)) to 40% solvent B over 90 minutes. Mass spectrometer settings for the MS analysis were: capillary voltage of 3.2 kV, cone voltage of 20 V, collision energy of 5.0 V and source temperature of 80° C. Mass spectra were acquired over the mass range 200 – 2000 Da. For calibration, an orthogonal reference spray (LockSpray) of a solution of Glu1-Fibrinopeptide B (500 fmol/µL) in water/acetonitrile 80:20 (v/v) and 0.1% formic acid, having a reference mass of 785.8496 (2+) was used.

LC-nanoChip-ESI ion trap MS. The analyses were performed using an Agilent 6340 XCT Ultra Ion Trap (Santa Clara, CA) equipped with an HPLC Chip Cube MS interface, an Agilent 1100 1200 nanoHPLC and an electron transfer dissociation module. Ion trap-MS/MS analyses were performed as follows: 20 µL injections of the tryptic digests dissolved in 0.1% formic acid were loaded onto a 40 nL enrichment column followed by a 43 mm x 75 µm analytical column, packed with ZORBAX 300SB C18 particles (Agilent, Santa Clara, CA). Linear gradients of 3-50% (0.1% FA) were performed over 50 min at a flow rate of 500 nL/min. The parameter settings for positive ion ESI-MS were as follows: capillary voltage – 2000 V, end plate offset – 500 V, capillary exit – 180 V, nebulizer – 2 psi, dry gas – 4 L/min, dry gas temperature – 325 °C. For MS/MS (ETD or CID), automated data dependent acquisitions of the six most abundant ions were employed. For CID, the fragmentation amplitude was 0.80 V. For ETD analyses, the accumulation time of the fluoranthene gas was 40 msec, and the reaction time was typically 100 or 150 ms.

Expressions of lysyl oxidase family

MC cells, EV and Sh clones were plated at a density of $2x10^5$ cells/35mm dish and after 48 hours, total RNA was extracted with TRIzol reagent (Invitrogen), and two µg of total RNA was converted into cDNA by using the Omniscript Reverse Transcriptase kit (Qiagen, Valencia, CA, USA). Quantitative real-time PCR was performed using the sequence specific primer for lysyl oxidase (LOX) (*LOX*; ABI assay number: Mm00495386_m1), LOXL1 (*LOXL1*; Mm00241410_m1), LOXL2 (*LOXL2*; Mm00804740_m1), LOXL3 (*LOXL3*; Mm00442953_m1), LOXL4 (*LOXL4*; Mm00446385_m1). The reactions were prepared and analyzed in triplicates by the ABI Prism 7000 Sequence detection system (Applied

Biosystems, Foster city, CA, USA). The mRNA expression level of LH3 relative to glyceraldehyde 3-phosphate dehydrogenase (GAPDH; ABI assay number 4308313) was analyzed by the $2^{-\Delta\Delta CT}$ method (171).

Collagen cross-links analysis

MC, Sh (Sh1-1, 1-2 and 1- 3), and EV clones were cultured in α-MEM, 10% FBS, and 50 µg/ml ascorbic acid. After 2 weeks of culture, cells/matrices were scraped, thoroughly washed with PBS and cold de-ionized distilled water (DDW), and lyophilized. The samples were prepared for collagen cross-links analysis as previously described (173). Briefly, two mg of dried samples was suspended in 0.15 M *N*-trismethyl-2-aminoethanesulfonic acid and 0.05 M Tris-HCl buffer (pH 7.4) and reduced with standardized NaB^3H_4. The reduced samples were hydrolyzed with 6N HCl *in vacuo*, flushed with N_2, incubated at 110°C for 22 hrs, dried, dissolved in distilled water, and filtered. An aliquot of the hydrolysate was subjected to amino acid analysis to determine Hyp content, and the hydrolysates with known amounts of Hyp were analyzed for cross-links on a cation-exchange column (AA-911; Transgenomic., Omaha, NE, USA) linked to a fluorescence detector (FP1520; Jasco Spectroscopic, Tokyo, Japan) and a liquid scintillation analyzer (500TR series; Packard Instrument, Meriden, CT, USA) as described previously (2). The cross-link precursor aldehydes (i.e., Hyl^{ald} and Lys^{ald}), the major reducible cross-links (i.e., dehydrodihydroxylysinonorleucine/its ketoamine [deH-DHLNL], dehydrohydroxylysinonorleucine/its ketoamine [deH-HLNL]), and dehydrohistidino-hydroxymerodesmosine (deH-HHMD) were analyzed as their reduced forms, that is, dihydroxynorleucine (DHNL), hydroxynorleucine (HNL), DHLNL, HLNL, and HHMD,

respectively. The level of mature non-reducible cross-links, pyridinoline (Pyr) and deoxypyridinoline (d-Pyr) were measured with the fluorescence detector and quantified based on the available standard. All cross-links and precursor aldehydes were quantified as moles per mole of collagen (33).

The glycosylated immature bivalent cross-links (GG-DHLNL, G-DHLNL, GG-HLNL or G-HLNL) were analyzed by subjecting the reduced cell/matrix layer to base hydrolysis with 2N NaOH, in sealed polypropylene tubes flushed with N_2 gas, and incubated at 110°C for 22 hours. The hydrolysate was acidified to a ~pH 3 with 2N HCl, and filtered. Without the available glycosylated immature cross-links standards, the MC base hydrolysate was applied to the cation-exchange column mentioned above. The separation of the potential reducible, glycosylated and the known non-glycosylated immature cross-links were observed. Brief acid hydrolysis of the base hydrolysate with 6N HCl at 110°C for 20 minutes was performed to cleave the sugar attachments for further identification of the types and elution time of the different cross-links species (GG-,G-, or un-modified). Next, an aliquot of the reduced cell/matrix base hydrolysate of the Sh clones and controls were separated as mentioned above. The amounts of the glycosylated and non-glycosylated immature reducible cross-links were quantified and shown in percentages. The amount as moles / mole collagen cannot be calculated solely from base hydrolysate analysis, since Hyp is partially degraded in the base hydrolysis condition, thus the levels of total cross-links from acid hydrolysis is needed. As for the mature cross-links, Pyr and its glycosylated forms are unstable under the base hydrolysis condition, rendering its quantification not feasible.

In vitro collagen cross-links maturation assay

The lyophilized cell/matrix layers collected at 2 weeks from the Sh clones and controls, MC and EV, were each weighed at 2 mg/time point, for 4 time points (Wk0, Wk1, Wk2 and Wk4). The cell/matrices, except the ones assigned for Wk0, were incubated in air, in 1ml PBS supplemented with 0.7mM βAPN (ß-aminopropionitrile), two drops of toluene and placed in a 37°C incubator, avoiding light exposure. At the end of each time point, the samples were removed from the incubator, centrifuged, and lyophilized. When the samples from all the time points were collected, they were reduced with standardized NaB^3H_4, acid hydrolyzed and analyzed for the levels of immature reducible cross-links (DHLNL and HLNL) and mature non-reducible cross-links (Pyr and d-Pyr) as described above.

Picrosirius red staining

The organization and maturation of type I collagen produced by the controls and Sh clones were assessed by Picrosirius Red Staining (189-190). The cells and clones were plated at the density of $2X10^5$ cells onto coverslips placed in 35mm dishes and cultured in the medium supplemented with 50μg/ml ascorbic acid for 4 weeks after confluence. The cell/matrix layers were then washed twice with PBS, fixed with 10% formalin for 30 minutes, washed twice with PBS and incubated with 2ml of Picrosirius Red reagent (0.1% Sirius Red in saturated picric acid; Electron Microscopy Science, Hatfield, PA, USA) at room temperature for 30 minutes. The stained matrices were rinsed with DDW and followed by overnight wash at room temperature. The coverslips were lifted from the dishes and dipped sequentially in increasing concentration of ethanol (50-60-70-90-100%) and xylene. The coverslips with the attached matrices were embedded onto the glass slide and left to dry

at room temperature for overnight. The samples were observed under a polarized light microscope (Olympus America Inc., Pittsburgh, PA, USA) and photographed.

In vitro mineralization assay

MC cells, EV, and Sh clones were plated at the density of $2X10^5$cells/35 mm dishes and cultured in α-MEM containing 10% FBS, 100 units/ml penicillin, and 100μg/ml streptomycin. When confluence is reached, the cells were maintained in the mineralization medium containing 50μg/ml ascorbic acid and 2mM β-glycerophosphate, and cultured up to 4 weeks. At the end of each week, the cell/matrices layers from each sample were washed with PBS twice, fixed with 100% methanol and stained with 1% Alizarin Red S (Sigma-Aldrich, St Louis, MO, USA), which binds to calcium in the mineralized nodules deposited. The rate and the extent of mineral deposition among the clones with varying levels of LH3 enzyme were compared.

Statistical analyses

Statistical analyses were performed using Jmp®8.0 software (SAS Institute Inc., Cary, NC, USA). Statistical differences were determined by Kruskal-Wallis one-way analysis of variance and means comparison by Student's t test. The data were presented as means ± SD and a *p* value less than 0.05 was considered significant.

Results

Identification of specific glycosylation sites by mass spectrometry

For glycopeptide analysis, the α1(I) and α2(I) chains were separated by SDS-PAGE, digested with trypsin and analyzed without further enrichment by HPLC-chip nano ESI-ion trap MS. The MS and MS/MS spectra were acquired using data dependent acquisition of the six most abundant precursor ions. The glycopeptides [76-90] containing the glycosylated residue Hyl 87 were observed as triply protonated ions of *m/z* 558.7 and 612.7, corresponding to glycoforms G-Hyl and GG-Hyl, respectively. The extracted ion chromatograms (EIC) of the two ions in the α1 tryptic digest isolated from the MC and Sh1 collagen are shown in Figure 4.1A and 4.1C. The averaged mass spectra over the chromatographic elution time of glycopeptides [76-90] observed in the MC and Sh1 collagen sample are shown in Figure 4.1B and 4.1D. The major glycopeptide observed in the MC α1 tryptic digest as the ion of *m/z* 612.7 (3+), was assigned to peptide [76-90] containing GG-Hyl and oxidized methionine. The glycoform containing G-Hyl, at *m/z* 558.7 (3+), was observed with lower relative abundance (Figure 4.1B). In contrast, in the α1 chain isolated from the Sh1 collagen, the most abundant glycoform attached at residue 87 was assigned to G-Hyl, while GG-Hyl was found considerably decreased compared to MC (Figure 4.1C and 4.1D). Structural characterization of these glycopeptide ions by electron transfer dissociation (ETD) confirmed the assignment based on the presence of fragment ions arising from cleavage of the N-Cα bond retaining the glycan moiety (Figure 4.2 A-B). In the full scan mass spectrum, additional ions arising from the loss of the N-terminal residues Gly-Leu from the glycosylated peptide [76-90] were observed as peaks at *m/z* 556.0 (3+) and 502.1 (3+),

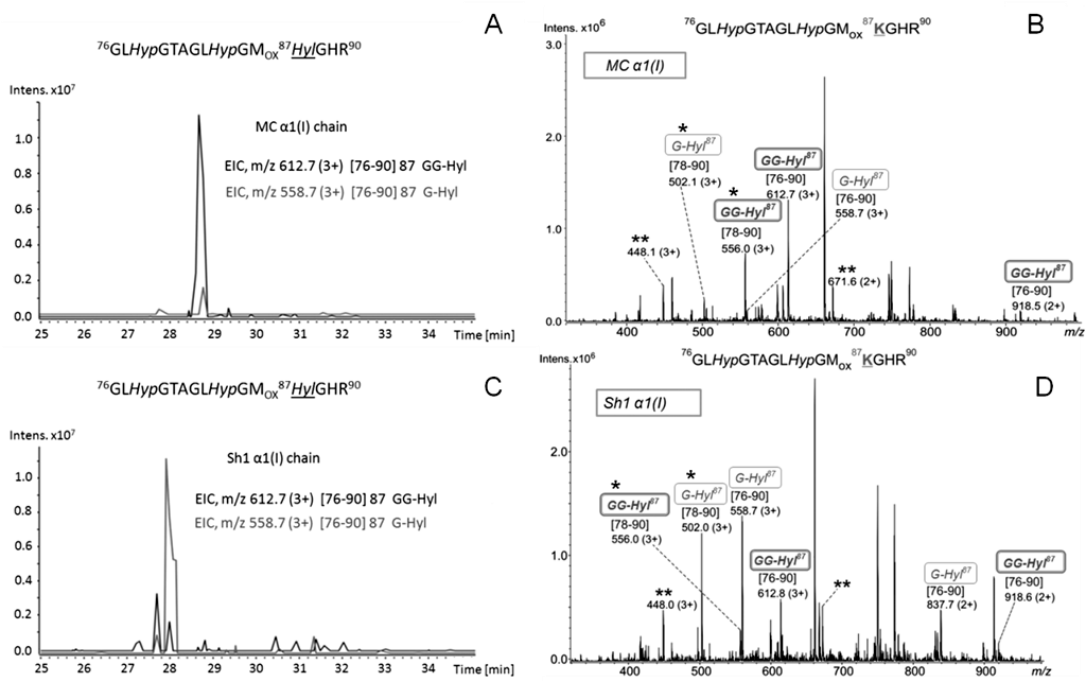

Figure 4.1) Mass spectrometry characterization of MC and Sh1 α1(I) glycopeptides [76-90]. Extracted ion chromatograms (EIC) of ions *m/z* 612.7 (3+) and 558.7 (3+) in the tryptic digest of collagen α1(I)-chain isolated from (A.) MC collagen, and (C.) Sh1 collagen, showing the relative abundance of glycopeptides [76-90] containing G-Hyl and GG-Hyl. Full scan mass spectra averaged over the chromatographic elution time of glycopeptide ions of *m/z* 612.7 and 558.7 in (B.) MC collagen, and (D.) Sh1 collagen. The ions indicated with asterisks represent in-source decomposition products arising from (*) peptide bond cleavage between residues Leu-Hyp at the N-terminus, and (**) glycosydic bond cleavage. The data was obtained by LC-chip nano ESI ion trap.

and were assigned to residues [78-90] modified with GG-Hyl and G-Hyl, respectively. Furthermore, the ion of *m/z* 448.1(3+), arising from the neutral loss of galactose from the peptide fragment [78-90], was observed in the MS spectrum as well. The three triply protonated ions of *m/z* 556.0, 502.0 and 448.1 are, most likely, formed by in-source decomposition of the glycosylated species [76-90], because they have identical chromatographic elution profiles with the intact glycopeptides observed as *m/z* 612.7 and 558.7. However, overall in MC collagen, the glycopeptides [76-90] containing GG-Hyl are more abundant than that containing G-Hyl while in Sh collagen, the majority of the

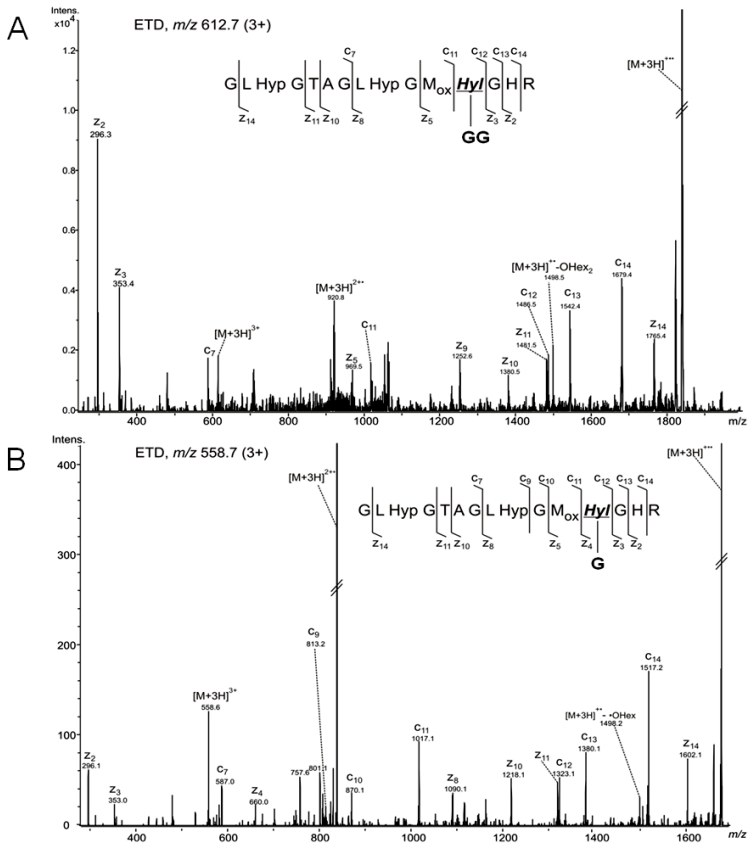

Figure 4.2) Structural characterization of α1(I) glycopeptides [76-90]. Electron transfer dissociation (ETD) spectra acquired on the ion trap of the ions of (A.) *m/z* 612.7 (3+), and (B.) *m/z* 558.7 (3+), corresponding to the α1(I) glycopeptides [76-90] containing GG-Hyl, and G-Hyl respectively.

glycopeptides [76-90] contain G-Hyl. The relative changes in the glycosylation pattern of site Hyl 87 observed in Sh1 collagen can be attributable to the suppression of LH3.

From MC collagen, Lys residue 87 on the α2 chain was exclusively observed modified by lysine hydroxylation in peptide [76-87] as the doubly protonated molecule of *m/z* 619.81. The identity of the peptide [76-87] was confirmed by MS/MS (Figure 4.3). No species containing unmodified (α2-Lys 87) or glycosylated α2-Hyl 87 was observed. Similar modification profile was seen at α2-87 of Sh1 collagen. It is noteworthy that this residue was

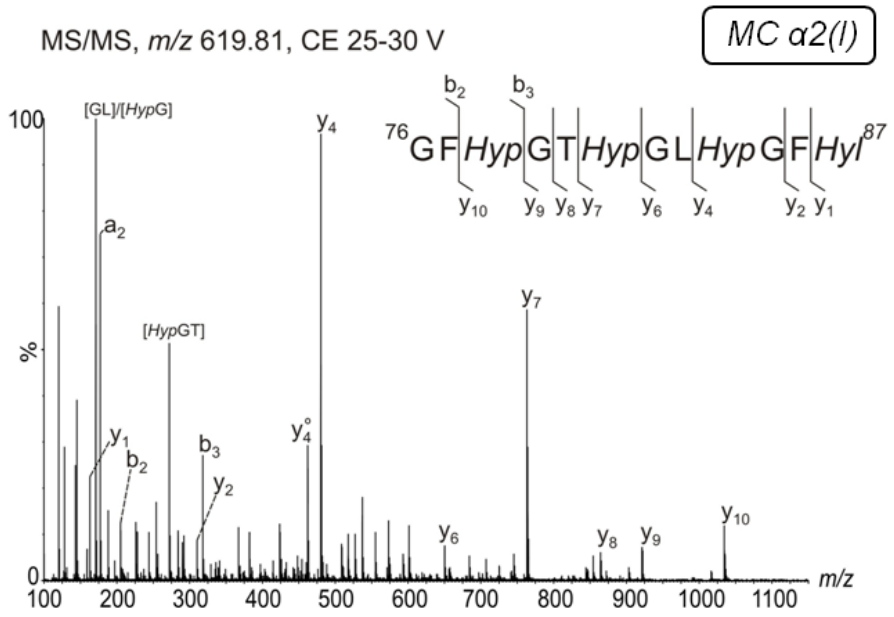

Figure 4.3) Structural characterization of α2(I) peptide [76-87]. Collision induced dissociation (CID) spectrum of the ion of *m/z* 619.81 (2+), assigned to the α2(I) peptide [76-87], containing hydroxyproline as the C-terminal residue. The spectrum was obtained on the QTof Premier with a collision energy ramp from 25 to 30 V.

found glycosylated in collagen isolated from bovine bone (data not shown). However, two amino acid substitutions proximal to Hyl 87 in the mouse α2 chain, Ile89Val and Arg90Lys, may be responsible for the observed differences in glycosylation status of α2-Hyl 87.

Early pioneer work in the area of collagen glycosylation documented the existence of additional glycosylation sites in type I collagen isolated from different sources (116-117) (reviewed in Table 1.2). In the LC-MS/MS analysis of MC-collagen digested with α-chymotrypsin and trypsin on the QTof Premier, glycosylated structures contained at Hyl 174 on α1, and Hyl 173 on α2 chains were observed. Residue 174 on MC α1 chain was observed in peptides [145-174] and [145-183], the latter containing one miscleavage at Hyl 174. The deconvoluted mass spectrum over the chromatographic time window in which the molecular

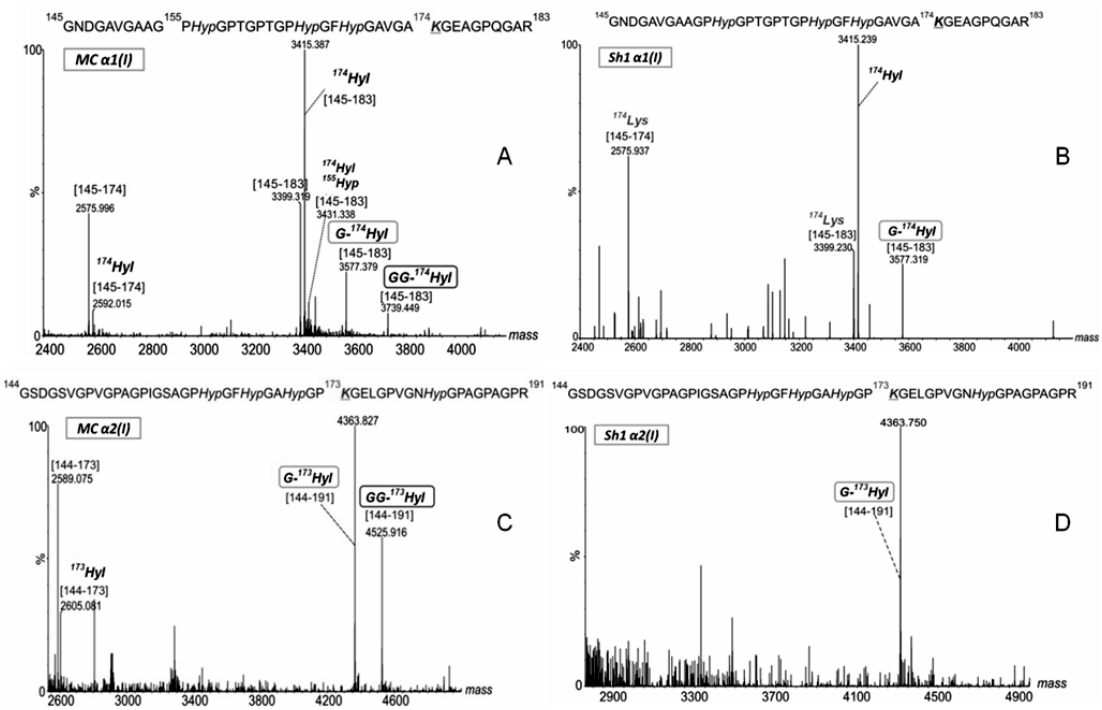

Figure 4.4) Mass spectrometry characterization of modifications at α1(I)-174 and α2(I)-173 in MC and Sh1 collagen. Deconvoluted mass spectra showing the site-specific microheterogeneity derived from lysine hydroxylation and hydroxylysine glycosylation of : *(i)* residue 174 on α1(I) chain in (A.) MC collagen, and (B.) Sh1 collagen, and *(ii)*, residue 173 on α2(I) in (C.) MC collagen, and (D.) Sh1 collagen. The amino acid sequence of the observed tryptic peptides are indicated at the top of each spectrum. The data was acquired on a Waters QTof Premier instrument.

species containing residue 174 eluted, is shown in Figure 4.4A. The extended molecular microheterogeneity of the species bearing residue 174 is attributable to lysine hydroxylation and partial glycosylation of hydroxylysine. The major modification observed at residue 174 is represented by lysine hydroxylation, while glycosylation of Hyl 174 is relatively low.

However, within the glycosylated structures, G-Hyl ($[M+H]^+$obs.: 3577.379 Da) was found with higher relative abundance than GG-Hyl ($[M+H]^+$ obs: 3739.449 Da), in contrast to the glycosylation pattern observed for MC α1-Hyl 87. The structural characterization of glycopeptides [145-183] was confirmed by collision induced dissociation on the QTof (data

not shown). Additional heterogeneity of peptides [145-183] was due to hydroxylation of Pro 155, representing, most likely, 3-hydroxyproline. In the Sh1 collagen, isolated from the LH3-suppresed clone, similar heterogeneity of lysine hydroxylation to the wild type was observed. However, the only glycosylated structure detected at Hyl 174 was G-Hyl, while GG-Hyl was present at the noise level (Figure 4.4B). This again indicates that the major function of LH3 is to transfer glucose residues to the G-Hyl residues in collagen.

The residue homologous to α1-Lys 174 is residue 173 on the α2 chain, observed in peptides [144-173] and [144-191]. The deconvoluted mass spectrum over the chromatographic time window in which the molecular species containing residue 173 eluted, is shown in Figure 4.4C, for the MC collagen. Similar to residue α1-174, Lys 173 was observed in peptide [141-173] both unmodified ($[M+H]^+$ obs: 2589.075 Da), as well as hydroxylated ($[M+H]^+$ obs: 2605.081 Da). In contrast to α1-Hyl 174, α2-Hyl 173 was found to have higher glycosylation occupancy, as the relative levels of glycosylated Hyl are higher compared to those of free Hyl. The glycosylated peptides [144-191] were observed as G-Hyl ($[M+H]^+$ obs:$^+$ 4363.827 Da) and GG-Hyl $[M+H]^+$ obs: 4525.916 Da). Their identities were confirmed by MS/MS (data not shown). Drastic changes in the glycosylation status of site 173 on α2 were found in the Sh1 collagen isolated from the LH3-suppresed clone (Figure 4.4D), because no glycopeptide ion corresponding to [144-191] bearing GG-Hyl was detected.

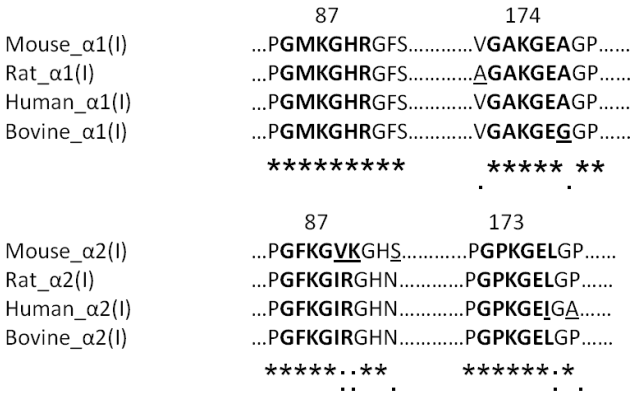

Figure 4.5) Protein sequence alignment of α1(I) and α2(I) chains from various species. Sequence alignment at the identified glycosylation sites, α1-87, α1-174, α2-87 and α2-173, among mouse, rat, human and bovine. Note the sequence differences at residues 89 and 90 of mouse α2 chain comparing to the other species.

By comparing the amino acid sequences of the α1 and α2 chains, exclusively at the regions of the identified glycosylation sites, among different species, it is apparent that in the α1 chain there is a high degree of sequence homology around the Lys residue (K) found to be glycosylated (residue 87 and 174) (Figure 4.5A). The sequence comparison of α2 chain, on the contrary, showed amino acid differences adjacent to the Lys residue 87 of the mouse sequence (GFKG**VK** vs GFKGIR) but not at residue 173 (Figure 4.5B). The differences shown may indicate the requirement of specific sequences for the galactosyltransferase reaction.

Collagen cross-links analysis

The collagen cross-links produced by MC cells and the Sh clones at 2 weeks culture composed mainly of immature reducible bivalent crosslinks (DHLNL and HLNL), and in a small amount mature trivalent cross-links (Pyr). Deoxypyridinoline (d-Pyr) was not detected in the cultures of MC cells or EV and Sh clones. The actual amounts of DHLNL, HLNL, Pyr, and the cross-links forming aldehydes (total aldehydes = DHLNL + HLNL + 2XPyr) are depicted in Table 4.1. From the LH3-suppressed Sh clones, it is evident that the less

Table 4.1 The levels of immature reducible cross-links (DHLNL and HLNL) and mature non-reducible cross-links (Pyr) from MC cells, EV and Sh clones. (moles/mole collagen)

Cell/Clones	*DHLNL*	*HLNL*	*Pyr*	*Total Ald*
MC	1.08 (0.26)	0.21 (0.03)	0.047 (0.008)	1.39 (0.24)
Sh1-1	0.69 (0.07) *	0.17 (0.02)	0.026 (0.005) *	1.04 (0.15)
Sh1-2	0.66 (0.03) *	0.15 (0.05)	0.023 (0.005) *	0.85 (0.01)
Sh1-3	0.66 (0.08) *	0.17 (0.02)	0.016 (0.005)*	0.88 (0.10)
EV	1.06 (0.26)	0.24 (0.01)	0.044 (0.018)	1.25 (0.46)

Values represent mean (SD) from three independent experiments. *, significantly different from controls MC and EV ($p < 0.05$). Total Ald = DHLNL + HLNL + 2XPyr, DHLNL, dihydroxylysinonorleucine; HLNL, hydroxylysinonorleucine; Pyr, pyridinoline; and Ald, aldehydes.

glucosylated type I collagen showed a decrease in the level of cross-links formation. The decrease of DHLNL, formed from telopeptide Hyl^{ald} and helical Hyl residue, was statistically significant ($p < 0.05$), while the level of HLNL was less affected. Moreover, the levels of Pyr from the Sh clones were also found to be significantly lower ($p < 0.05$). Altogether, the levels of total Ald (DHLNL + HLNL + 2XPyr) in the Sh clones showed a decreasing trend when compared to the controls, even though not statistically significant. However, it is apparent that the significant changes are seen with the immature cross-link, DHLNL and its maturation product, Pyr.

Expressions of LOX and its isoforms

The lowered levels of total aldehydes, hence decreased collagen cross-links formation observed in the Sh clones may be contributed by the effect of LH3 suppression on the lysyl oxidase (LOX) enzymes, catalyzing the Lys and Hyl aldehydes formation which are the precursors of intermolecular cross-links. However, real-time PCR analysis showed a slight upregulation of LOX and its isoforms (LOXL1, LOXL3 and LOXL4) expressions in the Sh

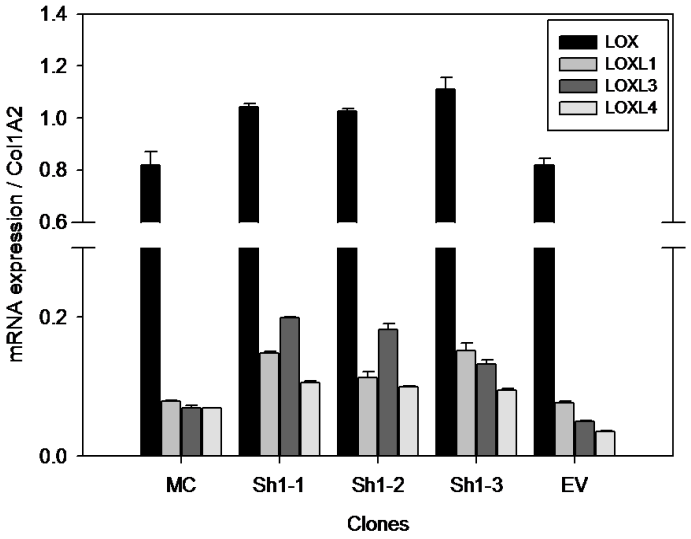

Figure 4.6) mRNA expression of LOX family (LOX, LOXL1, LOXL3, and LOXL4) relative to Col1A2, in the Sh clones (Sh1-1,-2,and -3) and controls (MC and EV). LOX, lysyl oxidase; LOXL, lysyl oxidase like. Error bar indicates mean ± SD of triplicate analyses.

clones compared to the controls (Figure 4.6), rendering this cause unlikely. The expression of LOXL2 was not detected in both the controls and Sh clones. Since the amount of LOX enzymes are not lowered, thus it could be suggested that their functions may be impaired.

Glycosylated cross-links analysis

Glycosylation of the immature bivalent cross-links were analyzed from the base hydrolysate of the reduced cell/matrix layer of MC cells and the EV and Sh clones. The identities of the potential glycosylated cross-links peaks were indirectly confirmed by brief acid hydrolysis of the base hydrolysate to observe the conversion from the GG- or G-modified form into the unmodified cross-links i.e. DHLNL or HLNL. In Figure 4.7A, the chromatogram of cross-links analysis from MC base hydrolysate showed the separation of the unmodified DHLNL and HLNL (based on the available standards) and a major peak eluted at 26 minutes, possibly the glycosylated cross-links (123,191). The result from the acid hydrolysis of the MC base hydrolysate is shown in Figure 4.7B. There was a significant decrease of the peak eluted at 26 minutes with simultaneous increases in the peaks at 33 minutes and 47 minutes, which is the known elution time for unmodified DHLNL,

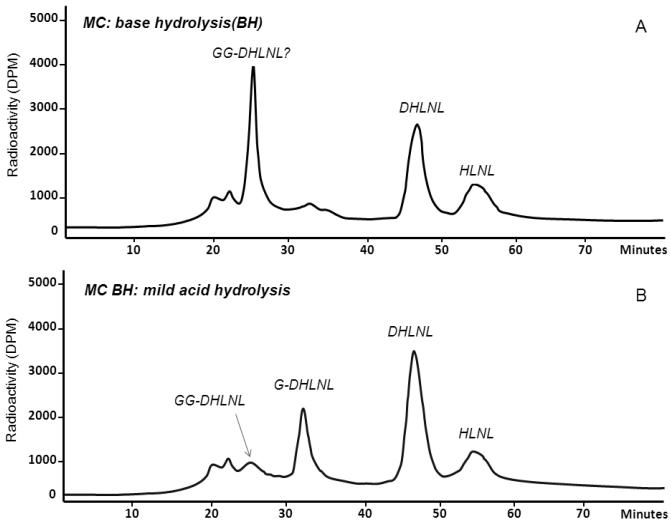

Figure 4.7) Ion-exchange chromatogram of potential glycosylated immature reducible cross-links. A. Chromatogram of base hydrolysate of reduced MC cell/matrix layer showed a high abundance of radioactive compound eluted around 26 min. B. Chromatogram of the brief acid hydrolysis of the above base hydrolysate showing a significant decrease in the abundance of the peak eluted at 26 min with simultaneous increase of the peaks at 33 and free DHLNL at 47 min.

GG-: glucosylgalactosyl-, G-: galactosyl-, DHLNL: dihydroxylysinonorleucine, HLNL:

suggesting the occurrence of the conversions. From this result, GG-DHLNL and G-DHLNL could possibly be assigned to the peaks at 26 and 33 minutes respectively. As for HLNL, no significant alteration was observed in the base followed by brief acid hydrolysis condition, suggesting that they may be originated from Lys^{ald} X non-glycosylated Hyl or Hyl^{ald} X Lys. Unfortunately, glycosylated Pyr cannot be analyzed by this method since it was reported that base hydrolysis could destroy at least 90% of the 3-hydroxypyridinium cross-links (Pyr and d-Pyr)(192).

Table 4.2 shows the quantification of GG-DHLNL, G-DHLNL and free DHLNL from the MC, EV and Sh clones from two independent experiments. The levels are shown in moles/mole collagen, calculated from the percentages of each DHLNL species (GG-, G-, and free DHLNL) obtained from the analyses of base hydrolysates and the total DHLNL from acid hydrolysates. From Table 4.1, we have shown that the level of total DHLNL from the Sh clones are significantly lower comparing to the controls ($p < 0.05$). By quantifying the levels of different DHLNL species, it is evident from the Sh clones that, there is a decrease in the

Table 4.2 Levels of glycosylated and free DHLNL. (moles/mole collagen)

Cell/Clones	*GG-DHLNL*	*G-DHLNL*	*free DHLNL*
MC	0.503 (0.089)	0.046 (0.034)	0.635 (0.135)
Sh1-1	0.216 (0.030)	0.112 (0.061)	0.368 (0.013)
Sh1-2	0.178 (0.005)	0.131 (0.012)	0.357 (0.015)
Sh1-3	0.210 (0.035)	0.147 (0.068)	0.339 (0.018)
EV	0.406 (0.001)	0.157 (0.203)	0.567 (0.114)

Values represent mean (SD) from two independent experiments. GG-: glucosylgalactosyl-, G-: galactosyl-, DHLNL: dihydroxylysinonorleucine.

level of GG-DHLNL, comparing to MC and EV, with concomitant increase in the level of G-DHLNL when comparing to MC. The levels of free DHLNL in the Sh clones are lower when compared to MC and EV. The difference between the two batches of culture was observed exclusively in EV as shown by the high standard deviation. For that reason, extra sets of cultures are being analyzed for confirmation of the glycosylated DHLNL profiles.

From the MS analysis shown above, only one cross-linking site was shown to be fully glycosylated, α1-Hyl 87, therefore, the glycosylated cross-linking profile shown in Table 4.2 would most likely be derived from this particular site. The MS analysis of Sh1 collagen demonstrated a significant change in the site occupancy from GG-Hyl to G-Hyl residue (GG-Hyl < G-Hyl) (Figure 4.1). Surprisingly, from the cross-links analysis of the Sh clones, higher abundance of GG-DHLNL compared to G-DHLNL (GG-DHLNL > G-DHLNL) was observed (Table 4.2), suggesting that at α1-Hyl 87, the GG-Hyl residue may be the preferential form for the formation of immature bivalent cross-links, rather than G-Hyl. Free DHLNL, the major form found, could be derived from α2-Hyl 87 (mainly Hyl) or potentially the other cross-linking site residue 930, at which Hyl glycosylation has never been reported.

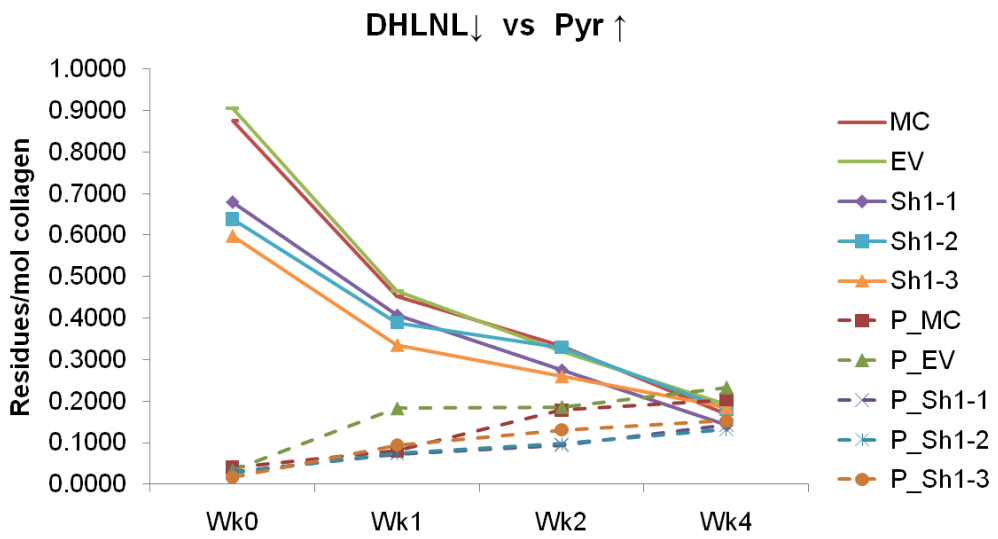

Figure 4.8) In vitro cross-links maturation assay. 3µg of lyophilized cell matrix layers from the culture of MC, EV and the Sh clones were incubated (see methods). At week 1, 2 and 4 of incubation, the samples were reduced with NaB^3H_4 and analyzed for the levels of immature reducible (DHLNL and HLNL) and mature non-reducible (Pyr) cross-links. DHLNL levels changes are shown with solid lines while those of Pyr (P_) are shown with dashed lines.

In vitro cross-links maturation assay

It is well established that Pyr is the maturational product from the condensation reaction between two DHLNL residues (34,36), resulting in a trivalent cross-link formed between Hyl^{ald} X Hyl^{ald} X Hyl. To directly measure the rate of cross-links maturation without the constant formation of new immature cross-links, we have utilized the in vitro cross-links maturation assay previously used in several reports (26,29,193-194) with some modifications. Figure 4.8 shows the changes in the levels of DHLNL and Pyr over time (Wk 0-1-2-4). After the first week of incubation of MC and EV cell/matrices, there is a fast decrease in the level of DHLNL which gradually decline thereafter. At the same time, interestingly, DHLNL in all the Sh clones were decreased at a slower rate, when compared to the controls. As for the formation of Pyr, it is apparent that in both the controls and Sh clones

Table 4.3 The level changes of Pyr and DHLNL during 4 weeks incubation and potential mechanism (moles/mole collagen)

Samples	Δ Pyr *Wk4-Wk0*	Pyr forming DHLNL *ΔPyr X2*	GG-DHLNL *Wk0*	Sum *(ΔPyr X2) + GG-DHLNL*	ΔDHLNL *Wk0-Wk4*
MC	0.18 (0.02)	0.35 (0.04)	0.503 (0.089)	0.858 (0.13)	0.757 (0.07)
Sh1-1	0.15 (0.06)	0.30 (0.11)	0.216 (0.030)	0.520 (0.14)	0.484 (0.07)
Sh1-2	0.19 (0.12)	0.38 (0.24)	0.178 (0.005)	0.556 (0.24)	0.452 (0.01)
Sh1-3	0.20 (0.10)	0.41 (0.19)	0.210 (0.035)	0.620 (0.23)	0.447 (0.05)
EV	0.19 (0.01)	0.38 (0.02)	0.406 (0.001)	0.789 (0.02)	0.705 (0.01)

Value represent mean (SD) from two independent incubations. ΔPyr: total pyridinoline formation, ΔDHLNL: total dihydroxylysinonorleucine decrease, GG-: glucosylgalactosyl-.

the increase of Pyr is clearly unproportional to the decrease of DHLNL. Two independent incubations were performed and the results were of the same trend. This unidentified loss of DHLNL (not forming Pyr) has been observed in previous studies (27-28) and to date the mechanism is still unknown.

Table 4.3 demonstrates the total level decrease of DHLNL (ΔDHLNL) and increase of Pyr (ΔPyr) during the 4 weeks incubation. The levels of Pyr-forming DHLNL (ΔPyrX2) were calculated and found that they were lower than the levels decline of DHLNL (ΔDHLNL). In bone, it was previously shown that for immature bivalent cross-links, GG-DHLNL is the predominant form over G-DHLNL, while for the mature trivalent cross-links, interestingly, G-Pyr is the major form and GG-Pyr is undetectable (reviewed in Table 1.1). So far the fate of GG-DHLNL in the maturation process is still unknown. From these evidences, it is possible that the unidentified loss of total DHLNL observed in the cross-links maturation assay shown in this (Figure 4.8) and previous studies (28) may be attributed to the unknown fate of GG-DHLNL. Moreover, the different levels of GG-DHLNL between the controls and Sh clones may be a possible cause of the different rate decrease of DHLNL

shown in Figure 4.8. In Table 4.3, when the levels of GG-DHLNL at Wk0 (non-incubated) were added to the levels of Pyr-forming DHLNL during 4 wks incubation (ΔPyrX2) from each sample (Sum), the outcome was relatively close to the total level decrease of DHLNL (ΔDHLNL) in both the Sh clones and controls. Altogether, the results shown here may suggest the possibility that mainly G-DHLNL and free DHLNL are the preferred forms to mature into G-Pyr and free Pyr respectively, which is consistent with previous reports showing mainly the presence of G-Pyr and free Pyr in bone tissue (38,121-122). The mechanism behind the disappearance of GG-DHLNL is yet to be determined.

Collagen matrix organization and maturation

The effect of lowered level of type I collagen glycosylation, from LH3 suppression, on collagen matrix organization and maturation was assessed by Picrosirius red staining. The Sh clones and the controls were cultured with supplementation of 50µg/ml ascorbic acid for 4 weeks followed by the staining of the matrices with 1% picrosirius red and observed under a polarized light microscope. As shown in Figure 4.9, the collagen matrix laid down by MC cells and EV clone is organized, mature and thick, as observed by the intense red staining of the collagen fibers. On the other hand, the collagen fibers of the Sh clones are thinner and less matured, as seen from the yellowish-orange staining. Two independent experiments were performed and analyzed at both 2 and 4 weeks time points, and the results were of similar trend (data not shown).

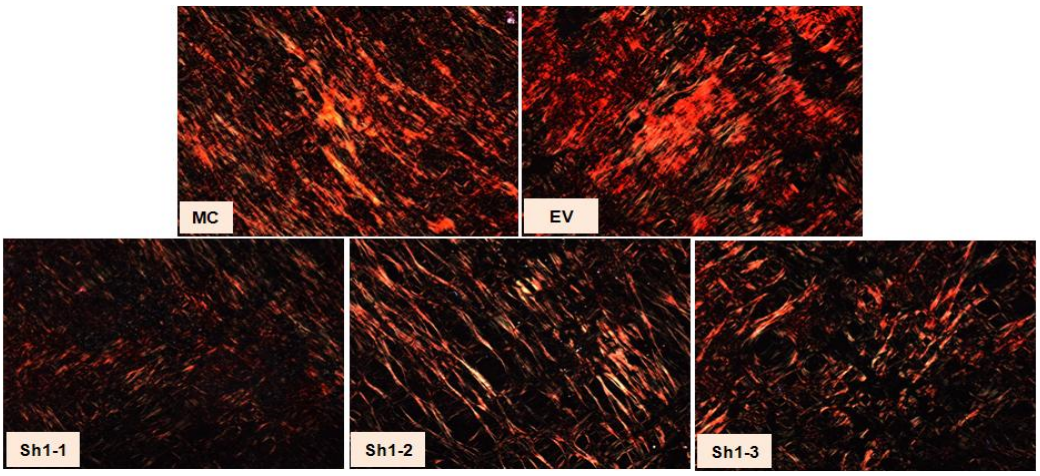

Figure 4.9) Picrosirius red staining. MC cells, EV and Sh clones were cultured for 4 weeks and the matrices were stained with picrosirius red observed under polarized light microscope. One representative image of each cell type is shown at 10X magnification. MC, MC3T3-E1 cells; EV clone, empty vector transfected clone; Sh clones, single cell-derived clones stably transfected with short hairpin containing plasmid targeting *Plod3.*

In vitro matrix mineralization assay

The results of *in vitro* mineralization assay, by Alizarin Red Staining, are shown in Figure 4.10. In the controls, MC and EV, the formation of mineralized nodules, is already observed at 2-3 weeks of culture and the number and size of nodules are gradually increases overtime. On the contrary, matrix mineralization in the Sh clones is markedly delayed. At 4 weeks, the cultures of Sh1-1 and Sh1-2 clones show almost no mineralized nodules formation, while a few is observed in Sh1-3 clone. Three independent experiments were performed with similar results observed.

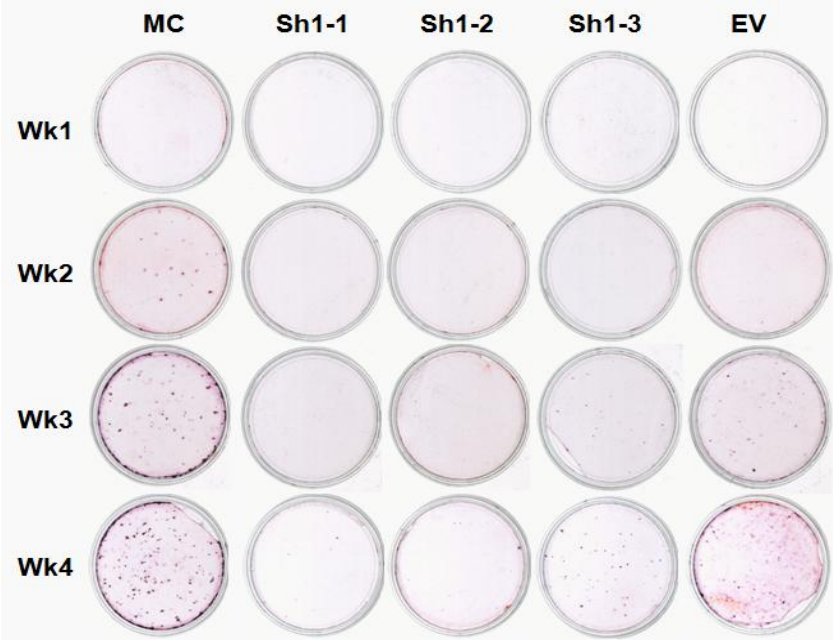

Figure 4.10) In vitro mineralization assay. MC cells, EV and Sh clones were cultured in mineralization medium for 1, 2, 3 and 4 weeks. At the end of each week, the cell/matrices were stained with Alizarin Red S and the mineralized nodules were observed. MC, MC3T3-E1 cells; EV clone, empty vector transfected clone; Sh clones, single cell-derived clones stably transfected with short hairpin containing plasmid targeting *Plod3*; Wk, Week.

Discussion

Type I collagen is a member of the fibrillar collagens with uniquely low levels of Hyl glycosylation (127,150). Sequencing of collagen and identifying its specific glycosylation sites have been studied in collagen from various tissues, species and reported in several publications over the 45-year span (reviewed in Table 1.2). The specific glycosylation sites in type I collagen identified, in most species, were those at residues 87 and 174 in α1 chain and residues 87 and 174 in α2 chain (113-117). In Chapter 3, we have shown that suppression of LH3 in mouse osteoblasts resulted in a decrease in the total levels of GG-Hyl with a simultaneous increase in the levels of G-Hyl in type I collagen, as determined by amino acid analysis. In the present study, we have employed high performance liquid chromatography – tandem mass spectrometry analyses of the glycopeptides formed by proteolytic degradation to characterize site specific glycosylation changes in type I collagen. One challenging aspect of collagen glycosylation analysis by mass spectrometry was derived from the fact that several glycosylated Hyl residues were contained in peptides larger than 3500 Da, which rendered these difficult to observe on the ion trap. However, from the MS analyses of the tryptic peptides, three Hyl residues were found glycosylated in MC collagen, i.e. Hyl 87 and Hyl 174 on α1, and Hyl 173 on α2 chain. In peptide [76-90] the glycosylated Hyl 87 was not cleaved by trypsin, suggesting that the substrate specificity is drastically altered by the presence of the carbohydrate units, which is in agreement with previous studies (195). Moreover, the absence of the cleavage at α1-Hyl 87 may also indicate that this site is fully glycosylated since enzymatic cleavage C-terminal of hydroxylysine was still observed in non-glycosylated peptides. While the major glycoform of the fully glycosylated α1-Hyl 87

in MC was determined as GG-Hyl, residues α1-Hyl 174 and α2-Hyl 173 were observed partially occupied, with the major glycoform determined as G-Hyl. In the LH3-suppressed clone, GG-Hyl was solely detected, with lower abundance, on α1-Hyl 87, while for α1-Lys 174 and α2-Lys 173 this glycoform was only present at the noise level. These sites were consistent with those identified previously (113-117) despite the numbering difference among species at residue 174 or 173 of α2 chain. Unlike earlier reports (reviewed in table 1.2), residue α2-87 in MC collagen was mainly detected as non-glycosylated Hyl. This result and the differences in the amino acids sequence in this particular region between mouse and the other species, may implicate sequence specific requirements for the galactosyltransferase enzyme function. Our results demonstrated, possibly for the first time, the specific molecular loci and forms of Hyl glycosylation in mouse bone collagen. These site-specific glycosylation changes, from LH3 suppression, may help explain the potential role of each glycosylated Hyl in the collagen physiology. In Sh1 collagen, the detection of GG-Hyl at α1-Hyl 87, but not at any other sites, indicates that G-Hyl 87 represents the primary substrate for the glucosylation activity of LH3.

Collagen cross-linking has been known to be essential for the mechanical stability of the extracellular matrix of the connective tissues. In mineralized tissues, there is a uniquely lower abundance of mature trivalent cross-links, Pyr and d-Pyr, and higher levels of immature bivalent cross-links, DHLNL and HLNL. Besides, Lys^{ald} derived cross-links are rarely observed, which indicates a high degree of telopeptidyl Hyl hydroxylation (196). In type I collagen, the two major cross-linking sites (residues 87 and 930) in the helical region of the α chains have been extensively studied and characterized (197-198). For residue 930, to date and to our best knowledge, there has been no direct evidence identifying the Hyl

glycosylation at this position. On the contrary, previous reports have shown that the residue 87-derived DHLNL (both from α1 or α2 chain) in bone were found to be mainly di-glycosylated (118-120,168). Also from these sites, a few HLNL (Lys^{ald} X Hyl) was detected and found to be glycosylated especially in the α2-derived HLNL (120) (reviewed in Table 1.1). Previously, Robins et al have shown that, in bone, 70% of HLNL is derived from Hyl^{ald} X Lys (199), hence no glycosylation. In this study, we have shown from mouse bone collagen that, the only cross-linking site to be glycosylated is α1-87 but not α2-87. Altogether, it could be suggested that the presence of glycosylated HLNL in mouse bone collagen would be very minimal, if any. As for the mature trivalent cross-links, in the MC cell culture system, only Pyr is present not d-Pyr, indicating that the helical Lys residues involved in the formation of pyridinium cross-links are fully hydroxylated.

The results from our study demonstrated that decreased G-Hyl glucosylation, mediated by LH3 suppression, result in a significant decrease in the formation of total aldehydes from the Sh clones, when compared to the controls, MC and EV. Interestingly, the expressions of LOX family (LOX, LOXL1, LOXL3, and LOXL4), the enzymes catalyzing the formation of cross-link precursors (Lys^{ald} or Hyl^{ald}) in the Sh clones were slightly upregulated compared to controls. Without the effect on the quantity of the LOX enzymes, the decreased formation of cross-links may result from their impaired functions. It was reported that LOX is more active towards quarter-staggered, native collagen fibrils and was suggested that the intermolecular interaction between collagen molecules are important for the enzyme activity (20). Studies have shown that the binding sites for LOX in type I collagen are in the triple helical region (21), potentially in the area with highly conserved sequences (Hyl-Gly-His-Arg), the cross-linking site residue 87, where it can catalyze the

formation of Lys or Hyl aldehydes in the telopeptides of the adjacent collagen molecule (reviewed in (22)). Collectively, there is a high possibility that the decrease in G-Hyl glucosylation at the major cross-linking site (α1-87) of the Sh collagen could affect the binding and/or function of the LOX enzymes, since it may alter the surface physico-chemical characteristics of the collagen molecules.

In this study, we have demonstrated the levels of glycosylated DHLNL (GG-DHLNL and G-DHLNL), potentially derived exclusively from α1-87, from the cultures of MC, EV and the Sh clones. Interestingly, in the Sh clones, the levels of GG-DHLNL are higher than G-DHLNL, while MS analysis at α1-87 showed a reverse in the relative abundance of GG-Hyl to G-Hyl (GG-Hyl < G-Hyl). Previous reports have also shown the higher levels of GG-DHLNL over G-DHLNL in bovine (119) and human (120) bone (reviewed in Table 1.1). Altogether, these data may indicate that at this particular site, GG-Hyl could be the preferred form for bivalent cross-links formation.

In regards to the maturation of immature bivalent into mature trivalent cross-links (DHLNL into Pyr, the major forms in bone), the most adopted mechanism is the spontaneous condensation reaction between two DHLNL resulting in the links between two Hyl^{ald} and one Hyl residue (34). By incubating bone collagen in vitro, faster decrease was observed in the level of DHLNL compared to HLNL, especially in the first 2 wks of incubation. The authors have suggested that the difference in the glycosylation content between them may have contributed to the different rate of disappearance during the maturation process (27). Moreover, Eyre et al had analyzed the level of immature and mature cross-links in human bone tissue of different ages, and showed a fast decrease in the level immature bivalent cross-links while the level of Pyr formation was unproportionally low. The data shown represented

the cross-links maturation in vivo, and prompted speculations of novel maturational products (28). In this study, in order to study the role of collagen glycosylation in cross-link maturation, we have employed the in vitro incubation assay and showed that, during the course of incubation there was a faster decrease in DHLNL of the controls, MC and EV, when compared to the Sh clones. We have speculated that the unproportional decline of total DHLNL may be from the loss of GG-DHLNL to an unknown mechanism. This speculation is based on the following observations: - 1. The decrease rate of DHLNL appeared to correlate with the level of GG-DHLNL (faster decrease with higher levels of GG-DHLNL), 2. In bone, GG-DHLNL is the predominant glycosylated immature cross-links, while G-Pyr is the major mature form present (38,121-122)(reviewed in Table 1.1), 3. From Table 4.3 of in vitro incubation study, when the amount of GG-DHLNL is added to the level of Pyr-forming DHLNL, the outcome is relatively close to the total level decrease of DHLNL. Further studies are thus warranted to uncover the mechanism and significance of this unidentified loss of DHLNL.

The organization and maturation of collagenous matrix of the Sh clones, exhibiting lowered level of GG-Hyl, has been characterized by picrosirius red staining. In comparison to the controls, MC and EV, the Sh clones laid down less matured and thinner collagen fibers. This result is not in accordance with our previous study (Chapter 3) and others (8-10,187) showing that collagen with lower level of glycosylation form fibrils with larger diameter, *in vitro.* This discrepancy may result from the fact that the *in vitro* fibril formation assay performed only account for the difference in the intrinsic factor in the collagen molecules between the Sh clones and controls, i.e. the levels of Hyl glucosylation. In the cell culture system, on the contrary, other potential extrinsic factors for collagen fibrillogenesis,

such as other minor collagen types (reviewed in (18)) or proteoglycans (12-13,200) may have contributed to the formation of thinner collagen fibers in the Sh clones. It also has been shown previously that type I collagen with lower levels of glycosylation was less deposited into the matrix (101,125), and was more susceptible to mammalian collagenase digestion (9). Consistent with former reports, our data provide indications that collagen molecules with optimal levels of modifications would be deposited into the matrix and stabilized by the formation of cross-links while collagen with lowered level of glucosylation (↓GG-Hyl and ↑G-Hyl), were less cross-linked and possibly susceptible to collagenase digestion. Consequently, the impaired matrices formed by the less glucosylated type I collagen, from the Sh clones, were apparently not a suitable template for the mineralization process, as seen from the significant delay in mineralized nodules formation. This data is consistent with our previous publications showing that small collagen fibrils are not good templates for mineral deposition (48-49).

In conclusion, by suppressing LH3 in mouse osteoblasts, we have developed a valuable tool to study the biological significance of the LH3-mediated glucosylation of G-Hyl in type I collagen. The glycosylation sites and forms were identified by MS and their alterations from LH3 suppression were correlated with the phenotypic changes in cross-links formation, extracellular matrix organization and mineralization. The results shown in our study and previous evidences may have indicated the biological roles of the glucose units in GG-Hyl residues, at the major cross-linking residue α1-87, in the mechanisms of intermolecular cross-links formation and maturation. Firstly, the glucose units appeared to have a positive regulatory role in the binding and/or function of LOX enzymes, since suppression of glucosylation resulted in a significant decrease in the level of total aldehydes

(total cross-links) formation. Secondly, because higher levels of GG-DHLNL were formed over G-DHLNL, even from the Sh clones with lower levels of GG-Hyl residues, therefore GG-Hyl may be the more preferred form in the formation of DHLNL. Thirdly, for the spontaneous conversion of DHLNL into Pyr, since GG-Pyr was barely found in bone collagen unlike other tissues (38,121-122), we have proposed that GG-DHLNL was likely involved in the rapid decrease of DHLNL observed in the in vitro incubation study. However, its fate and the selective condition or mechanism pertaining to specific tissues remained to be further investigated. Even so, the results in this study have shown, for the first time, the biological significance of the LH3-mediated glucosylation on the quality of type I collagen matrix formation and mineralization.

Chapter 5

Study III

GLYCOSYLTRANSFERASE 25 DOMAIN 1 GALACTOSYLATES HYDROXYLYSINE RESIDUES IN BONE TYPE I COLLAGEN

Marnisa Sricholpech, Megumi Yokoyama, Sun Min Lee, Hideaki Nagaoka and Mitsuo Yamauchi

NC Oral Health Institute, School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA,

Abstract

Glycosylation of collagens occurs through a series of enzymatic reactions including the hydroxylation of helical lysine (Lys) residues, galactosylation of hydroxylysine (Hyl) residues and finally, glucosylation of galactosylhydroxylysine (G-Hyl) residues forming glucosylgalactosylhydroxylysine (GG-Hyl) residues. We have shown in Chapter 3 that, in type I collagen, the multifunctional enzyme lysyl hydroxylase 3 (LH3) mainly catalyzes the glucosylation reaction not the Lys hydroxylation or subsequent galactosylation. Glycosyltransferase 25 domain 1 and 2 (Glt25d1 and 2) are a recently identified novel family of enzymes that may function as collagen galactosyltransferase. In our previous study (Chaptor 3), we reported that Glt25d1 was the only isoform expressed in mouse osteoblast cell line MC3T3-E1 (MC). In this study, in order to elucidate the function of Glt25d1 in type I collagen, we generated MC cell-derived clones stably suppressing Glt25d1 by use of short hairpin RNA technology and analyzed the collagen synthesized. The results showed that the levels of both G-Hyl and GG-Hyl in type I collagen from the Sh clones were significantly decreased in comparison to those of controls. The data indicates that Glt25d1 is the major collagen galactosyltransferase for bone type I collagen.

Introduction

Collagens and proteins containing the collagenous triplet repeats (Gly-X-Y), were shown to possess unique post-translational modifications including the hydroxylation of proline (Pro), and lysine (Lys) and the glycosylation of hydroxylysine (Hyl) (5). The biological significance of glycosylated Hyl in fibrillar collagens is still unclear however it was shown to be associated with the control of collagen fibrillogenesis (8-11,112), cross-linking (38,118,121-124,168), remodeling (126-133) , and collagen-cell interaction (148-149) (*reviewed in section 1.6.3-1.6.8*). Besides, it is shown to be essential for the biosynthesis and secretion of highly glycosylated non-fibrillar collagens (69-70,99). For collagen-like proteins, the glycosylated Hyl residues in adiponectin and mannose-binding lectin (MBL) appeared to be involved in the oligomerization and secretion of the proteins (reviewed in (81)).

Hydroxylysine glycosylation is modulated by the group of enzymes known as collagen glycosyltransferases which transfer the carbohydrate units (D-galactose, or D-glucose) O-glycosidically to hydroxyl group of specific Hyl residues. The reactions involve two essential steps, firstly, the galactosylation of Hyl by hydroxylysyl galactosyltransferase (GT) (EC 2.4.1.50) forming galactosylhydroxylysine (G-Hyl), and secondly, the glucosylation of G-Hyl by galactosylhydroxylysyl glucosyltransferase (GGT) (EC 2.4.1.66) forming glucosylgalactosylhydroxylysine (GG-Hyl) (103). These enzymes were previously characterized from whole chick embryo until a decade ago when these functions were discovered in the multifunctional collagen modifying enzyme, lysyl hydroxylase 3 (LH3) (79-80,108). Characterization of LH3 deficient mice showed an embryonic lethal phenotype at E9.5 due to the impaired type IV and VI collagen biosynthesis and the severity of the

phenotype showed a significant correlation to the level of GGT activity (69-70,99) *(reviewed in section 1.6.2).*

As for the GT activity of LH3, the data has been conflicting (70,79-81) and the biological significance was unclear until most recently, a novel family of collagen galactosyltransferase enzyme encoded by the *GLT25D1* and *GLT25D2* has been discovered and characterized, showing the capabilities of transferring galactose units, but not glucose, to collagens and MBL (81). Unlike LH3, GLT25D1 and GLT25D2, was shown to be strictly maintained in the ER by the ER retention motif, RDEL (81,111). Among them, *GLT25D1* exhibited broad expression in fetal and adult human tissues, and was suggested to be the main collagen galactosyltransferase enzyme, while *GLT25D2* was only detected in a few cell types (81). So, there is a high possibility that Hyl galactosylation would be mainly modulated by GLT25D1 and subsequent glucosylation by LH3.

From our previous studies (Chapter 3 and 4) in mouse osteoblast cell-line (MC3T3-E1 (MC)), we have demonstrated that for type I collagen, LH3 mainly catalyzes the final transfer of glucose units to G-Hyl and thus leaving the regulation of the intermediate step, galactosylation of Hyl, unclear. However, since then we have also found that in MC cells, *Glt25d1* is the only isoform expressed, not *Glt25d2*. This had led us to hypothesize that Glt25d1 may galactosylate Hyl residues while LH3 glucosylates G-Hyl residues. In order to elucidate the functional involvement of Glt25d1 in the sequential glycosylation process of Hyl in bone type I collagen, we have employed the short hairpin RNA technology, successfully used for the suppression of LH3 in MC cells (Chapter 3). Single cell-derived clones stably suppressing Glt25d1 were generated and their synthesized type I collagen analyzed for the levels of Hyl glycosylations. The results shown here, though preliminary,

are supportive of the hypothesis that Glt25d1 modulates the GT activity in bone type I collagen.

Experimental Procedures

Cell Lines and Culture Conditions

MC3T3-E1 subclone #4 (MC), a well characterized nontransformed mouse osteoblast-like cell line (170) was purchased from American Type Culture Collection (CRL-2593). The cells were grown in α-MEM (Invitrogen, Carlsbad, CA, USA) containing 10% FBS (Invitrogen) and supplemented with 100 U/ml penicillin G sodium, 100 μg/ml streptomycin sulfate. The cells were cultured in a 5% CO_2 atmosphere at 37ºC and the medium was changed twice a week.

Generation of short hairpin expressing plasmid targeting Glt25d1

The short hairpin expression plasmids were generated by the same method described in detail in Chapter 3. Briefly, the target sequences for suppression of mouse *Glt25d1* were designed (si#1 GGAATAATGAGCAGGTTAA [start position 1891], si#2 AGGCAGAGGTGCAAATGTA [985], si#3 TGTCAGAATACAAGAGTCA [1762]) and converted into short hairpin oligonucleotide sequences specific for pSilencer2.1 vector. The two complimentary 55-mer oligonucleotides for each target were ordered, annealed then ligated into the pSilencer2.1-U6/neo vector (Ambion, Austin, TX, USA), linearized with restriction enzymes digestion, and transformed into bacteria competent cells. The clones with 100% nucleotide sequence accuracy (Eton Bioscience, Durham, NC, USA) for each target (G_Sh1, 2 and 3) were selected and tested for their suppression efficiency by transient transfection into MC cells using FuGENE6 transfection reagent (Roche, Indianapolis, IN, USA). The non-transfected MC cells and those transfected with the original pSilencer2.1-

U6/neo plasmid (EV; encoding a hairpin siRNA sequence not found in any genome database) were used as controls. After 48 hours, total RNA was extracted, converted into cDNA by using the Omniscript Reverse Transcriptase kit (Qiagen, Valencia, CA, USA). Quantitative real-time PCR was performed using the sequence specific primer for Glt25d1 (*Glt25d1*; Mm00600638_m1). The reactions were prepared and analyzed in triplicates by the ABI Prism 7000 Sequence detection system (Applied Biosystems, Foster city, CA, USA). The mRNA expression level of Glt25d1 relative to glyceraldehyde 3-phosphate dehydrogenase (GAPDH; ABI assay number 4308313) was analyzed by the $2^{-\Delta\Delta CT}$ method (171). The suppression efficiency among the different ShRNA target sites was then compared in percentage relative to the controls.

Generation and characterization of MC single cell-derived clones stably suppressing Glt25d1 (G_Sh clones)

MC cells were transfected with the selected pSilencer2.1-U6/neo-G_Sh plasmid by using FuGENE6 transfection reagent (Roche), and isolation of single cell-derived clones stably suppressing Glt25d1 was performed as described in Chapter 3. The selected G_Sh clones were then characterized by comparing the level of *Glt25d1* expression to those of the EV clone (generated and characterized in Chapter 3) and the non-transfected MC cells, by using quantitative real-time PCR. Three to six Sh clones with different levels of *Glt25d1* suppression were then characterized for their effects on the mRNA expression of LH1 (*Plod1*; ABI assay number: Mm00599925_m1), LH2 (*Plod2*; Mm00478767_m1), LH3 (*Plod3*; Mm00478798_m1), and type I collagen α2 chain (*Col1A2*; Mm00483888_m1).

Purification of type I collagen

Type I collagen without telopeptides was purified from the 2-week cell/matrix layers of MC cells, EV and G_Sh clones with the same methods, described in Chapter 3.

Quantification of GG-Hyl, G-Hyl and Free Hyl

Lyophilized collagen samples from MC cells, EV and G_Sh clones were acid and base hydrolyzed, and analyzed by amino acid analysis as described in detail in Chapter 3.

Results

Generation of clones stably suppressing Glt25d1

The pSilencer2.1-U6/neo-G_Sh plasmids (G_Sh1, 2 and 3) were transiently transfected into MC cells and the expressions of *Glt25d1* and *Col1A2* were analyzed by quantitative real-time PCR. From the cells transfected with the short hairpin constructs (G_Sh1, 2 and 3), the expression of *Glt25d1* relative to *GAPDH* was 36%, 39% and 35% respectively, of 100% expression in MC non-transfected cells. The expressions of Glt25d1 relative to Col1A2 are shown in Figure 5.1. Transfection of the G_Sh1 construct resulted in a lowered level of Glt25d1 relative to Col1A2, while the G_Sh2 and 3 constructs showed off target effects on *Col1A2* expression. The effects led to higher expression levels of Glt25d1 relative to Col1A2, when compared to the controls, MC and EV. Since our aim is to analyze the effects of gene manipulation on the levels of collagen post-translational modification, the relative expression of the gene to *Col1A2* is an important factor to be taken into consideration. Therefore the target sequence #1 (G_Sh 1), with the least off target effect on

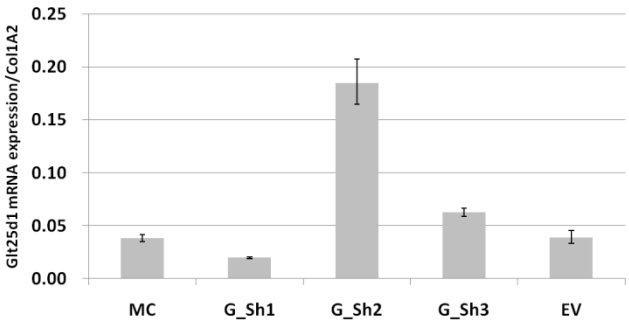

Figure 5.1) Gene expression of *Glt25d1* from transient transfection of short hairpin constructs targeting *Glt25d1*. pSilencer2.1-U6/neo-G_Sh1, 2, 3 and EV plasmids were individually transfected into MC cells. After 48 hours of culture, real time PCR was performed to determine the suppression levels of *Glt25d1* and the expression of *Col1A2*. The *Glt25d1* mRNA expression was normalized to the expression of *Col1A2*.

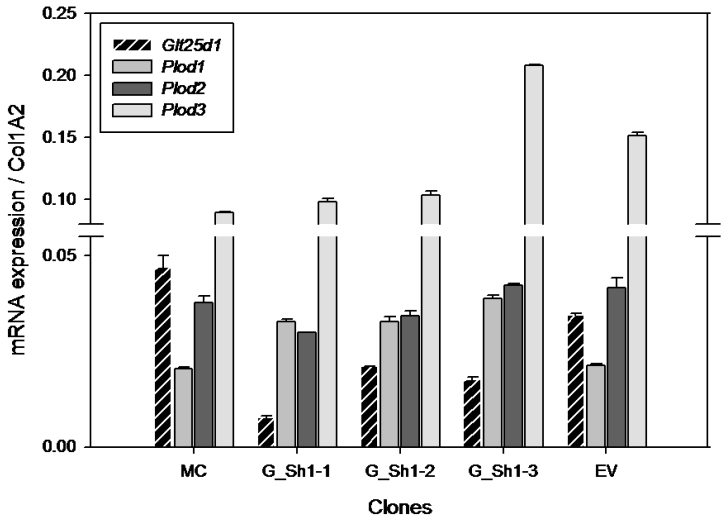

Figure 5.2) Gene expression of *Glt25d1* and *Plod* family in single cell-derived clones stably suppressing Glt25d1 (G_Sh clones) After 48 hours of culture, real time PCR was performed to determine the suppression levels of *Glt25d1* in the single cell-derived G_Sh clones and its potential effects on *Plod1, Plod2* and *Plod3* expression. The mRNA expression was normalized to the expression of *Col1A2.*

Col1A2, was selected for further generation of single cell-derived clones stably suppressing Glt25d1 by constant selection with G418.

Characterization of G_Sh1-derived clones stably suppressing Glt25d1

Several clones were initially isolated and characterized based on their cell morphology and growth pattern compared to controls, MC and EV. Three representative clones (G_Sh1-1, -2, and -3) showing similar growth patterns to the controls, were selected for further phenotypic characterization. Real-time PCR analysis was performed to assess the levels of *Glt25d1* and the effects of its suppression on the expression of *Plod* family (*Plod1, Plod2* and *Plod3* encoding LH1, LH2 and LH3 respectively), relative to *Col1A2*. In comparison to MC non-transfected control (100%), the expression of *Glt25d1*in the G_Sh 1-1,-2 and -3 clones were 16%, 45% and 37% respectively. Figure 5.2 demonstrated the effect

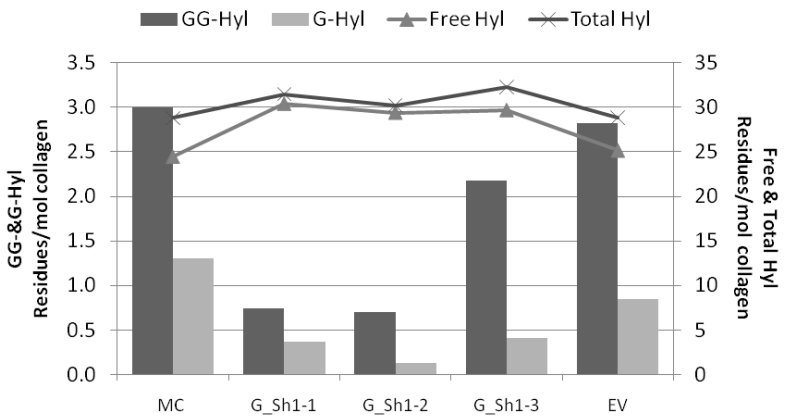

Figure 5.3) The modification levels of Hyl residues in the purified type I collagen synthesized by G_Sh1 clones and controls. The levels of total Hyl and each Hyl species (GG-Hyl, G-Hyl and free Hyl) in the purified type I collagen synthesized by G_Sh1 clones and controls, MC and EV. The levels are shown as residues/mol collagen.

of *Glt25d1* suppression on the levels of *Plod* family. In the G_Sh clones, the expression of *Plod1* was ~1.5-2 fold higher, while the expression of *Plod2* was essentially comparable to MC and EV. On the contrary, *Plod3* was upregulated in the G_Sh as well as in the EV clones, ranging from 1.1-2.3 fold increase, suggesting the effect of transfection as previously discussed in Chapter 3.

Quantification of GG-Hyl, G-Hyl and Free Hyl

Type I collagen without telopeptides purified from MC cells, EV and G_Sh clones (G_Sh1-1, -2 and -3) was subjected to acid and base hydrolysis and analyzed by the newly established amino acid analysis program described in Chapter 3. Figure 5.3 shows the levels of total Hyl and each Hyl species (GG-Hyl, G-Hyl and free Hyl) expressed as residues/mol collagen. The levels of total Hyl in all the G_Sh collagens are slightly higher than those of MC and EV. The observation in G_Sh collagen of significant decrease in the level of Hyl glycosylation (GG-Hyl + G-Hyl) with a simultaneous increase in the level of free Hyl

indicates the decreased transfer of galactose units to Hyl residues. The reduced levels of G-Hyl residues in the G_Sh collagens were further glucosylated by the endogenous LH3 and resulted in the formation of GG-Hyl residues at levels lower than those of MC and EV. In the G_Sh clones, the levels of GG-Hyl residues appeared to be correlated to the expression levels of LH3 gene, however in EV, the level of GG-Hyl was slightly lower than in MC, despite the higher LH3 expression. Potentially, the lower *Glt25d1* expression in EV may contribute to the lower levels of G-Hyl residues to function as substrates for the GGT activity of LH3.

Discussion

Prior to the discovery of the novel collagen galactosyltransferase enzymes, GLT25D1 and GLT25D2, the biological significance of the GT activity identified in LH3 has been in question. Conflicting results from several studies have been reported showing low level of activity (79-80) or undetectable (70,81). In discovering the novel enzymes, Schegg et al have characterized the enzymatic functions by means of GT and GGT activity assays and showed that GLT25D1 and D2 were capable of transferring the galactose units to collagen substrates and mannose-binding lectin, while LH3 was strictly a GGT enzyme (81). In Chapter 3, we have shown that, firstly, LH3 mainly modulates the GGT but not the GT activity for type I collagen. Secondly, we have also shown that Glt25d1 is the only isoform expressed in mouse osteoblasts which may suggest its potential role as a GT enzyme.

Therefore, in this preliminary study, we have attempted to uncover the biological function of Glt25d1 for type I collagen. Type I collagen purified from the single cell-derived clones stably suppressing Glt25d1 was shown to be, overall, less glycosylated with concomitant increase in the level of free Hyl. This clearly indicates that by suppressing Glt25d1, the transfer of galactose units to Hyl was impaired; hence the higher levels of free Hyl and less G-Hyl substrate for the GGT activity of LH3. Moreover, this is another confirmation of the insignificance of the GT activity of LH3, since upregulation of *Plod3* expression was observed in the G_Sh clones; however the galactosylation level was still very low. Unfortunately, the Glt25d1 protein levels in the controls and G_Sh clones could not be effectively investigated due to lack of an antibody with good sensitivity and specificity to the protein. Several attempts to optimize the Western blot condition using two commercially available Glt25d1 antibodies have so far been unsuccessful. Further phenotypic

characterization of MC cell-derived G_Sh clones in regards to collagen fibrillogenesis, cross-linking profile, extracellular matrix formation and mineralization are still ongoing.

In conclusion, the preliminary results shown here are in good support of our hypothesis that in bone type I collagen, Glt25d1 catalyzes the transfer of galactose units to Hyl residues and subsequently, LH3 modulates the final transfer of glucose units to G-Hyl residues. To date, our studies would be the first to elucidate the enzymatic regulatory mechanisms of Hyl galactosylation and subsequent glucosylation in type I collagen from osteoblasts culture by direct measurement of the G-Hyl and GG-Hyl respectively.

Chapter 6

Concluding Remarks

For type I collagen, until now, the enzymatic controls of the formation and the biological functions of the glycosylated Hyl (G-Hyl and GG-Hyl) residues had remained elusive. The GT and GGT activities were shown to be modulated by LH3 (79-80) until the recent identification of GLT25D1 and D2 as novel collagen GT enzymes (81). To uncover their roles for type I collagen, we have suppressed the LH3 encoding *Plod3* and *Glt25d1* (the only isoform expressed) in mouse osteoblast cell line, MC3T3-E1 (MC) and generated single cell-derived clones (Sh and G_Sh clones respectively). By characterizing the clones, we have discovered the following:

1. The major function of LH3 is to transfer glucose units to G-Hyl residues in the α chains of bone type I collagen as determined by amino acid analyses of collagen from the Sh clones and the HPLC-based glycosyltransferase activity assays using purified recombinant LH3-V5/His tagged protein.
2. The less glucosylated type I collagen, from the Sh clones, showed different in vitro fibril formation kinetics when compared to the controls, MC and EV.
3. By utilizing mass spectrometry (MS), specific molecular loci of Hyl glycosylation and forms (G-Hyl and GG-Hyl) in the α chains of MC collagen were identified at residues α1-87, α1-174 and α2-173. In the collagen from the LH3-suppressed clone (Sh1

4. collagen), the presence of GG-Hyl was significantly diminished while G-Hyl was the major form observed; ultimately confirming the major role of LH3 as a glucosyltransferase enzyme.
5. From collagen cross-links analyses, the levels of total cross-links (total aldehydes) formation were significantly lower in the Sh clones, despite the higher expression of LOX enzymes. It may suggest that by suppressing LH3 the less glucosylated collagen may not be a good substrate for the LOX enzymes; hence the lower levels of aldehydes.
6. The levels of total glycosylated immature bivalent cross-links (mainly DHLNL) in the Sh clones were significantly lower than those in the controls, with higher levels of GG-DHLNL over G-DHLNL unlike the higher levels of G-Hyl over GG-Hyl at α1-87 observed in the MS analysis of Sh collagen. These data may indicate that, at this major cross-linking site, GG-Hyl is the preferred form for DHLNL formation.
7. From in vitro cross-links maturation study, there was a faster decline in the levels of total DHLNL in the controls than in the Sh clones while the formation of Pyr are overall unproportionally low. The difference in the levels of glucosylation between the controls and Sh clones, along with previous reports showing only the existence of G-Pyr in bone (38,121-122), may point to the possibility that GG-DHLNL could be involved in the unidentified loss of total DHLNL, hence the different rate observed between the Sh clones and controls.
8. Thin collagen fibers were formed in the cultures of Sh clone suppressing LH3, when compared to the controls, as shown by picrosirius red staining. In addition, these thin collagen fibers may not be suitable templates for mineral deposition since significant delay in the mineralized nodules formation was observed from Alizarin Red S staining.

These phenomena may result from the possibility that collagen with lower levels of glucosylation may be less stabilized in the extracellular matrix, since the levels of cross-links in the Sh clones are significantly lower, or it could be subjected to enzymatic degradation digestion (9).

9. Glt25d1 is the major galactosyltransferase enzyme for type I collagen, shown from the characterization of collagen from G_Sh clones by amino acid analysis.

Collectively, the findings from this study clearly demonstrated that, for bone type I collagen, the galactosylation of Hyl residues is modulated be Glt25d1 and the subsequent glucosylation of G-Hyl is catalyzed by LH3. The glucose units in the specific GG-Hyl residues appeared to have positive regulatory roles in the formation of intermolecular cross-links and normal collagen template for the mineralization process. These findings may have given new insights to the biological significance of type I collagen *O*-glycosylation in bone physiology.

The association of bone type I collagen overglycosylation with the poor bone quality phenotype has been reported in several human osteogenic disorders such as osteogenesis imperfecta (OI) (156,160), post-menopausal osteoporosis (155,158,165-166), osteosarcoma and osteofibrous dysplasia (159), but the mechanisms are only partially understood. Mutations in the α chains of type I collagen, in some subtypes of OI, impaired the folding of α chains into triple helix and higher levels of post-translational modifications were observed (reviewed in (201). Moreover, impaired function of the collagen chaperone complex, P3H/CRTAP/CyPB complex, is shown to be the cause of recessive OIs with overmodified collagen and normal *COL1A1* and *COL1A2* sequences (202). However, the mechanism

behind the association between overmodified/overglycosylated collagen and poor bone quality is still unclear.

In order to elucidate the effect of collagen overglycosylation on the mechanism of mineralization, based on what we have known from this study, future experiments can be perfomed, *in vitro,* to simulate the overglycosylation phenotype in bone type I collagen by utilizing the gain-of-function approach in MC cell. By overexpressing Glt25d1, it is likely that the increased level of G-Hyl residues will be subjected to further glucosylation by the potent LH3, endogenously expressed. Alternatively, a combination of LH1 and Glt25d1 overexpression may result in a higher degree of glycosylation in collagen, comparing to transfection of Glt25d1 alone, since total Hyl in the helical region may be increased from the higher level of LH1. If successfully generated, these overexpressed clones showing overglycosylated collagen may be subjected to *in vitro* mineralization assay and/or *in vivo* transplantation model (17,172) to evaluate the mineralized nodules or the quality of the bone tissue formed. Eventually, the thorough understanding of the regulatory mechanisms of the biosynthesis of type I collagen and its formation of extracellular matrix, the basic structure in bone tissue, may lead to the development of effective interventions to improve the quality of bone for the patients.

Bibliography

1. Gordon, M. K., and Hahn, R. A. (2010) *Cell Tissue Res* **339**, 247-257

2. Carter, E. M., and Raggio, C. L. (2009) *Curr Opin Pediatr* **21**, 46-54

3. Hulmes, D. J. S. (2008) Collagen Diversity, Synthesis and Assembly. in *Collagen* (Fratzl, P. ed.), Springer US. pp 15-47

4. Franzke, C. W., Bruckner, P., and Bruckner-Tuderman, L. (2005) *J Biol Chem* **280**, 4005-4008

5. Shinkai, H., and Yonemasu, K. (1979) *Biochem J* **177**, 847-852

6. Yamauchi, M. (2002) Collagen Biochemistry: An Overview. in *Advance in Tissue Banking*, World Scientific Publishing Co. Pte. Ltd. pp 445-500

7. Kadler, K. E., Holmes, D. F., Trotter, J. A., and Chapman, J. A. (1996) *Biochem J* **316 (Pt 1)**, 1-11

8. Amudeswari, S., Liang, J. N., and Chakrabarti, B. (1987) *Coll Relat Res* **7**, 215-223

9. Yang, C. L., Rui, H., Mosler, S., Notbohm, H., Sawaryn, A., and Muller, P. K. (1993) *Eur J Biochem* **213**, 1297-1302

10. Notbohm, H., Nokelainen, M., Myllyharju, J., Fietzek, P. P., Muller, P. K., and Kivirikko, K. I. (1999) *J Biol Chem* **274**, 8988-8992

11. Torre-Blanco, A., Adachi, E., Hojima, Y., Wootton, J. A., Minor, R. R., and Prockop, D. J. (1992) *J Biol Chem* **267**, 2650-2655

12. Vogel, K. G., Paulsson, M., and Heinegard, D. (1984) *Biochem J* **223**, 587-597

13. Pogany, G., Hernandez, D. J., and Vogel, K. G. (1994) *Arch Biochem Biophys* **313**, 102-111

14. Ameye, L., and Young, M. F. (2002) *Glycobiology* **12**, 107R-116R

15. Reed, C. C., and Iozzo, R. V. (2002) *Glycoconj J* **19**, 249-255

16. Corsi, A., Xu, T., Chen, X. D., Boyde, A., Liang, J., Mankani, M., Sommer, B., Iozzo, R. V., Eichstetter, I., Robey, P. G., Bianco, P., and Young, M. F. (2002) *J Bone Miner Res* **17**, 1180-1189

17. Mochida, Y., Parisuthiman, D., Pornprasertsuk-Damrongsri, S., Atsawasuwan, P., Sricholpech, M., Boskey, A. L., and Yamauchi, M. (2009) *Matrix Biol* **28**, 44-52

18. Viguet-Carrin, S., Garnero, P., and Delmas, P. D. (2006) *Osteoporos Int* **17**, 319-336

19. Lucero, H. A., and Kagan, H. M. (2006) *Cell Mol Life Sci* **63**, 2304-2316

20. Siegel, R. C., Fu, J. C., and Chang, Y. (1976) *Adv Exp Med Biol* **74**, 438-446

21. Cronlund, A. L., Smith, B. D., and Kagan, H. M. (1985) *Connect Tissue Res* **14**, 109-119

22. Bailey, A. J., Paul, R. G., and Knott, L. (1998) *Mech Ageing Dev* **106**, 1-56

23. Kagan, H. M. (2000) *Acta Trop* **77**, 147-152

24. Avery, N. C., and Bailey, A. J. (2008) Restraining Cross-Links Responsible for the Mechanical Properties of Collagen Fibers: Natural and Artificial. in *Collagen* (Fratzl, P. ed.), Springer US. pp 81-110

25. Bailey, A. J., and Shimokomaki, M. S. (1971) *FEBS Lett* **16**, 86-88

26. Robins, S. P., Shimokomaki, M., and Bailey, A. J. (1973) *Biochem J* **131**, 771-780

27. Robins, S. P., and Bailey, A. J. (1977) *Biochim Biophys Acta* **492**, 408-414

28. Eyre, D. R., Dickson, I. R., and Van Ness, K. (1988) *Biochem J* **252**, 495-500

29. Eyre, D. R. (1981) *Crosslink Maturation in Bone Collagen*, Elsevier North Holland, Inc., New York

30. Fujimoto, D., Moriguchi, T., Ishida, T., and Hayashi, H. (1978) *Biochem Biophys Res Commun* **84**, 52-57

31. Ogawa, T., Ono, T., Tsuda, M., and Kawanishi, Y. (1982) *Biochem Biophys Res Commun* **107**, 1252-1257

32. Eyre, D. R., Paz, M. A., and Gallop, P. M. (1984) *Annu Rev Biochem* **53**, 717-748

33. Yamauchi, M., and Katz, E. P. (1993) *Connect Tissue Res* **29**, 81-98

34. Eyre, D. R., and Oguchi, H. (1980) *Biochem Biophys Res Commun* **92**, 403-410

35. Robins, S. P., and Duncan, A. (1983) *Biochem J* **215**, 175-182

36. Yamauchi, M., and Mechanic, G. (1988) Cross-linking of Collagen. in *Collagen* (Nimni, M. E. ed.), CRC Press, Boca Raton. pp 157-172

37. Kuypers, R., Tyler, M., Kurth, L. B., Jenkins, I. D., and Horgan, D. J. (1992) *Biochem J* **283 (Pt 1)**, 129-136

38. Hanson, D. A., and Eyre, D. R. (1996) *J Biol Chem* **271**, 26508-26516

39. Brady, J. D., and Robins, S. P. (2001) *J Biol Chem* **276**, 18812-18818

40. Landis, W. J. (1999) *Gravit Space Biol Bull* **12**, 15-26

41. Gajjeraman, S., Narayanan, K., Hao, J., Qin, C., and George, A. (2007) *J Biol Chem* **282**, 1193-1204

42. Siperko, L. M., and Landis, W. J. (2001) *J Struct Biol* **135**, 313-320

43. Lees, S. (2003) *Biophys J* **85**, 204-207

44. Vashishth, D. (2007) *Curr Osteoporos Rep* **5**, 62-66

45. Camacho, N. P., Landis, W. J., and Boskey, A. L. (1996) *Connect Tissue Res* **35**, 259-265

46. Grabner, B., Landis, W. J., Roschger, P., Rinnerthaler, S., Peterlik, H., Klaushofer, K., and Fratzl, P. (2001) *Bone* **29**, 453-457

47. Paschalis, E. P., Shane, E., Lyritis, G., Skarantavos, G., Mendelsohn, R., and Boskey, A. L. (2004) *J Bone Miner Res* **19**, 2000-2004

48. Mochida, Y., Duarte, W. R., Tanzawa, H., Paschalis, E. P., and Yamauchi, M. (2003) *Biochem Biophys Res Commun* **305**, 6-9

49. Pornprasertsuk, S., Duarte, W. R., Mochida, Y., and Yamauchi, M. (2005) *J Bone Miner Res* **20**, 81-87

50. Chen, C. C., and Boskey, A. L. (1985) *Calcif Tissue Int* **37**, 395-400

51. Hoshi, K., Kemmotsu, S., Takeuchi, Y., Amizuka, N., and Ozawa, H. (1999) *J Bone Miner Res* **14**, 273-280

52. Myllyla, R., Wang, C., Heikkinen, J., Juffer, A., Lampela, O., Risteli, M., Ruotsalainen, H., Salo, A., and Sipila, L. (2007) *J Cell Physiol* **212**, 323-329

53. Yeowell, H. N., and Walker, L. C. (1999) *Matrix Biol* **18**, 179-187

54. Mercer, D. K., Nicol, P. F., Kimbembe, C., and Robins, S. P. (2003) *Biochem Biophys Res Commun* **307**, 803-809

55. Pornprasertsuk, S., Duarte, W. R., Mochida, Y., and Yamauchi, M. (2004) *J Bone Miner Res* **19**, 1349-1355

56. Valtavaara, M., Papponen, H., Pirttila, A. M., Hiltunen, K., Helander, H., and Myllyla, R. (1997) *J Biol Chem* **272**, 6831-6834

57. Passoja, K., Rautavuoma, K., Ala-Kokko, L., Kosonen, T., and Kivirikko, K. I. (1998) *Proc Natl Acad Sci U S A* **95**, 10482-10486

58. Ruotsalainen, H., Sipila, L., Kerkela, E., Pospiech, H., and Myllyla, R. (1999) *Matrix Biol* **18**, 325-329

59. Schneider, V. A., and Granato, M. (2006) *Neuron* **50**, 683-695

60. Schneider, V. A., and Granato, M. (2007) *Matrix Biol* **26**, 12-19

61. Norman, K. R., and Moerman, D. G. (2000) *Dev Biol* **227**, 690-705

62. Takaluoma, K., Lantto, J., and Myllyharju, J. (2007) *Matrix Biol* **26**,396-403

63. Risteli, M., Niemitalo, O., Lankinen, H., Juffer, A. H., and Myllyla, R. (2004) *J Biol Chem* **279**, 37535-37543

64. Wang, C., Valtavaara, M., and Myllyla, R. (2000) *DNA Cell Biol* **19**, 71-77

65. Uzawa, K., Yeowell, H. N., Yamamoto, K., Mochida, Y., Tanzawa, H., and Yamauchi, M. (2003) *Biochem Biophys Res Commun* **305**, 484-487

66. Zuurmond, A. M., van der Slot-Verhoeven, A. J., van Dura, E. A., De Groot, J., and Bank, R. A. (2005) *Matrix Biol* **24**, 261-270

67. Bank, R. A., Robins, S. P., Wijmenga, C., Breslau-Siderius, L. J., Bardoel, A. F., van der Sluijs, H. A., Pruijs, H. E., and TeKoppele, J. M. (1999) *Proc Natl Acad Sci U S A* **96**, 1054-1058

68. Takaluoma, K., Hyry, M., Lantto, J., Sormunen, R., Bank, R. A., Kivirikko, K. I., Myllyharju, J., and Soininen, R. (2007) *J Biol Chem* **282**, 6588-6596

69. Ruotsalainen, H., Sipila, L., Vapola, M., Sormunen, R., Salo, A. M., Uitto, L., Mercer, D. K., Robins, S. P., Risteli, M., Aszodi, A., Fassler, R., and Myllyla, R. (2006) *J Cell Sci* **119**, 625-635

70. Rautavuoma, K., Takaluoma, K., Sormunen, R., Myllyharju, J., Kivirikko, K. I., and Soininen, R. (2004) *Proc Natl Acad Sci U S A* **101**, 14120-14125

71. Uzawa, K., Grzesik, W. J., Nishiura, T., Kuznetsov, S. A., Robey, P. G., Brenner, D. A., and Yamauchi, M. (1999) *J Bone Miner Res* **14**, 1272-1280

72. van der Slot, A. J., Zuurmond, A. M., Bardoel, A. F., Wijmenga, C., Pruijs, H. E., Sillence, D. O., Brinckmann, J., Abraham, D. J., Black, C. M., Verzijl, N., DeGroot, J., Hanemaaijer, R., TeKoppele, J. M., Huizinga, T. W., and Bank, R. A. (2003) *J Biol Chem* **278**, 40967-40972

73. Eyre, D., Shao, P., Weis, M. A., and Steinmann, B. (2002) *Mol Genet Metab* **76**, 211-216

74. Suokas, M., Myllyla, R., and Kellokumpu, S. (2000) *J Biol Chem* **275**, 17863-17868

75. Suokas, M., Lampela, O., Juffer, A. H., Myllyla, R., and Kellokumpu, S. (2003) *Biochem J* **370**, 913-920

76. Salo, A. M., Wang, C., Sipila, L., Sormunen, R., Vapola, M., Kervinen, P., Ruotsalainen, H., Heikkinen, J., and Myllyla, R. (2006) *J Cell Physiol* **207**, 644-653

77. Myllyla, R., Risteli, L., and Kivirikko, K. I. (1975) *Eur J Biochem* **58**, 517-521

78. Oikarinen, A., Anttinen, H., and Kivirikko, K. I. (1977) *Biochem J* **166**, 357-362

79. Heikkinen, J., Risteli, M., Wang, C., Latvala, J., Rossi, M., Valtavaara, M., and Myllyla, R. (2000) *J Biol Chem* **275**, 36158-36163

80. Wang, C., Luosujarvi, H., Heikkinen, J., Risteli, M., Uitto, L., and Myllyla, R. (2002) *Matrix Biol* **21**, 559-566

81. Schegg, B., Hulsmeier, A. J., Rutschmann, C., Maag, C., and Hennet, T. (2009) *Mol Cell Biol* **29**, 943-952

82. Yeowell, H. N., Walker, L. C., Farmer, B., Heikkinen, J., and Myllyla, R. (2000) *Hum Mutat* **16**, 90

83. Yeowell, H. N., Allen, J. D., Walker, L. C., Overstreet, M. A., Murad, S., and Thai, S. F. (2000) *Matrix Biol* **19**, 37-46

84. Walker, L. C., Marini, J. C., Grange, D. K., Filie, J., and Yeowell, H. N. (1999) *Mol Genet Metab* **67**, 74-82

85. Pousi, B., Heikkinen, J., Schroter, J., Pope, M., and Myllyla, R. (2000) *Mutat Res* **432**, 33-37

86. Yeowell, H. N., and Walker, L. C. (1997) *Proc Assoc Am Physicians* **109**, 383-396

87. Ha, V. T., Marshall, M. K., Elsas, L. J., Pinnell, S. R., and Yeowell, H. N. (1994) *J Clin Invest* **93**, 1716-1721

88. Eyre, D. R., and Glimcher, M. J. (1972) *Proc Natl Acad Sci U S A* **69**, 2594-2598

89. Pasquali, M., Still, M. J., Vales, T., Rosen, R. I., Evinger, J. D., Dembure, P. P., Longo, N., and Elsas, L. J. (1997) *Proc Assoc Am Physicians* **109**, 33-41

90. Giunta, C., Randolph, A., Al-Gazali, L. I., Brunner, H. G., Kraenzlin, M. E., and Steinmann, B. (2005) *Am J Med Genet A* **133A**, 158-164

91. Walker, L. C., Overstreet, M. A., Willing, M. C., Marini, J. C., Cabral, W. A., Pals, G., Bristow, J., Atsawasuwan, P., Yamauchi, M., and Yeowell, H. N. (2004) *Am J Med Genet A* **131**, 155-162

92. Walker, L. C., Teebi, A. S., Marini, J. C., De Paepe, A., Malfait, F., Atsawasuwan, P., Yamauchi, M., and Yeowell, H. N. (2004) *Mol Genet Metab* **83**, 312-321

93. Miyake, N., Kosho, T., Mizumoto, S., Furuichi, T., Hatamochi, A., Nagashima, Y., Arai, E., Takahashi, K., Kawamura, R., Wakui, K., Takahashi, J., Kato, H., Yasui, H., Ishida, T., Ohashi, H., Nishimura, G., Shiina, M., Saitsu, H., Tsurusaki, Y., Doi, H., Fukushima, Y., Ikegawa, S., Yamada, S., Sugahara, K., and Matsumoto, N. (2010) *Hum Mutat* **31**, 966-974

94. Malfait, F., Syx, D., Vlummens, P., Symoens, S., Nampoothiri, S., Hermanns-Le, T., Van Laer, L., and De Paepe, A. (2010) *Hum Mutat* Epub Date 09-16-10

95. Giunta, C., Elcioglu, N. H., Albrecht, B., Eich, G., Chambaz, C., Janecke, A. R., Yeowell, H., Weis, M., Eyre, D. R., Kraenzlin, M., and Steinmann, B. (2008) *Am J Hum Genet* **82**, 1290-1305

96. Ha-Vinh, R., Alanay, Y., Bank, R. A., Campos-Xavier, A. B., Zankl, A., Superti-Furga, A., and Bonafe, L. (2004) *Am J Med Genet A* **131**, 115-120

97. Hyry, M., Lantto, J., and Myllyharju, J. (2009) *J Biol Chem* **284**, 30917-30924

98. Kelley, B. P., Malfait, F., Bonafe, L., Baldridge, D., Homan, E., Symoens, S., Willaert, A., Elcioglu, N., Van Maldergem, L., Verellen-Dumoulin, C., Gillerot, Y., Napierala, D., Krakow, D., Beighton, P., Superti-Furga, A., De Paepe, A., and Lee, B. (2010) *J Bone Miner Res* Epub Date 09-13-10

99. Sipila, L., Ruotsalainen, H., Sormunen, R., Baker, N. L., Lamande, S. R., Vapola, M., Wang, C., Sado, Y., Aszodi, A., and Myllyla, R. (2007) *J Biol Chem* **282**, 33381-33388

100. Salo, A. M., Cox, H., Farndon, P., Moss, C., Grindulis, H., Risteli, M., Robins, S. P., and Myllyla, R. (2008) *Am J Hum Genet* **83**, 495-503

101. Risteli, M., Ruotsalainen, H., Salo, A. M., Sormunen, R., Sipila, L., Baker, N. L., Lamande, S. R., Vimpari-Kauppinen, L., and Myllyla, R. (2009) *J Biol Chem* **284**, 28204-28211

102. Spiro, R. G. (1967) *J Biol Chem* **242**, 4813-4823

103. Kivirikko, K. I., and Myllyla, R. (1979) *Int Rev Connect Tissue Res* **8**, 23-72

104. Spiro, R. G., and Spiro, M. J. (1971) *J Biol Chem* **246**, 4899-4909

105. Spiro, M. J., and Spiro, R. G. (1971) *J Biol Chem* **246**, 4910-4918

106. Anttinen, H., Myllyla, R., and Kivirikko, K. I. (1977) *Eur J Biochem* **78**, 11-17

107. Myllyla, R., Risteli, L., and Kivirikko, K. I. (1976) *Eur J Biochem* **61**, 59-67

108. Wang, C., Risteli, M., Heikkinen, J., Hussa, A. K., Uitto, L., and Myllyla, R. (2002) *J Biol Chem* **277**, 18568-18573

109. Wang, C., Kovanen, V., Raudasoja, P., Eskelinen, S., Pospiech, H., and Myllyla, R. (2009) *J Cell Mol Med* **13**, 508-521

110. Colley, K. J., and Baenziger, J. U. (1987) *J Biol Chem* **262**, 10290-10295

111. Liefhebber, J. M., Punt, S., Spaan, W. J., and van Leeuwen, H. C. (2010) *BMC Cell Biol* **11**, 33

112. Batge, B., Winter, C., Notbohm, H., Acil, Y., Brinckmann, J., and Muller, P. K. (1997) *J Biochem (Tokyo)* **122**, 109-115

113. Butler, W. T., and Cunningham, L. W. (1966) *J Biol Chem* **241**, 3882-3888

114. Cunningham, L. W., and Ford, J. D. (1968) *J Biol Chem* **243**, 2390-2398

115. Morgan, P. H., Jacobs, H. G., Segrest, J. P., and Cunningham, L. W. (1970) *J Biol Chem* **245**, 5042-5048

116. Aguilar, J. H., Jacobs, H. G., Butler, W. T., and Cunningham, L. W. (1973) *J Biol Chem* **248**, 5106-5113

117. Fietzek, P. P., and Kuhn, K. (1976) *Int Rev Connect Tissue Res* **7**, 1-60

118. Eyre, D. R., and Glimcher, M. J. (1973) *Biochem Biophys Res Commun* **52**, 663-671

119. Robins, S. P., and Bailey, A. J. (1974) *FEBS Lett* **38**, 334-336

120. Eriksen, H. A., Sharp, C. A., Robins, S. P., Sassi, M. L., Risteli, L., and Risteli, J. (2004) *Bone* **34**, 720-727

121. Gineyts, E., Garnero, P., and Delmas, P. D. (2001) *Rheumatology (Oxford)* **40**, 315-323

122. Robins, S. P. (1983) *Biochem J* **215**, 167-173

123. Henkel, W., Rauterberg, J., and Stirtz, T. (1976) *Eur J Biochem* **69**, 223-231

124. Moro, L., Romanello, M., Favia, A., Lamanna, M. P., and Lozupone, E. (2000) *Calcif Tissue Int* **66**, 151-156

125. Spiro, R. G. (1969) *J Biol Chem* **244**, 602-612

126. Cunningham, L. W., Ford, J. D., and Segrest, J. P. (1967) *J Biol Chem* **242**, 2570-2571

127. Segrest, J. P., and Cunningham, L. W. (1970) *J Clin Invest* **49**, 1497-1509

128. Rauch, F., Schnabel, D., Seibel, M. J., Remer, T., Stabrey, A., Michalk, D., and Schonau, E. (1995) *J Clin Endocrinol Metab* **80**, 1295-1300

129. Michalsky, M., Stepan, J. J., Wilczek, H., Formankova, J., and Moro, L. (1995) *Clin Chim Acta* **234**, 101-108

130. Bettica, P., Taylor, A. K., Talbot, J., Moro, L., Talamini, R., and Baylink, D. J. (1996) *J Clin Endocrinol Metab* **81**, 542-546

131. Pecile, A., Netti, C., Sibilia, V., Villa, I., Calori, G., Tenni, R., Coluzzi, M., Moro, G. L., and Rubinacci, A. (1996) *J Endocrinol* **150**, 383-390

132. Moro, L., Noris-Suarez, K., Michalsky, M., Romanello, M., and de Bernard, B. (1993) *Biochim Biophys Acta* **1156**, 288-290

133. Al-Dehaimi, A. W., Blumsohn, A., and Eastell, R. (1999) *Clin Chem* **45**, 676-681

134. Myllyla, R., Becvar, R., Adam, M., and Kivirikko, K. I. (1989) *Clin Chim Acta* **183**, 243-252

135. Kuutti-Savolainen, E. R. (1979) *Clin Chim Acta* **96**, 53-58

136. Kuutti-Savolainen, E. R., Anttinen, H., Miettinen, T. A., and Kivirikko, K. I. (1979) *Eur J Clin Invest* **9**, 97-101

137. Myllyla, R., Myllyla, V. V., Tolonen, U., and Kivirikko, K. I. (1982) *Arch Neurol* **39**, 752-755

138. Takala, T. E., Vuori, J., Anttinen, H., Vaananen, K., and Myllyla, R. (1986) *Pflugers Arch* **407**, 500-503

139. Virtanen, P., Viitasalo, J. T., Vuori, J., Vaananen, K., and Takala, T. E. (1993) *J Appl Physiol* **75**, 1272-1277

140. Murai, A., Miyahara, T., and Shiozawa, S. (1975) *Biochim Biophys Acta* **404**, 345-348

141. Cetta, G., Tenni, R., Zanaboni, G., De Luca, G., Ippolito, E., De Martino, C., and Castellani, A. A. (1982) *Biochem J* **204**, 61-67

142. Hamazaki, H., and Hotta, K. (1979) *J Biol Chem* **254**, 9682-9687

143. Hamazaki, H., and Hotta, K. (1980) *Eur J Biochem* **111**, 587-591

144. De Luca, G., Tenni, R., Lauria, A., Cetta, G., Salvini, R., Zanaboni, G., and Castellani, A. A. (1983) *Ital J Biochem* **32**, 418-430

145. Sternberg, M., and Spiro, R. G. (1979) *J Biol Chem* **254**, 10329-10336

146. Sweeney, S. M., Orgel, J. P., Fertala, A., McAuliffe, J. D., Turner, K. R., Di Lullo, G. A., Chen, S., Antipova, O., Perumal, S., Ala-Kokko, L., Forlino, A., Cabral, W. A., Barnes, A. M., Marini, J. C., and San Antonio, J. D. (2008) *J Biol Chem* **283**, 21187-21197

147. Volpin, D., and Veis, A. (1973) *Biochemistry* **12**, 1452-1464

148. Vogel, W., Gish, G. D., Alves, F., and Pawson, T. (1997) *Mol Cell* **1**, 13-23

149. Myers, L. K., Myllyharju, J., Nokelainen, M., Brand, D. D., Cremer, M. A., Stuart, J. M., Bodo, M., Kivirikko, K. I., and Kang, A. H. (2004) *J Immunol* **172**, 2970-2975

150. Miller, E. J. (1984) Collagen Chemistry. in *Extracellular Matrix Biochemistry* (Piez, K. A., Reddi, A.H. ed.), Elsevier Science Publishing Co., Inc. pp 41-78

151. Suarez, K. N., Romanello, M., Bettica, P., and Moro, L. (1996) *Calcif Tissue Int* **58**, 65-69

152. Toole, B. P., Kang, A. H., Trelstad, R. L., and Gross, J. (1972) *Biochem J* **127**, 715-720

153. Schofield, J. D., Freeman, I. L., and Jackson, D. S. (1971) *Biochem J* **124**, 467-473

154. Savolainen, E. R., Kero, M., Pihlajaniemi, T., and Kivirikko, K. I. (1981) *N Engl J Med* **304**, 197-204

155. Dominguez, L. J., Barbagallo, M., and Moro, L. (2005) *Biochem Biophys Res Commun* **330**, 1-4

156. Tenni, R., Valli, M., Rossi, A., and Cetta, G. (1993) *Am J Med Genet* **45**, 252-256

157. Brinckmann, J., Notbohm, H., Tronnier, M., Acil, Y., Fietzek, P. P., Schmeller, W., Muller, P. K., and Batge, B. (1999) *J Invest Dermatol* **113**, 617-621

158. Michalsky, M., Norris-Suarez, K., Bettica, P., Pecile, A., and Moro, L. (1993) *Biochem Biophys Res Commun* **192**, 1281-1288

159. Lehmann, H. W., Wolf, E., Roser, K., Bodo, M., Delling, G., and Muller, P. K. (1995) *J Cancer Res Clin Oncol* **121**, 413-418

160. Bateman, J. F., Mascara, T., Chan, D., and Cole, W. G. (1984) *Biochem J* **217**, 103-115

161. Pinnell, S. R., Fox, R., and Krane, S. M. (1971) *Biochim Biophys Acta* **229**, 119-122

162. Grazioli, V., Alfano, M., Stenico, A., and Casari, E. (1996) *FEBS Lett* **388**, 134-138

163. Yang, C., Niu, C., Bodo, M., Gabriel, E., Notbohm, H., Wolf, E., and Muller, P. K. (1993) *Biochem J* **289 (Pt 3)**, 829-835

164. Zhu, S., Zhu, J., Xiao, J., Ren, L., Liu, L., and Zhou, Y. (2004) *J Huazhong Univ Sci Technolog Med Sci* **24**, 427-429

165. Moro, L., Suarez, K. N., and Romanello, M. (1997) *Eur J Clin Chem Clin Biochem* **35**, 269-273

166. Moro, L., Bettica, P., Romanello, M., and Suarez, K. N. (1997) *Eur J Clin Chem Clin Biochem* **35**, 29-33

167. Zanaboni, G., De Luca, G., Faga, A., Salvini, R., and Castellani, A. A. (1987) *Eur Surg Res* **19**, 11-15

168. Eyre, D. R., and Glimcher, M. J. (1973) *Biochem J* **135**, 393-403

169. Cetta, G., De Luca, G., Tenni, R., Zanaboni, G., Lenzi, L., and Castellani, A. A. (1983) *Connect Tissue Res* **11**, 103-111

170. Wang, D., Christensen, K., Chawla, K., Xiao, G., Krebsbach, P. H., and Franceschi, R. T. (1999) *J Bone Miner Res* **14**, 893-903

171. Livak, K. J., and Schmittgen, T. D. (2001) *Methods* **25**, 402-408

172. Parisuthiman, D., Mochida, Y., Duarte, W. R., and Yamauchi, M. (2005) *J Bone Miner Res* **20**, 1878-1886

173. Yamauchi, M., and Shiiba, M. (2008) *Methods Mol Biol* **446**, 95-108

174. Williams, B. R., Gelman, R. A., Poppke, D. C., and Piez, K. A. (1978) *J Biol Chem* **253**, 6578-6585

175. Tenni, R., Rimoldi, D., Zanaboni, G., Cetta, G., and Castellani, A. A. (1984) *Ital J Biochem* **33**, 117-127

176. Garza, H., Bennett, N., Jr., and Rodriguez, G. P. (1996) *J Chromatogr A* **732**, 385-389

177. Yamauchi, M., London, R. E., Guenat, C., Hashimoto, F., and Mechanic, G. L. (1987) *J Biol Chem* **262**, 11428-11434

178. Askenasi, R. (1973) *Biochim Biophys Acta* **304**, 375-383

179. Lou, M. F., and Hamilton, P. B. (1971) *Clin Chem* **17**, 782-788

180. Wood, G. C., and Keech, M. K. (1960) *Biochem J* **75**, 588-598

181. Odell, V., Wegener, L., Peczon, B., and Hudson, B. G. (1974) *J Chromatogr* **88**, 245-252

182. Peczon, B. D., Antreasian, R., and Bucay, P. (1979) *J Chromatogr* **169**, 351-356

183. Domon, B., and Costello, C. E. (1988) *Glycoconjugate Journal* **5**, 397-409

184. Rautavuoma, K., Takaluoma, K., Passoja, K., Pirskanen, A., Kvist, A. P., Kivirikko, K. I., and Myllyharju, J. (2002) *J Biol Chem* **277**, 23084-23091

185. Guitton, J. D., Le Pape, A., Sizaret, P. Y., and Muh, J. P. (1981) *Biosci Rep* **1**, 945-954

186. Kivirikko, K. I., and Myllyla, R. (1982) *Methods Enzymol* **82 Pt A**, 245-304

187. Batge, B., Winter, C., Notbohm, H., Acil, Y., Brinckmann, J., and Muller, P. K. (1997) *J Biochem* **122**, 109-115

188. Bigi, A., Cojazzi, G., Panzavolta, S., Ripamonti, A., Roveri, N., Romanello, M., Noris Suarez, K., and Moro, L. (1997) *J Inorg Biochem* **68**, 45-51

189. Junqueira, L. C., Bignolas, G., and Brentani, R. R. (1979) *Histochem J* **11**, 447-455

190. Dayan, D., Hiss, Y., Hirshberg, A., Bubis, J. J., and Wolman, M. (1989) *Histochemistry* **93**, 27-29

191. Robins, S. P. (1982) *Methods Biochem Anal* **28**, 329-379

192. Eyre, D. (1987) Collagen Cross-linking Amino Acids. in *Methods in Enzymology* (Colowick, S. P., and Kaplan, N. O. eds.), ACADEMIC PRESS, INC., Orlando. pp 115-139

193. Davis, N. R., Risen, O. M., and Pringle, G. A. (1975) *Biochemistry* **14**, 2031-2036

194. Gineyts, E., Borel, O., Chapurlat, R., and Garnero, P. (2010) *J Chromatogr B Analyt Technol Biomed Life Sci* **878**, 1449-1454

195. Richards, A. A., Stephens, T., Charlton, H. K., Jones, A., Macdonald, G. A., Prins, J. B., and Whitehead, J. P. (2006) *Mol Endocrinol* **20**, 1673-1687

196. Knott, L., and Bailey, A. J. (1998) *Bone* **22**, 181-187

197. Light, N., and Bailey, A. J. (1985) *FEBS Lett* **182**, 503-508

198. Light, N. D., and Bailey, A. J. (1981) *Stabilization of Bone and Dentin Collagen*, Elsevier North Holland, Inc., New york

199. Robins, S. P., and Bailey, A. J. (1975) *Biochem J* **149**, 381-385

200. Sini, P., Denti, A., Tira, M. E., and Balduini, C. (1997) *Glycoconj J* **14**, 871-874

201. Di Lullo, G. A., Sweeney, S. M., Korkko, J., Ala-Kokko, L., and San Antonio, J. D. (2002) *J Biol Chem* **277**, 4223-4231

202. Marini, J. C., Cabral, W. A., and Barnes, A. M. (2010) *Cell Tissue Res* **339**, 59-70