Amino Acids (2008) 35:703–710
DOI 10.1007/s00726-008-0073-2

REVIEW ARTICLE

The importance of proline residues in the structure, stability and susceptibility to proteolytic degradation of collagens

Stephen M. Krane

Received: 16 November 2007 / Accepted: 7 February 2008 / Published online: 23 April 2008

Abstract Collagens are among proteins that undergo several post-translational modifications, such as prolyl hydroxylation, that occur during elongation of the nascent chains in the endoplasmic reticulum. The major structural collagens, types I, II and III, have large, uninterrupted triple helices, comprising three polyproline II-like chains supercoiled around a common axis. The structure has a requirement for glycine, as every third residue, and is stabilized by the high content of proline and 4-hydroxyproline residues. Action of prolyl hydroxylases is critical. Spontaneous or targeted genetic defects in prolyl hydroxylases can be lethal or result in severe osteogenesis imperfecta. Prolines, as determinants of substrate specificity and susceptibility, also play a role in degradation of collagen by collagenolytic matrix metalloproteinases (MMPs). Targeted mutations in mice in the collagenase cleavage domain have profound effects on collagen turnover and the function of connective tissues. Prolines are thus critical determinants of collagen structure and function.

Keywords Prolylhydroxylases · Collagen structure · Osteogenesis imperfecta · Collagenases · Matrix metalloproteinases

S. M. Krane (✉)
Department of Medicine, Harvard Medical School
and the Massachusetts General Hospital, Center for Immunology
and Inflammatory Diseases, Building 149, 13th Street,
Room 8301, Boston, MA 02129, USA
e-mail: krane.stephen@mgh.harvard.edu

Introduction

Collagens with uninterrupted triple helices, such as type I collagen, are characterized biochemically by distinctive amino acid composition, i.e., on a molar basis ~33% glycine, ~10% proline, ~10% 4-hydroxyproline and ~0.1% 3-hydroxyproline. This review will focus on the unique roles of prolines and modified prolines in determining the structure and biological stability as well as proteolytic degradation of type I collagen.

Collagen structure, proline and 4-hydroxyproline

The structure of the collagens reflects the amino acid composition. Thus, there is in the triple helical domain a requirement for glycine (Gly) as every third residue and the high content of prolines (Pro) stabilizes the polyproline-II-like helices characteristic of collagen sequences (Myllyharju and Kivirikko 2004). Hydroxylation of Pro in the 4-position (4-Hyp, i.e., 4(*R*)-hydroxl-L-proline (Schumacher et al. 2006) further stabilizes the helical structure. This 4-hydroxylation takes place during elongation of the nascent polypeptide chains in the endoplasmic reticulum. 4-Hyp was isolated and its structure determined at the turn of the last century (Fischer 1902). In an extraordinary paper published over four decades later (Stetten and Schoenheimer 1944), it was reasoned that, "perhaps a part of the hydroxyproline is metabolized by way of proline". They then synthesized *l* (−) Pro where "the carbon skeleton was marked by stably bound deuterium and the amino group by N^{15}" and they fed the doubly labeled Pro to rats and measured enrichment of the isotopes in Hyp from carcass proteins. They concluded that 4-Hyp is formed *after* incorporation of Pro into these proteins

(collagens). We now know from amino acid and gene sequencing that the helical region of these collagens comprises uninterrupted repeats of Gly-X-Y tripeptides (Piez 1984; Myllyharju and Kivirikko 2004). It has been established by amino acid sequencing that whereas Pro is found in either the -X- or -Y- position of the Gly-X-Y tripeptide repeat, 4-Hyp is found only in the -Y- position. 4-Hyp is an abundant modification in types I and II collagens, with ~ 85–90 residues/1000 amino acids (~40% of the total Pro + 4-Hyp). Any substitutions for Gly residues that result from mutations in the types I, II, III or V collagen genes produce an unstable structure that results in human diseases that primarily affect type I collagen in bone (forms of "brittle bone disease", i.e., osteogenesis imperfecta), type II collagen in growth plate or articular cartilage (chondrodysplasias) or types III or V collagens in skin and other "soft "connective tissues (forms of Ehlers-Danlos syndrome) (Myllyharju and Kivirikko 2004; Byers 2000).

The triple helical structure of the fibril/fiber-forming collagens thus reflects the high content of Pro and 4-Hyp residues with a regular and uninterrupted repeat of Gly-X-Y triplets. The denaturation temperature (T_D) of collagen in solution is usually measured as the melting temperature transitition (T_m), and the T_D of native, insoluble collagens is usually measured as the shrinkage temperature (T_s). In general, the stability of the helical structure is regulated by the content of Pro *plus* 4-Hyp. These numbers have physiological meaning. It has been shown (Piez 1984) that the T_m of different animal collagens is proportional to the upper limit of environmental temperature to which the animal or tissue is exposed. For example, the T_m of skin collagen from codfish that live in cold waters is ~14°C, whereas the T_m of hog intestine collagen and the cuticle collagen of the parasite, *Ascaris*, which inhabits the hog intestinal lumen, is ~40°C. Furthermore, based on a detailed analysis of these and other collagens from different species (Josse and Harrington 1964), the best correlation with T_D is with total pyrrolidine content (Pro + 4-Hyp) and not with either Pro or 4-Hyp content alone. More recently, analyses have been made of the refolding of thermally denatured model collagen-like peptides that have Gly-X-Y triplet repeats with differing residues in the -X- and -Y- positions (Ackerman et al. 1999). The results showed a strong dependence of the folding rate on the identity of the "guest" triplet. For example, all triplets of the form Gly-X-4-Hyp promoted rapid folding of these model peptides, whereas triplets of Gly-Pro-X and Gly-X-Y had much slower folding. In vivo, there are additional mechanisms that direct an assembly of these complex macromolecules (Khoshnoodi et al. 2006). The component α chains have non-collagenous (NC) domains at their N- and C-terminal ends. Type I collagen comprises two α1 chains and one α2 chain in the triple helix; it has now been shown that the α1 C-NC domain contains sufficient information for directing homotrimer assembly, whereas the α2 C-NC domain is required for directing heterotrimer assembly (Khoshnoodi et al. 2006). Disulfide bonding in the α2 C-NC domain is critical for this process. Before the nascent chains fully fold, the prolyl residues in the -Y- position are hydroxylated and the newly formed 4-Hyp residues serve to stabilize the helix (Myllyharju and Kivirikko 2004).

The enzymes that catalyze the formation of collagen 4-Hyp are the collagen prolyl 4-hydroxylases (P4Hs) members of a family of 2-oxyglutarate and ferrous iron-dependent dioxygenases (Myllyharju and Kivirikko 2004). The P4Hs are $\alpha_2\beta_2$ tetramers that localize in the endoplasmic reticulum; the α subunits are the catalytic hydroxylases whereas the chaperone, protein disulfide isomerase (PDI), comprises the β subunits. There are three isoforms of the α subunits in humans (Fig. 1). Spontaneous mutations of collagen prolyl 4-hydroxylases have not yet been reported and a targeted null mutation of one isoform in mice is embryonic lethal (Myllyharju and Kivirikko 2004). The collagen prolyl 4-hydroxylases all require as substrates unfolded collagen chains, molecular oxygen, 2-oxyglutarate, ferrous iron and ascorbate; substrate affinities vary with the enzyme isoform. Other prolyl 4-hydroxylases function in regulating responses to hypoxia by hydroxylating key prolyl residues in the sequence of the hypoxia-inducible transcription factor, HIFα. The amino acid sequence specificity that determines 4-hydroxylation of prolyl residues in HIFα is different from that of the collagen P4Hs (Kaelin 2005). The HIFα hydroxylases, however, do have other substrate requirements that are similar to those of the collagen prolyl 4-hydroxylases, although they do not hydroxylate Pro residues in collagens, have different affinities and do not complex with PDI β subunits. Prolyl hydroxylation of HIFα alters binding to the von Hippel Lindau tumor suppressor protein (pVHL) and

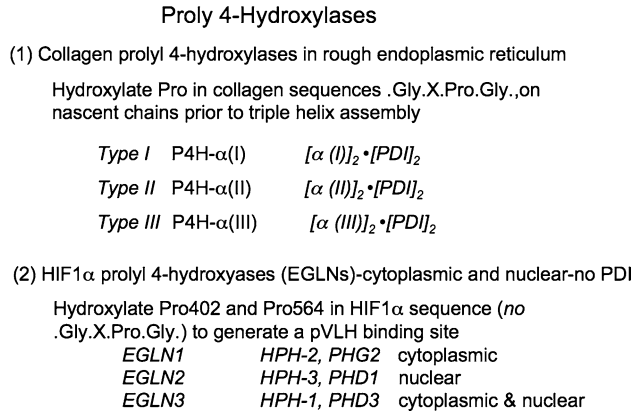

Fig. 1 Prolyl 4-Hydroxylases

thus regulates its activity (Kaelin 2005). These highly specific 4-Pro hydroxylations of regulatory proteins such as HIFα illustrate the important biological roles of the Pro components of other proteins besides structural proteins such as collagens.

Collagen 3-hydroxyproline

The least abundant hydroxylated imino acid in collagens with uninterrupted triple helices is 3-hydroxyproline (3-Hyp, i.e., 3(*S*)-hydroxl-L-proline; Schumacher et al. 2006; Ogle et al. 1962; Irreverre et al. 1962). In type I collagen, there is only 1 residue of 3-Hyp/1000 amino acids, ~1 % that of 4-Hyp, and the single 3-Hyp is found only in the -X- position of a Gly-X-Y triplet with the sequence Gly-3-Hyp-4-Hyp, at residue 986 in the human α1(I) chain. Potential biological functions and metabolism of 3-Hyp were not explored until a prolyl 3-hydroxylase (P3H) was isolated, cloned and characterized in chick embryos and termed P3H1 (Vranka et al. 2004). Other potential members of the P3H family were subsequently identified (Fig. 2). The structure of chick P3H1 indicated that it is the orthologue of a previously described endoplasmic reticulum protein called leprecan (Wassenhove-McCarthy and McCarthy 1999). The *Leprecan* cDNA cloned from the mouse encoded a protein with structural features in common with the other collagen hydroxylases sufficient to consider it another member of the family of 2-oxyglutarate- and ferrous iron-dependent dioxygenases (Vranka et al. 2004). In addition, it was demonstrated by Vranka et al. (2004) that P3H1 specifically binds to denatured collagen in a complex containing at least two other proteins, cyclophilin B (CYPB) and a protein previously isolated as a "cartilage-associated protein" or CRTAP (Tonachini et al. 1999; Morello et al. 1999).

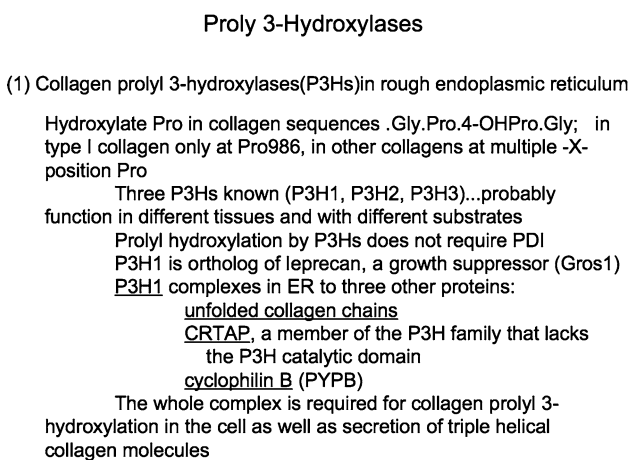

Fig. 2 Prolyl 3-Hydroxylases

Although CRTAP could also be a member of the P3H family, it lacks the catalytic dioxygenase domain and therefore cannot function enzymatically as a collagen prolyl 3-hydroxylase. Interestingly, neither CYPB nor CRTAP are required for full prolyl 3-hydroxylase activity of P3H (Vranka et al. 2004).

Clinical observations have been fundamental in further elucidation of the biology of 3-Hyp. Brittle bone disease or osteogenesis imperfecta (OI) is clinically, biochemically and genetically heterogeneous (Byers 2000; Marini et al. 2007b; Morello et al. 2006). Most forms of OI are dominantly inherited and caused by mutations in the genes that encode type I collagen (*COL1A1* and *COL1A2*). Ward et al. (2002) showed that a clinical variant of OI, termed type VII, however, did not map to *COL1A1* or *COL1A2* and had a recessive rather than a dominant inheritance. Subsequently, elucidation of the role of the *P3H1/CRTAP/CYBP* complex in collagen biology began with the mapping of a kindred with OI type VII to a locus on chromosome 3p22.3 (Ward et al. 2002). Morello et al. (2006) reasoned that mutations in *CRTAP*, included in this region on chromosome 3p22.3, could be a cause of OI type VII in view of its binding to denatured type I collagen and its potential role in proly3-hydroxylation (Vranka et al. 2004). A mutation was then identified in *CRTAP* at a splice site that could potentially lead to unstable *CRTAP* mRNA and decrease *CRTAP* protein (Morello et al. 2006). DNA was also examined from members of an additional family with clinically normal parents and four children, all affected with a severe form of OI considered clinically to be type II OI (Morello et al. 2006). Analysis of *CRTAP* sequences in this group of patients with OI showed a homozygous single base pair deletion that caused a frameshift mutation in exon 4; both asymptomatic parents were carriers of the mutation. Analysis of collagen in medium conditioned by fibroblasts from affected individuals in both kindreds using mass spectroscopy showed Pro, but not 3-Hyp in residue 986, in type I collagen α1(I) chains, despite observations that CRTAP is not a prolyl 3-hydroxylase. In addition, several features of the human mutation including severe osteoporosis and decreased proly3-hydroxylation were found in collagens from *Crtap*-null mice (Morello et al. 2006).

Subsequent observations have helped to clarify several issues. Patients were identified with severe OI previously classified as OI II or III, usually lethal in the first year of life, that also did not map to *COL1A1* or *COL1A2* (Barnes et al. 2006). Several of these patients had mutations in *CRTAP* with absent CRTAP on Western blots and 3-Hyp levels (α1(I) chain residue 986) 0–20% control, consistent with the earlier observations of Morello et al. (2006). Important for this discussion is the later demonstration of null mutations in both alleles of the P3H1/leprecan gene (*P3H1/LEPRE1*) with little or no P3H1 protein made by

fibroblasts and marked decreases in α1(I) chain 3-Hyp content in the remaining seven probands (Cabral et al. 2007). The affected individuals with the *P3H1/LEPRE1* mutations had severe OI with short limbs and markedly distorted bone structure. Both collagen α1(I) and α2(I) chains synthesized by cultured fibroblasts from affected individuals with mutations in *CRTAP* or *P3H1/LEPRE1* had increased lysyl hydroxylation and glycosylation indicating longer retention of nascent chains during elongation in the endoplasmic reticulum. These observations are consistent with a critical role for prolyl 3-hydroxylation in collagen helix formation or stabilization. Based on studies by Jenkins et al. (2003) of model synthetic polyproline-II-like helices, however, 3-Hyp *destabilizes* the triple helical structure, whereas the more abundant 4-Hyp residues in the -Y- positions of the collagen triple helical repeats *stabilize* the triple helical structure (Myllyharju and Kivirikko 2004; Kar et al. 2006). Another study of model collagen peptides concluded that the presence of 3(*S*)Hyp in the -X- position of the Gly-X-Y triplet repeat does not lead to "large structural alterations" in the collagen triple helix (Schumacher et al. 2006).

Vranka et al. (2004) had postulated that a complex of proteins comprising P3H1, CYPB, CRTAP and possibly other larger complexes interact with as yet unfolded procollagen chains in vivo to achieve a fully folded and assembled collagen molecule within the cell. *CYPB*-deficiency, which has not yet been reported, might also result in a phenotype of severe OI with decreased collagen prolyl 3-hydroxylation. Therefore, deficiency in collagen 3-Hyp might simply be a marker for the "dysfunctional" complex. 3-Hyp is more abundant in collagens other than types I, II, III and V, such as type IV collagen, a major component of basement membranes and a determinant of their physiological function. It is possible that the structure of type IV collagens might also be disturbed if the P3H1/CYPB/CRTAP complex were defective. The product of another P3H gene might be required, however. Thus, deficiency in 3-Hyp in the bone collagens could be a marker for the deficiency of P3H1, i.e., the critical biological lesion is the defective chaperone complex that includes P3H1 (the enzyme) as well as CRTAP (the helper protein; Marini et al. 2007a). This possibility could be addressed as suggested previously (Krane 2006) by experiments involving knock-in of a *P3h1* lacking the catalytic dioxygenase sequences, but retaining the capacity to form complexes with Cypb and Crtap.

Collagen prolines and susceptibility to collagenases

Aspects of the structure of collagens determined by the location of prolyl and 4-hydroxyprolyl residues also influence degradation of these extracellular matrix macromolecules by proteolytic enzymes. The so-called collagenases are matrix metalloproteinases (MMPs) (Nagase and Woessner 1999; Nagase et al. 2006; Puentes and López-Otín 2004; Stamencovic 2003; Sternlicht and Werb 2001; Vu and Werb 2000). In humans, five collagenases have been identified and characterized: MMP1, MMP2, MMP8, MMP13 and MMP14. In mice, *Mmp2, Mmp8, Mmp13 and Mmp14* are expressed in various cells and tissues and their function has been elucidated in part by targeted mutagenesis. An ortholog of *Mmp1* in mice has been cloned, but is expressed only at low levels. The first four collagenases are secreted from cells as inactive zymogens (procollagenases), whereas MMP14 is a transmembrane proteinase (also known as MT1-MMP) with the catalytic portion in an extracellular location; MMP14 is activated intracellularly. Each of these activated collagenases in vitro cleaves native, fibril-forming collagens with uninterrupted triple helices at a single corresponding site in the major triple helical domain of type I collagen, for example, between a Gly775 and Ile776 in the α1 chains and a Gly775 and Leu776 in the α2 chain. There are several Gly–Ile and Gly–Leu sequences throughout the triple helical portion of type I collagen and only those at the 755/776 locus are attacked. To explain these observations, it had been proposed three decades ago that the triple helical structure in the region around the collagenase cleavage site is less stable, ascribable to "the arrangement of the amino acids proline and hydroxyproline" (Brown et al. 1977). All of the reported sites in type I collagen where known ligands bind have been mapped and three major ligand-binding regions (MLBRs) have been identified (DiLullo et al. 2002; Marini et al. 2007). One of these, MLBR2, extends from helical position 691–823 and includes sites for binding of $\alpha1\beta1$ and $\alpha2\beta1$ integrins, fibronectin and cartilage oligomeric protein (COMP) as well as the site important for cleavage by MMPs (Lauer-Fields et al. 2000; Lauer-Fields et al. 2001). As can be seen in Fig. 3, the two Gly-X-Y triplets around the cleavage site (775–776) in type I collagen α1 and α2 chains lack Pro residues in the -Y- position.

MMPs can hydrolyze small octapeptide peptides substrates with sequences similar to those around the cleavage

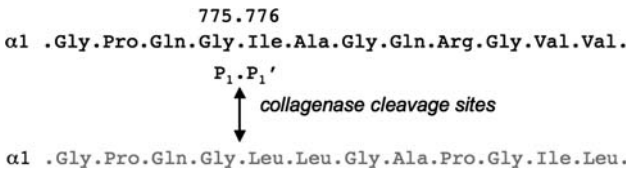

Fig. 3 Amino acid sequences in the region of the colllagenase cleavage sites in the component α chains of type I collagen. Amino acids are numbered from the Gly of the first triplet in the major helix. (Wu et al. 1990)

site in native type I collagen (Fields et al. 1987). Substitution of Pro for Gln at the P_2 site markedly reduces the relative rate of hydrolysis. To understand the cleavage of *native* triple helical collagen, we introduced several mutations around the cleavage site into large, cloned fragments of murine genomic *Col1a1* and transfected the mutant clones into Mov13 fibroblasts that do not express endogenous *Col1a1* due to a retroviral insertion (Wu et al. 1990). Mov13 fibroblasts, however, which continue to express *Col1a2* and translate *Col1a2* mRNA, nevertheless secrete no type I collagen since α2(I) chains cannot form homotrimers (Khoshnoodi et al. 2006). Following transfection of Mov13 fibroblasts with mutated *Col1a1*, collagens are secreted that comprise one endogenous α2(I) chain and two mutated α1(I) chains. The secreted collagens can then be tested for susceptibility to cleavage by collagenases (Wu et al. 1990). As shown schematically in Fig. 4, type I collagen containing an Ile776Pro substitution in the α1(I) chains at P_1' was not cleaved by MMP1. Furthermore, type I collagen containing a double substitution of Gln774Pro (P_2) and Ala777Pro (P_2') leaving P_1' intact was also not cleaved providing direct evidence for the possible role of the less stable helical structure in this region of native collagen in determining where and whether the collagen is susceptible to proteolysis. We learned from these studies in cell culture that the presence of two mutant uncleavable α1(I) chains in the trimer prevents the cleavage of the wt α2(I) chain and that a misaligned wt α2(I) chain prevents the collagenase cleavage of two cleavable α1(I) chains.

We then introduced the mutation encoding the double substitution of Gln774Pro (P_2) and Ala777Pro (P_2') (Fig. 4, Mutant IV and Fig. 5) into the endogenous Col1a1 gene by gene targeting that permitted the mice to develop and survive to >1 year of age. The homozygous collagenase-resistant mice (termed Col1a1$^{tm1\ Jae}$ or r/r mice) did develop fibrosis of the dermis and some impairment of postpartum involution of the uterus (Liu et al. 1995). As predicted from the earlier in vitro studies (Wu et al. 1990), collagen extracted from skin and bone of r/r mice was completely resistant to proteolysis at the helical site by all the mammalian collagenases tested including MMP1, MMP2, MMP8, MMP13 and MMP14 and by a collagenase from *Xenopus laevis* gene 11 termed *XCL3* (Jung et al. 2004). We also extracted collagen from the skin of heterozygous (r/+) mice that enabled us to show that a single mutant (r) α1(I) chain in the heterotrimer prevents collagenase cleavage of the single wt α1(I) chain and the wt α2(I) (Chiusaroli et al. 2003). The collagenase-resistance mutation thus behaves as dominant negative. Important insights into the mechanism of MMP cleavage of helical collagen have been obtained by the demonstration that collagenases, through their C-terminal so-called pexin domains, first bind and locally unwind the triple helical structure before hydrolyzing the peptide bonds (Chung et al. 2004). For these experiments, a catalytically inactive mutant of MMP1 (Glu200Ala) was utilized in combination with other noncollagenolytic active MMPs that by themselves do not cleave native collagen. Thus, one could predict that the collagen from the r/+ or r/r mice with the substitutions of Gln774Pro and Ala777Pro in one or two α1(I) chains would be resistant to this unwinding and therefore resistant to proteolysis. Dr. H. P. Bächinger of the Portland OR Shriners Hospital kindly measured the T_m of collagen that we extracted from the skin or tail of r/r and wt mice using circular dichroism; there was no difference in T_m of collagens from r/r and wt mice. It could thus not be shown that addition of the two Pro residues in the Y position in two triplets of two α1(I) chains altered the

MMP1Cleavage

	775.776	
WT	.Gly.Pro.Gln.Gly.Ile.Ala.Gly.Gln.	+
MutI	.Gly.Pro.Gln.Gly.**Pro**.Ala.Gly.Gln.	0
MutII	.Gly.Pro.**Pro**.Gly.Ile.**Pro**.Gly.Gln.	0
MutIV	.Gly.Pro.**Pro**.Gly.**Met**.**Pro**.Gly.Gln.	0

Fig. 4 Mutations encoded in Col1a1 genomic clones expressed in Mov13 fibroblasts. Mov13 fibroblasts were transfected with genomic Col1a1 clones with mutations as indicated. Type I collagen in medium proteins containing wt or mutant α1(I) chains was prepared by limited digestion of medium proteins with pepsin after labeling cells with [3H] proline. To test susceptibility of the labeled collagens to collagenase digestion, the labeled collagens were incubated at 20°C for 18–24 h with purified activated rheumatoid synovial collagenase (MMP1) and products resolved by SDS-PAGE and analyzed by fluorography (Wu et al. 1990)

wild type

	775.776
α1(I)	α1 .Gly.Pro.Gln.Gly.Ile.Ala.Gly.Gln.
	α1 .Gly.Pro.Gln.Gly.Ile.Ala.Gly.Gln.
α2(I)	α2 .Gly.Pro.Gln.Gly.Leu.Leu.Gly.Ala.

r mutation

	775.776
α1(I)	α1 .Gly.Pro.**Pro**.Gly.**Met**.**Pro**.Gly.Gln.
	α1 .Gly.Pro.**Pro**.Gly.**Met**.Pro.Gly.Gln.
α2(I)	α2 .Gly.Pro.Gln.Gly.Leu.Leu.Gly.Ala.

Fig. 5 Sequences around the collagenase cleavage site in wild type and collagenase-resistant type I collagen based on the nucleotide sequence of the mutated genomic Co1ia1 clone. The Pro residues at 774 and 777 would be at least partially hydroxylated to 4-Hyp in vivo

overall helical structure of the mutated collagen molecules, even though the Pro (4-Hyp) substitutions could still alter the local helical stability (not measurable by this analysis) in the region of the collagenase cleavage site.

Further studies of the r/r mice revealed other features of the abnormal phenotype, particularly in the skeleton. For example, there was evidence of decreased generation of functional osteoclasts (bone-resorbing cells) with decreased bone resorption (Zhao et al. 1999). In addition, there were striking differences noted in the number of empty osteocyte lacunae in calvariae and long bones from r/r compared to wt mice as young as 4 weeks of age, and increased further as the animals aged (Zhao et al. 2000). In this study, the number of persisting *TUNEL*-positive (apoptotic) osteocytes as well as periosteal cells in the calvariae from r/r mice also increased. Osteocytes from wt mice normally release collagenase that act on the surrounding extracellular matrix, evidenced by staining with Mab 9A4, which recognizes a cryptic epitope in the larger N-terminal collagenase-cleavage fragment of type I collagen (Zhao et al. 2000). Staining with the Mab was not detected in the periphery of osteocytes in calvariae from r/r mice. Thus, collagenase normally cleaves type I collagen at the helical locus in vivo, but collagenase does not cleave r/r collagen at this locus in vivo.

Paradoxically, increases in calvarial thickness were first noted in r/r mice at 4–6 weeks of age and by 12 months of age, were ~twice that in +/+ mice. The pattern of in vivo labeling with calcein, a fluorescent marker that localizes to newly mineralized bone surfaces, was consistent with increased bone formation (Zhao et al. 2000). In the long bones from wt mice, endosteal surfaces were smooth with little deposition of new endosteal bone, but, in r/r mice, endosteal surfaces contained abundant new, endocortical, trabecular bone surrounding active marrow spaces. In r/r mice, calcein labeling was observed at both periosteal and endosteal surfaces with most of the fluorescence in the new bone at endosteal surfaces; in wt mice only limited and scattered labeling was found on endosteal surfaces. In view of these observations, we postulated that normally osteocytes (as well as osteoblasts and their precursor cells) utilize signals generated by collagen cleavage to maintain their viability and, if such signals are not induced, they undergo apoptosis and their lacunae empty. The increased bone deposition in untreated r/r mice is possibly accounted for by loss of normal negative signals from osteocytes (Zhao et al. 2000). A good candidate for such a negative signal derived from osteocytes is sclerostin also known as SOST that has since been identified (van Bezooijen et al. 2004). Subsequently, in studies using Bodian staining, it was evident that the canalicular network through which osteocytes communicate with each other and with bone forming cells at bone surfaces are strikingly disrupted in the r/r mice and mimic findings in *Mmp2*-null mice (Inoue et al. 2006).

Abnormal soft tissue remodeling was also found in the collagenase-resistant mice. Wound healing is impaired (Beare et al. 2003). In other studies, more collagen accumulates in experimentally induced atherosclerotic plaques in r/r mice (Fukumoto et al. 2004) and recovery from CCl4-induced liver fibrosis is delayed in r/r mice with persistent fibrosis and altered pattern of the component cells (Issa et al. 2003). Whether such pathology is due to absence of collagen resorption and/or an altered pattern of collagen deposition has not been established. In several of these examples, altered cellular proliferation and migration were evident in the r/r mice suggesting that the failure to resorb collagen may account for only a part of the phenotype. Indeed, some cells may use collagenases not only to resorb collagen in the surrounding extracellular matrix but also to nick a chain to reveal a cryptic binding site essential for some signaling event such as a cellular survival signal. In this regard, it has been demonstrated that human keratinocytes (HCat cells) can migrate on a normal type I collagen matrix in a process that involves cell binding to the matrix and induction of expression of *MMP1* (Pilcher et al. 1997). This migration of HCat cells is abrogated by synthetic MMP inhibitors and does not take place on a substrate of r/r collagen.

Conclusions and future perspectives

The structure of collagens requires a glycine residue at every third residue (Gly-X-Y triplets) of the polypeptides that comprise the triple helix and substitutions for glycine disrupt the structure and cause disease such as osteogenesis imperfecta. Prolines and 4-hydroxyprolines, abundant amino acids in collagens, play roles as stabilizers of this collagen helical structure. The collagen prolyl 4-hydroxylases act on prolines in growing not yet folded nascent chains in the rough endoplasmic reticulum first anchored through their C terminal non-collagenous domains to form trimers. The major prolyl 3-hydroxylase turns out to be a critical component of a complex of molecular chaperones, that function in moving the newly synthesized collagen through the rough endoplasmic reticulum. The position of prolines and 4-hydroxyprolines in the collagen triplets also determines the susceptibility of collagens to cleavage at a unique locus by collagenases. Targeted mutagenesis in mice that alters the position of critical proline and 4-hydroxyproline residues results in interesting phenotypes. It is essential to elucidate the molecular mechanisms that underlie these phenotypes.

Acknowledgments Supported by NIH grant AR044815.

References

Ackerman MS, Bhate M, Shenoy N, Beck K, Ramshaw JA, Brodsky B. (1999) Sequence dependence of the folding of collagen-like peptides. Single amino acids affect the rate of triple-helix nucleation. J Biol Chem 274:7668–7673

Barnes AM, Chang W, Morello R, Cabral WA, Weis M, Eyre DR, Leikin S, Makareeva E, Kuznetsova N, Uveges TE, Ashok A, Flor AW, Mulvihill JJ, Wilson PL, Sundaram UT, Lee B, Marini JC (2006) Deficiency of cartilage-associated protein in recessive lethal osteogenesis imperfecta. N Engl J Med 355:2757–2764

Beare AHM, O'Kane S, Krane SM, Ferguson MWJ (2003) Severely impaired wound healing in the collagenase resistant mouse. J Invest Dermatol 120:153–163

Brown RA, Hukins DW, Weiss JB, Twose TM (1977) Do mammalian collagenase and DNA restriction endonucleases share a single mechanism for cleavage site recognition? Biochem Biophys Res Commun 74:1102–1108

Byers PH (2000) Osteogenesis imperfecta: perpectives and opportunities. Curr Opin Perdiatr 12:603–609

Cabral WA, Chang W, Barnes AM, Weis M, Scott MA, Leikin S, Makareeva E, Kuznetsova NV, Rosenbaum KN, Tifft CJ, Bulas DI, Kozma C, Smith PA, Eyre DR, Marini JC (2007) Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. Nat Genet 39:359–365

Chiusaroli R, Maier A, Knight MC, Byrne M, Calvi LM, Baron R, Krane SM, Schipani E (2003) Collagenase cleavage of type I collagen is essential for both basal and parathyroid hormone (PTH)/PTH-related peptide receptor-induced osteoclast activation and has differential effects on discrete bone compartments. Endocrinology 144:4106–4116

Chung L, Dinakarpandian D, Yoshida N, Lauer-Fields JL, Fields GB, Visse R, Nagase H (2004) Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. EMBO J 23:3020–3030

DiLullo GA, Sweeney SM, Körkkö J, Ala-Kokko L, SanAntonio JD (2002) Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. J Biol Chem 277:4223–4231

Fields GB, Van Wart HE, Birkedal-Hansen H (1987) Sequence specificity of human skin fibroblast collagenase. Evidence for the role of collagen structure in determining the collagenase cleavage site. J Biol Chem 262:6221–6226

Fischer E (1902) Über eine neue Aminosäure aus Leim. Chem Berichte 35:2660–2665

Fukumoto Y, Deguchi JO, Libby P, Rabkin-Aikawa E, Sakata Y, Chin MT, Hill CC, Lawler PR, Varo N, Schoen FJ, Krane SM, Aikawa M (2004) Genetically determined resistance to collagenase action augments interstitial collagen accumulation in atherosclerotic plaques. Circulation 110:1953–1959

Inoue K, Mikuni-Takagaki Y, Oikawa K, Itoh T, Inada M, Noguchi T, Park J-S, OnoderaT, Krane SM, Noda M, Itohara S (2006) A crucial role for Mmp-2 in osteocytic canalicular formation and bone metabolism. J Biol Chem 281:33813–33824

Irreverre F, Morita K, Robertson AV, Witkop B (1962) Isolation and synthesis of 3-hydroxy-L-proline. Biochem Biophys Res Commun 8:453–455

Issa R, Zhou X, Trim N, Millward-Sadler H, Krane S, Benyon C, Iredale J (2003) Mutation in collagen-1 that confers resistance to the action of collagenase results in failure of recovery from CCl4-induced liver fibrosis, persistence of activated hepatic stellate cells, and diminished hepatocyte regeneration. FASEB J 17:47–49

Jenkins CL, Bretscher LE, Guzei IA, Raines RT (2003) Effect of 3-hydroxyproline residues on collagen stability. J Am Chem Soc 125:6422–6427

Josse J, Harrington W (1964) Role of pyrrolidine residues in the structure and stabilization of collagen. J Mol Biol 9:269–287

Jung J-C, West-Mays JA, Stramer BM, Byrne MH, Scott S, Mody MK, Sadow PM, Krane SM, Fini ME (2004) Studies on activity and expression of *xenopus laevis* matrix metalloproteinases: identification of a novel role for the hormone prolactin in regulating collagenolysis in both amphibians and mammals. J Cell Physiol 201:155–164

Kaelin WG Jr (2005) Proline hydroxylation and gene expression. Annu Rev Biochem 74:115–128

Kar K, Amin P, Bryan MA, Persikov AV, Mohs A, Wang Y-H, Brodsky B (2006) Self-association of collagen triple-helical peptides into higher order structures. J Biol Chem 80:33282–33290

Khoshnoodi J, Cartailler J-P, Alvares K, Veis A, Hudson BG (2006) Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers. J Biol Chem 281:38117–38121

Krane SM (2006) Mutations in genes encoding components of a post-translational-modifying protein cause another collagen disease. BoneKEy-Osteovision 3(11):10–13

Lauer-Fields JL, Tuzinski KA, Shimokawa K, Nagase H, Fields GB (2000) Hydrolysis of triple-helical collagen peptide models by matrix metaloproteinases. J Biol Chem 275:13282–13290

Lauer-Fields JL, Broder T, Sritharan T, Chung L, Nagase H, Fields GB (2001) Kinetic analysis of matrix metalloproteinase activity using fluorogenic triple-helical substrates. Biochemistry 40:5795–5803

Liu X, Wu H, Byrne M, Jeffrey J, Krane S, Jaenisch R (1995) A targeted mutation at the known collagenase cleavage site in mouse type I collagen impairs tissue remodeling. J Cell Biol 130:227–237

Marini JC, Cabral WA, Barnes AM, Chang W (2007a) Components of the collagen prolyl 3-hydroxylation complex are crucial for normal bone development. Cell Cycle 6:1675–1681

Marini JC, Forlino A, Cabral WA, Barnes AM, San Antonio JD, Milgrom S, Hyland JC, Korkko J, Prockop DJ, De Paepe A, Coucke P, Symoens S, Glorieux FH, Roughley PJ, Lund AM, Kuurila-Svahn K, Hartikka H, Cohn DH, Krakow D, Mottes M, Schwarze U, Chen D, Yang K, Kuslich C, Troendle J, Dalgleish R, Byers PH (2007b) Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. Hum Mutat 28:209–221

Morello R, Tonachini L, Monticone M, Viggiano L, Rocchi M, Cancedda R, Castagnola P (1999) cDNA cloning, characterization and chromosome mapping of Crtap encoding the mouse cartilage associated protein. Matrix Biol 18:319–324

Morello R, Bertin TK, Chen Y, Hicks J, Tonachini L, Monticone M, Castagnola P, Rauch F, Glorieux FH, Vranka J, Bächinger HP, Pace JM, Schwarze U, Byers PH, Weis M, Fernandes RJ, Eyre DR, Yao Z, Boyce BF, Lee B (2006) CRTAP is required for prolyl 3- hydroxylation and mutations cause recessive osteogenesis imperfecta. Cell 127:291–304

Myllyharju J, Kivirikko KI (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet 20:33–43

Nagase H, Woessner JF Jr (1999) Matrix metalloproteinases. J Biol Chem 274:21491–21494

Nagase H, Visse R, Murphy G (2006) Structure and function of matrix metalloproteinases and TIMPs. Cardiovascular Res 69:562–573

Ogle JD, Arlinghaus RB, Logan MA (1962) 3-Hydroxyproline, a new amino acid of collagen. J Biol Chem 237:3667–3673

Piez KA (1984) Molecular and aggregate structures of the collagens. In: Piez KA, Reddi AH (eds) Extracellular matrix biochemistry. Elsevier, New York, pp 1–39

Pilcher BK, Dumin JA, Sudbeck BD, Krane SM, Welgus HG, Parks WC (1997) The activity of collagenase-1 is required for keratinocyte migration on a type I collagen matrix. J Cell Biol 137:1445–1457

Puentes XS, López-Otín C (2004) A genomic analysis of rat proteases and protease inhibitors. Genome Res 14:609–622

Schumacher MA, Mizuno K, Bächinger HP (2006) The crystal structure of a collagen-like polypeptide with 3(S)-hydroxyproline residues in the Xaa position forms a standard 7/2 collagen triple helix. J Biol Chem 281:27566–27574

Stamenkovic I (2003) Extracellular matrix remodelling: The role of matrix metalloproteinases. J Pathol 200:448–464

Sternlicht MD, Werb Z (2001) How matrix metalloproteinases regulate cell behavior. Annu Rev Cell Dev Biol 17:463–516

Stetten MR, Schoenheimer R (1944) The metabolism of *l*(-) proline studied with the aid of deuterium and isotopic nitrogen. J Biol Chem 153:113–132

Tonachini L, Morello R, Monticone M, Skaug J, Scherer SW, Cancedda R, Castagnola P (1999) cDNA cloning, characterization and chromosome mapping of the gene encoding human cartilage associated protein (CRTAP) Cytogenet Cell Genet 87:191–194

van Bezooijen RL, Roelen B A, Visser A, van der Wee-Pals L, de Wilt E, Karperien M, Hamersma H, Papapoulos SE, ten Dijke P, Lowik CW (2004) Sclerostin is an osteocyte-expressed negative regulator of bone formation, but not a classical BMP antagonist. J Exp Med 199:805–814

Vranka JA, Sakai LY, Bächinger HP (2004) Prolyl 3-hydroxylase 1, enzyme characterization and identification of a novel family of enzymes. J Biol Chem 279:23615–23621

Vu TH, Werb Z (2000) Matrix metalloproteinases:effectors in development and normal physiology. Genes Dev 14:2123–2133

Ward LM, Rauch F, Travers R, Chabot G, Azouz EM, Lalic L, Roughley PJ, Glorieux FH (2002) Osteogenesis imperfecta type VII: an autosomal recessive form of brittle bone disease. Bone 31:12–18

Wassenhove-McCarthy DJ, McCarthy KJ (1999) Molecular characterization of a novel basement membrane associated proteoglycan, leprecan. J Biol Chem 274:25004–25017

Wu H, Byrne MH, Stacey A, Goldring MB, Birkhead JR, Jaenisch R, Krane SM (1990) Generation of collagenase-resistant collagen by site-directed mutagenesis of murine proα1(I) collagen gene. Proc Natl Acad Sci USA 87:5888–5892

Zhao W, Byrne MH, Boyce BF, Krane SM (1999) Bone resorption induced by parathyroid hormone is strikingly diminished in collagenase-resistant mutant mice. J Clin Invest 103:517–524

Zhao W, Byrne MH, Wang Y, Krane SM (2000) Inability of collagenase to cleave type I collagen in vivo is associated with osteocyte and osteoblast apoptosis and excessive bone deposition. J Clin Invest 106:941–949