# Genotype-phenotype relations in patient-derived point mutations in collagen II

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

**Chakkalakal Anandan Salin**

aus Trichur, India

im September 2005

| | |
|---|---|
| Berichterstatter: | Herr Prof. Dr. Mats Paulsson |
| | Herr Prof. Dr. Helmut W. Klein |
| Tag der mündlichen Prüfung: | 2. Dezember 2005 |

*TO MY PARENTS*

# TABLE OF CONTENTS

# Abstract

Collagen II is the major collagen present in the extracellular matrix of cartilage, in addition it is found in the vitreous of the eye and it is also detected during early embryogenesis. Due to the complexity of the biosynthesis, assembly and secretion, collagen II is highly susceptible to mutations leading to disease states which are broadly classified as chondrodysplasias. Most of these mutations are substitutions of glycine in the Gly-X-Y repeats in the triple helical domain resulting in destabilization of the helix. Point mutations leading to arginine to cysteine substitution are interesting since they occur at either X or Y position and cause two different diseases termed Stickler syndrome and congenital spondyloepiphyseal dysplasia (SEDC) in association with osteoarthritis.

The present study aimed at determining the consequences of arginine to cysteine substitutions in either X or the Y position in the Gly-X-Y repeats and at either the N- or the C- terminus of the triple helix. The impact of these mutations on protein trafficking, secretion and cell survival was also analyzed.

Biochemical studies revealed great similarities between R75C, R134C and R704C collagens and the wild type molecules with the exception that electron micrographs of R75C collagen displayed kinks in the structure. R740C and R789C collagens accumulated in the cells and the R789C protein migrated faster on SDS gels. The R740C and R789C proteins were also susceptible to protease digestion and circular dichroism spectra were altered and showed lower Tm values than other collagen II variants. Due to the altered structure, R789C protein was more susceptible to MMP cleavage in the vicinity of the mutation causing the truncation of the protein. Additionally, electron micrographs revealed only scarce and thin filamentous structures in preparations of R740C and R789C protein. The biochemical results indicate that the R740C and R789C proteins have unstable triple helices and that this affects the overall protein structure.

Protein trafficking was monitored in HT1080 cells expressing the collagen II variants. Intracellular retention in the ER due to misfolding of the R740C and R789C proteins triggered an ER stress response including splicing of XBP-1 and induction and binding of BiP. Continuous accumulation of misfolded proteins in the ER caused apoptosis of the R740C and R789C expressing cells.

Substitution of arginine to cysteine in the X or Y position towards the C-terminus of the triple helix caused pronounced instability of the triple helix with a deleterious effect on the cells, while R704C and more N-terminal mutations did not cause any significant changes irrespective of being in the X or the Y position.

The different severities of patient phenotypes are due to a combination of structural factors, which may be synergistically augmented by genetic modifiers, additional unknown mutations and environmental factors.

# Zusammenfassung

Collagen II ist der Hauptbestandteil der extrazellulären Matrix des Knorpels. Außerdem wird Collagen II im Glaskörper des Auges exprimiert und konnte während der frühen embryonalen Entwicklung nachgewiesen werden. Collagen II ist durch seine komplexe Biosynthese, Assemblierung und Sekretion besonders anfällig für Mutationen, die im Menschen zu Krankheiten führen, die allgemein unter dem Begriff Chondrodysplasien zusammengefasst werden. Bei den meisten dieser Mutationen handelt es sich um den Austausch eines Glycins in den Gly-X-Y Triplets der tripelhelikalen Domäne, die zu einer Destabilisierung dieser Helix führen. Punktmutationen, die zu einem Austausch von Arginin zu Cystein führen, kommen sowohl in der X- als auch Y-Position vor und führen interessanterweise zu zwei unterschiedlichen Krankheitsbildern, dem Stickler Syndrom bzw. der congenitalen Spondyloepiphysiären Dysplasie verbunden mit Osteoarthrose.
In der vorliegenden Arbeit wurden Auswirkungen von Substitutionen eines Arginins zu einem Cystein sowohl in der X- als auch in der Y-Position bzw. im N- oder C-terminalen Bereich der Tripelhelix untersucht. Außerdem wurden Effekte dieser Mutationen auf den intrazellulären Proteintransport, die Sekretion und die Zellvitalität analysiert.
Biochemische Untersuchungen ergaben keinerlei wesentliche Unterschiede zwischen den Mutanten R75C, R134C und R704C im Vergleich mit dem Wildtyp Protein mit Ausnahme der elektronenmikroskopischen Aufnahmen, in denen die Mutation R75C zu charakteristischen Knicken in der Struktur des Collagens führte. Die Proteine R740C und R789C akkumulierten intrazellulär, und das Collagen R789C zeigte veränderte Laufeigenschaften im SDS Gel. Die Proteine R740C und R789C waren außerdem empfindlicher gegenüber Proteasen, und die Analyse mittels CD Spektroskopie ergab untypische Spektren und erniedrigte Schmelztemperaturen. Die durch die veränderte Struktur erhöhte Empfindlichkeit des Proteins R789C gegenüber dem Abbau durch MMPs in unmittelbarer Nähe der Mutation führte zu einer Verkürzung des Proteins durch Proteolyse.
In elektronenmikroskopischen Untersuchungen der Proteine R740C und R789C konnten nur kurze filamentöse Strukturen nachgewiesen werden. Diese Ergebnisse führen zu dem Schluss, dass die Proteine R740C und R789C lediglich instabile Tripelhelices ausbilden und die Mutationen so die Gesamtstruktur des Proteins beeinflussen.
Der intrazelluläre Transport der mutierten Proteine wurde in transfizierten HT1080 Zellen untersucht. Die fehlgefalteten Proteine R740C und R789C wurden im endoplasmatischen Retikulum (ER) zurückgehalten und lösten eine ER Stressantwort aus, die mit einem Splicing von XBP-1 und einer erhöhten Induktion und Bindung des Chaperons BiP einhergeht. Die andauernde Akkumulation dieser fehlgefalteten Proteine im ER führte schließlich zur Auslösung der Apopotse.
Der Austausch von Arginin zu Cystein im C-terminalen Bereich der Tripelhelix führte sowohl in der X- als auch der Y-Position zur Instabilität der Tripelhelix mit schädlichen Auswirkungen auf die exprimierenden Zellen. Im Gegensatz hierzu resultierten bereits die Mutation R704C und weiter N-terminal gelegene Mutationen in keinerlei signifikanten Veränderungen, unabhängig davon, ob sie in der X- bzw. Y-Position vorlagen.
Die in Patienten beschriebenen, sehr unterschiedlich ausgeprägten Krankheitsbilder sind wahrscheinlich durch eine Kombination sich synergistisch verstärkender struktureller und genetischer Faktoren sowie unbekannter Mutationen und Umwelteinflüsse zu erklären.

# Abbreviations

| | |
|---|---|
| AA | Acrylamide |
| AP | Alkaline phosphatase |
| APS | Ammonium persulphate |
| BCA | Bicinchoninic acid |
| BiP | Immunoglobulin heavy chain binding protein |
| BSA | Bovine serum albumin |
| BM-40 | Basement membrane protein, molecular weight 40 kD |
| CD | Circular Dichroism |
| COMP | Cartilage oligomeric matrix protein |
| DEPC | Diethylpyrocarbonate |
| DMEM | Dulbecco's minimum essential medium |
| DMSO | Dimethylsulfoxide |
| dNTP | Deoxyribonucleotide triphosphate |
| DTT | 1,4-dithiothreitol |
| EBNA | Ebstein barr virus nuclear antigen |
| EDTA | Ethylenediaminetetraacetic acid |
| EK | Enterokinase |
| ER | Endoplasmic reticulum |
| FBS | Fetal bovine serum |
| FITC | Fluorescein-5-isothiocyanate |
| GFP | Green fluorescence protein |
| His | Histidine |
| HRP | Horse radish peroxidase |
| IgG | Immunoglobulin G |
| kb | Kilobase pairs |
| kD | Kilodalton |
| MALDI-TOF-MS | Matrix assisted laser desorption time of flight |
| MED | Multiple epiphyseal dysplasia |
| MMP | Matrix metalloproteinase |
| MOI | Moieties of infection |
| mRNA | Messenger ribonucleic acid |
| NEM | N-Ethylmaleimide |

| | |
|---|---|
| NP-40 | Nonylphenylpolyethyleneglycol |
| NTA | Nitriloacetic acid |
| OD | Optical density |
| OMIM | Online mendelian inheritance in man |
| PDI | Protein disulfide isomerase |
| PMSF | Phenylmethylsulphonylfluoride |
| PAGE | Polyacrylamide gel electrophoresis |
| SC | Pepsin extracted collagen from bovine nasal cartilage (Sigma) |
| SCGE | Single cell gel electrophoresis |
| SDS | Sodium dodecyl sulphate |
| SEDC | Spondyloepiphyseal dysplasi congenita |
| TAE | Tris acetate EDTA |
| XBP | X-box DNA binding protein |

# 1 Introduction

## 1.1 Collagen

Collagens form a family of proteins that constitute the major structural components of the ECM, representing approximately one third of the body protein in man. The general definition of a collagen is "a protein consisting of three identical or related polypeptide chains, which are folded into at least one triple-helical domain and assemble into supramolecular aggregates mostly in the extracellular space". They are found in almost all tissues of the body, and are particular abundant in bone, skin, tendon, ligaments, cartilage and vessel walls (Myllyharju & Kivirikko 2004).

## 1.2 Classification of collagens

Collagens are characterized by the presence of triple-helical domains, which are either composed of three identical polypeptide chains (homotrimers) or two to three genetically distinct chains (heterotrimers). The collagen types are designated with Roman numerals in the order of their discovery, and the chains found in each collagen type are identified with Arabic numerals (Kivirikko 1993, Pihlajaniemi & Rehn 1995, Prockop & Kivirikko 1995, Myllyharju & Kivirikko 2004). Collagens are broadly classified into two groups on the basis of their primary structure, physicochemical properties and macromolecular assembly:

1. Fibrillar collagens (over 70% of total collagens; e.g. type I, II, III, V, XI), possessing a large, uninterrupted triple-helical domain capable of fibril formation. Their fibrils create the structural frameworks for many tissues.
2. The non-fibril-forming collagens, made up by basement membrane collagens (type IV and VII), transmembrane collagens (type XIII and XVII), fibril-associated collagens with interrupted triple helices (FACIT; type IX, XII, XIV, XVI and XIX), multiplexin collagens (type XV and XVIII), short-chain collagens (type VI, VIII and X) form a more heterogeneous group. A distinct feature of non-fibril-forming collagens is the presence of one or more non collagenous interruptions within the collagenous sequence. Members of this group form filaments, sheet-like structures, network-like elements and anchoring fibrils and can associate with collagen fibrils and membranes in a highly tissue-specific manner in variety of organs.

Fig 1.1 summarizes the classification within the collagen superfamily, the known collagen supramolecular assemblies and molecular and functional properties.

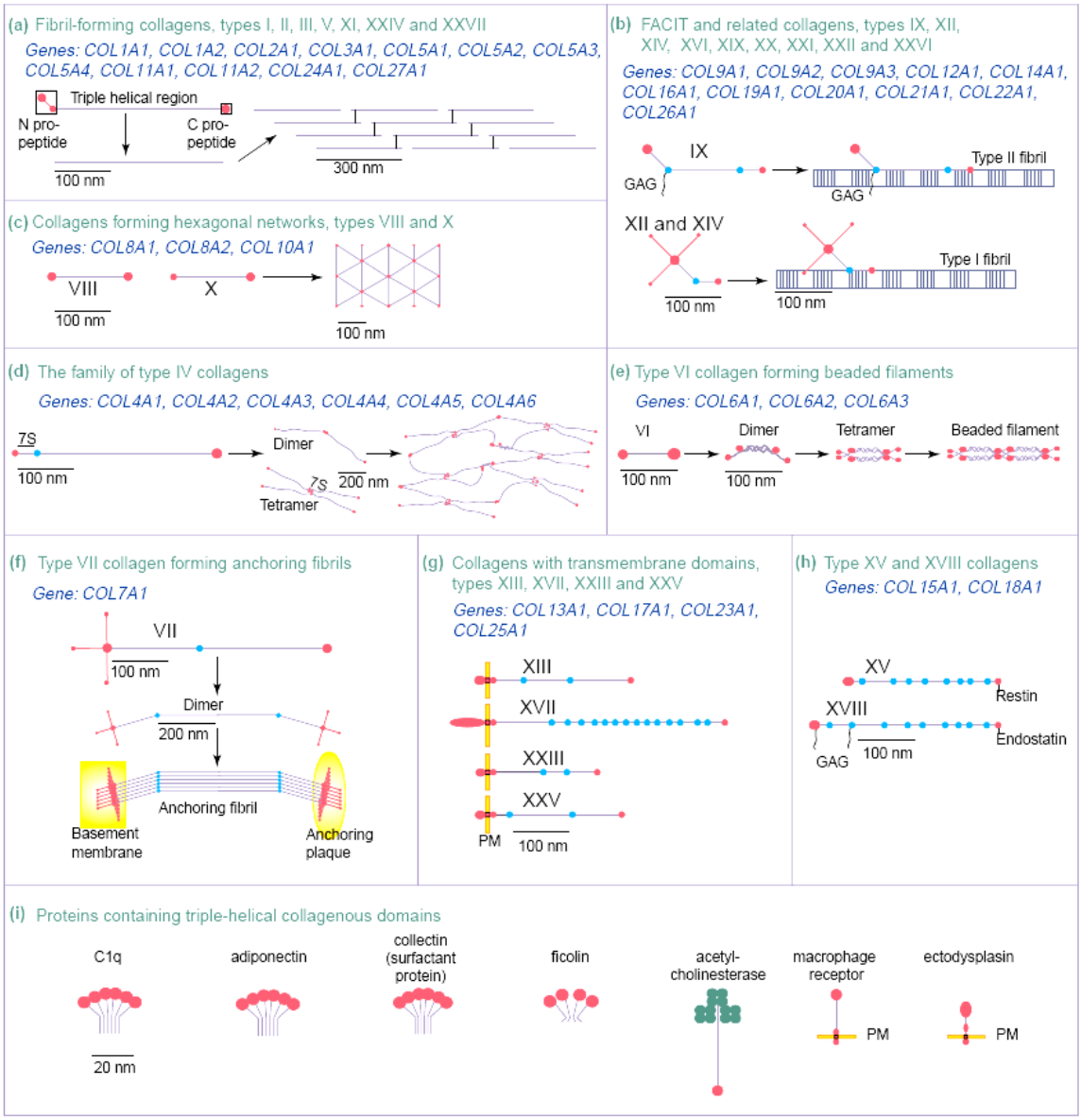


**Fig 1.1 Members of the collagen superfamily and their known supramolecular assemblies.** The collagen superfamily can be divided into nine families on the basis of the supramolecular assemblies and other features of its members: (a) fibril-forming collagens; (b) fibril-associated collagens with interrupted triple helices (FACITs), located at the surface of fibrils, and structurally related collagens; (c) collagens forming hexagonal networks; (d) the family of type IV collagens located in basement membranes; (e) type VI collagen, which forms beaded filaments; (f) type VII collagen, which forms anchoring fibrils at basement membranes; (g) collagens with transmembrane domains; and (h) the family of type XV and XVIII collagens. The supramolecular assemblies of families (g) and (h) are unknown and are therefore not shown (Myllyharju & Kivirikko 2004).

## 1.3 Genes and nomenclature

The genes encoding the fibril-forming collagens I-III show structural similarity but vary in size from 18 to 44 kb (Prockop & Kivirikko 1995, Yamada et al. 1980). They consist of 51-54 exons with the major triple-helical domain being encoded by 44 exons. The similarity in gene structure is also extended to propeptides. The N-propeptide is in each case encoded by six exons, whereas

the C-propeptides are encoded by four exons. The gene loci for the members of the collagen family have been given names beginning with COL in a human context and col in animals, followed by an Arabic numeral denoting the collagen. For example, COL1A1 is the gene locus for the human proα1(I) chain, and COL1A2 for the proα2(I) chain.

All the exons encoding the triple-helical domain are of sizes that are multiples of 9 bp (e.g. 45, 54, 99, 108 and 162 bp, coding for 5, 6, 11, 12 and 18 Gly-X-Y triplets, respectively), the most common size being 54 bp. It has been suggested that the ancestral gene for the fibril-forming collagens must have evolved by amplification of a 54 bp unit embedded into intron sequences (Yamada et al. 1980). All these exons start with a codon for glycine and end with a complete codon for an amino acid in the Y position.

## 1.4 Structure and stability of the collagen triple helix

In collagens three α chains are wound around each other into a right-handed triple helix. In some collagens all three chains in the molecule are identical and form homotrimers, while other types contain two or even three different chains leading to heterotrimeric molecules (Engel & Prockop 1991, Hulmes 1992). At the amino acid level the collagenous or triple helical domains contain of repeats consisting of three amino acids (Fig 1.2A). The presence of glycine as every third amino acid in the repeating Gly-X-Y sequence is essential because a larger amino acid would not fit in the center of the triple helix where the three chains come together (Fig 1.2B). Proline is frequently in the X position of the Gly-X-Y sequence and 4-hydroxyproline is frequently in the Y position. These two amino acids limit the rotation of the polypeptide chains. The triple helix is further stabilized by hydrogen bonds and water bridges, many of which require the presence of 4-hydroxyproline. Stabilization of the triple helix by glycines has been studied using collagen-like peptides (Bella et al. 1994, 1995). In addition, many glycine substitutions in the collagen triple helix have been identified in various heritable connective tissue disorders (Kivirikko 1993, Dalgleish 1997, 1998). These glycine substitutions result in a small local untwisting of the triple helix and reduce its thermal stability (Bella et al. 1994), suggesting that similar conformational changes may occur in the case of glycine substitutions at other positions. It is also thought that glycine substitutions may reduce the rate of folding of the triple helix, as a Gly to Cys substitution within the collagen domain of type I collagen prolongs the time needed to reach the triple-helical state (Raghunath et al. 1994).

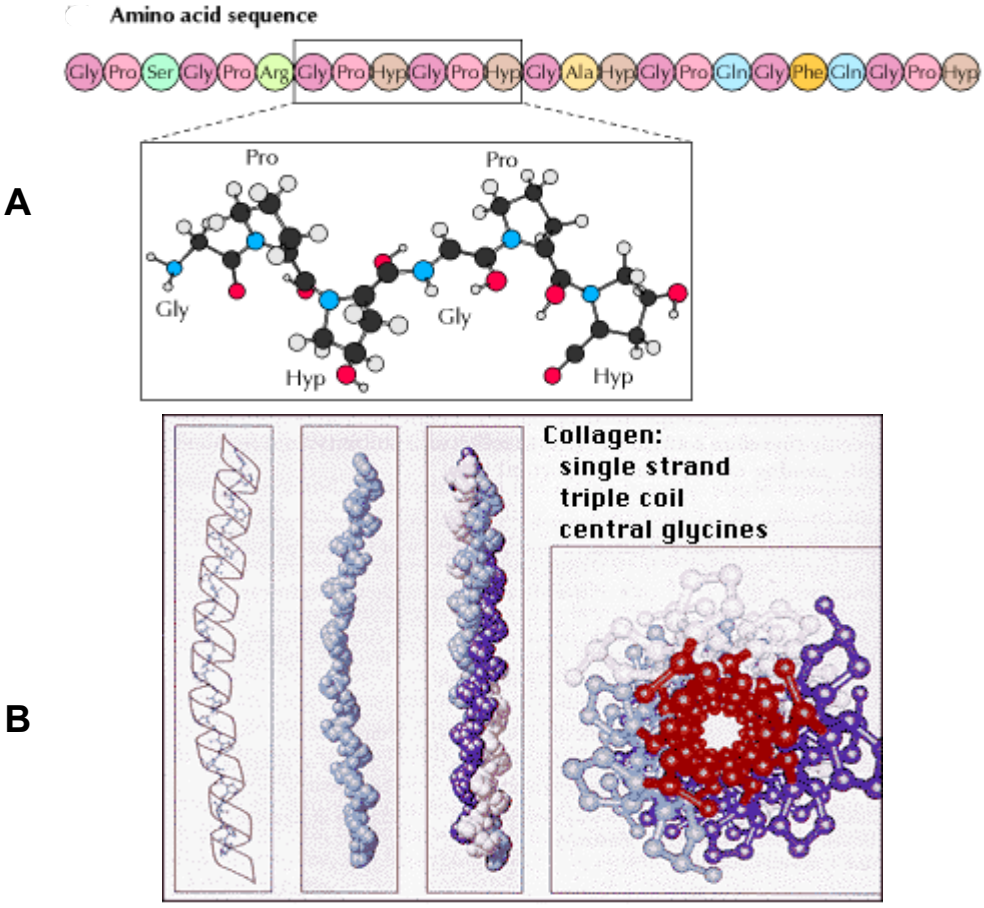


**Fig 1.2 Arrangement of amino acids in the Gly-X-Y repeat in the triple helical part of a fibrillar collagen.** (A) Characteristic Gly-X-Y repeat with the orientation of glycine. (B) The central fit of glycine in the structure of triple helix with all other amino acid being more peripheral (Cooper & Hausman 2004, Nelson & Cox 2005).

## 1.5 Biosynthesis of collagen

Collagen biosynthesis and assembly follows the normal pathway of a secretory protein. Fig 1.3 represents a schematic illustration of the intracellular and extracellular events during collagen biosynthesis, which are briefly described below.

**1. Synthesis:** Collagen chains are synthesized as longer precursors called procollagens. The growing polypeptide chains are co-translationally transported into the lumen of the rough endoplasmic reticulum (ER). In the ER, the procollagen chain undergoes a series of post-translational modifications, in particular hydroxylation and glycosylation.

**2. Hydroxylation**: Specific proline and lysine residues in the propeptide are hydroxylated by membrane-bound hydroxylases. These hydroxyproline residues are essential for the stability of triple helical molecules since underhydroxylated procollagens are not stable at normal body

temperature. Lysyl hydroxylase and proline hydroxylase catalyses the hydroxylation of lysine and proline residues in the procollagen chains. The hydroxylysine residues serve as attachment sites for carbohydrate units and are also needed for the formation of intermolecular cross-links during the collagen fibril formation. Ascorbate is a cofactor required for hydroxylase activity and a deficiency may lead to triple helix instability.

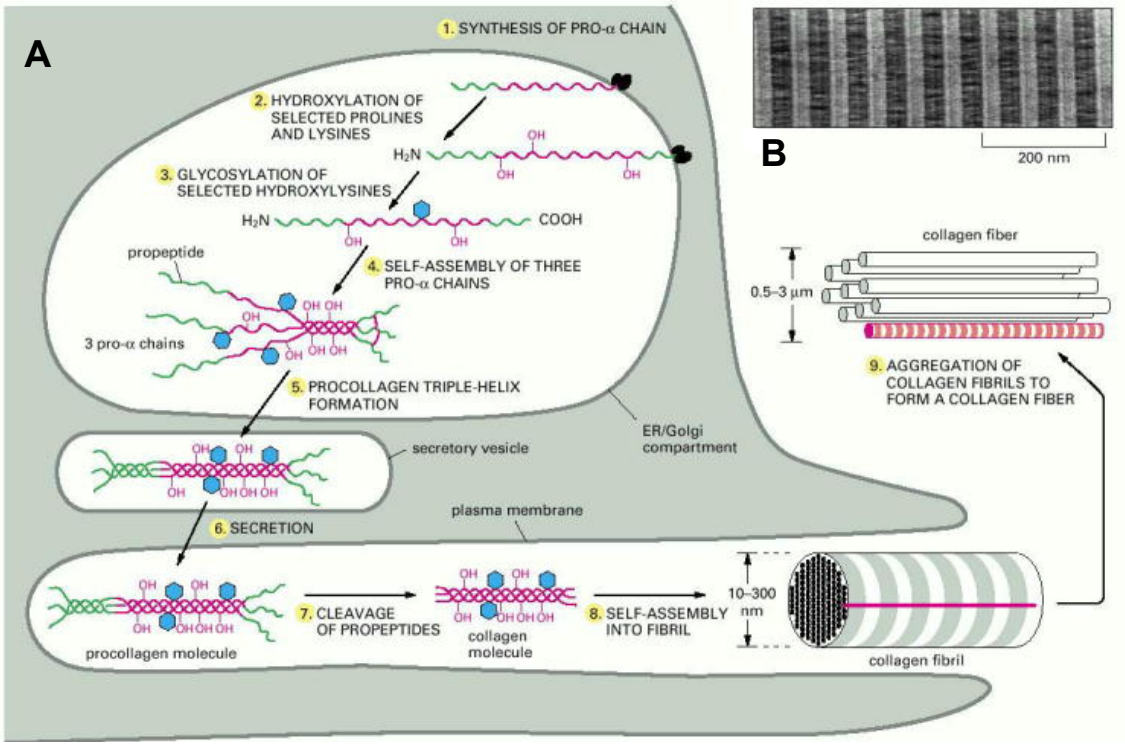


**Fig 1.3 Intracellular and extracellular events during collagen biosynthesis and fibrillogenesis.** (A) Procollagen synthesis, posttranslational modification, secretion and fibril assembly. (B) Electron micrograph of a negatively stained collagen fibril reveals its typical striated appearance (Alberts et al. 1994).

**3. Glycosylation:** Procollagens are further glycosylated in the ER and in the Golgi complex. Galactose and glucose residues are added to hydroxylysine residues and longer oligosaccharides are transferred to certain asparagine residues in the C-terminal propeptide.

**4. Procollagen assembly**: N- and C-propeptides first fold and intrachain disulphide bonds are formed within them (Bächinger et al. 1981, Doege & Fessler 1986). Three pro α chains then associate through non-covalent interactions between the folded C-propeptides followed by formation of interchain disulphide bonds catalyzed by protein disulphide isomerase (PDI) (Olsen et al. 1976, Forster & Freedman 1984, Noiva & Lennarz 1992). The formation of a triple-helical procollagen molecule begins by association of the C-propeptides of the three pro α chains

(Bächinger et al. 1981, Bulleid et al. 1996) and the triple helix is formed in a zipper-like manner proceeding from the C-terminal region towards the N-terminus (Engel & Prockop 1991).

**5. Quality control:** During the entire process molecular chaperones play an important role in regulating the correct folding and assembly of proteins (Fig 1.4). Chaperones bind to hydrophobic regions of unfolded proteins and prevent aggregation while the polypeptides are synthesized (Gething & Sambrook 1992, Becker & Craig 1994, Hartl et al. 1994). They also serve as folding quality controllers, preventing the secretion of misfolded proteins and promoting their degradation. BiP is a 78 kD protein located within the lumen of the ER and plays a general role in protein folding and oligomeric assembly (Haas & Wabl 1983). In certain osteogenesis imperfecta patients, mutant type I collagen remains bound to BiP (Chessler & Byers 1992). PDI also acts as a molecular chaperone during the assembly of procollagen chains (Wilson et al. 1998). PDI interacts specifically with the propeptides of monomeric type I procollagen chains and prevents their premature assembly or aggregation (Wilson et al. 1998).

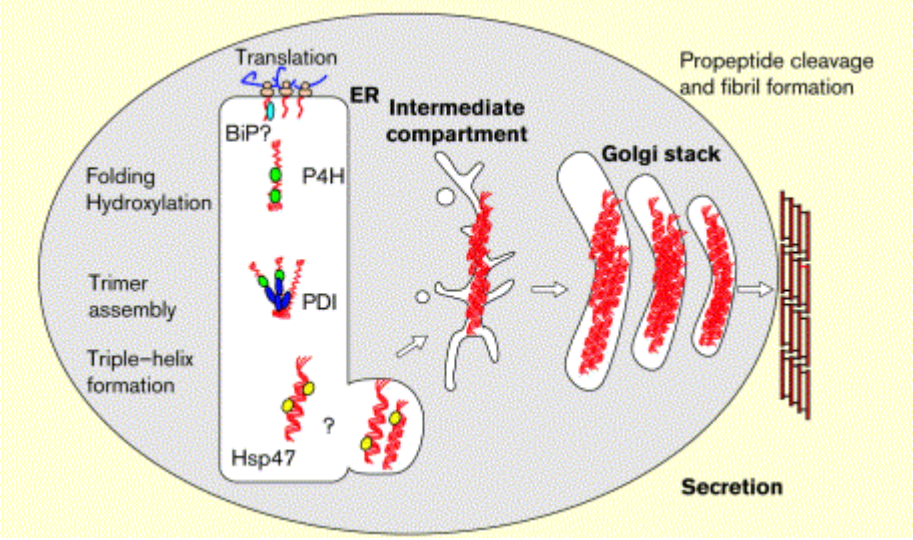


**Fig 1.4 Interactions of ER chaperones during procollagen assembly and secretion.** Procollagen interacts with a variety of ER resident proteins during folding and assembly within the ER. The carboxy-terminal propeptide may interact transiently with BiP (light blue) before it folds correctly and assembles covalently with two other pro-α-chains to form a trimer. Protein disulfide isomerase (PDI; blue) catalyzes this latter event. Proline hydroxylases (P4H; green) modifies selected proline and lysine residues in the triple helix-forming domains. These hydroxylated residues form hydrogen bonds that stabilize the helix once it forms. This triple helical form of procollagen is the preferred substrate of Hsp47 (yellow) (Hendershot & Bulleid 2000).

Hsp47 is an endoplasmic reticulum (ER)-resident molecular chaperone that is specific for collagen and plays a role in collagen maturation (Nagata & Hosokawa 1996, Nagata 2003). After binding to procollagen in the ER, Hsp47 is co-transported with procollagen from the ER to the

cis-Golgi or ER-Golgi intermediate compartment, where it dissociates from procollagen in a pH-dependent manner (Saga et al. 1987, Satoh et al. 1996).

**6. Secretion:** The mechanism responsible for transporting assembled procollagen molecules from the ER to the Golgi complex is poorly understood. Various studies have shown that triple-helical procollagen molecules form large electron-dense aggregates in a cis-Golgi compartment of fibroblasts and that these aggregates then move across the Golgi stacks without leaving the lumen of the Golgi cisternae (Bonfanti et al. 1998). Progressive maturation of the Golgi cisternae appears to be responsible for the transport through the Golgi complex (Mironov et al. 1997). If triple helix formation is prevented and random coil polypeptide chains accumulate within the cisternae of the rough endoplasmic reticulum, they are degraded in part or secreted at a delayed rate (Kivirikko et al. 1992). Detailed investigations using video-electron microscopy, serial-section 3D reconstruction and electron tomography have shown that procollagens are transported in saccular structures formed directly from protruding portions of the ER membrane (Mironov et al. 2003).

**7. Procollagen processing:** In the extracellular space the N- and C-propeptides of procollagens are cleaved by specific proteinases before the molecules start to self-assemble into fibrils (Prockop & Hulmes 1994, Canty & Kadler 2005) (Fig 1.5). These procollagen proteinases are endopeptidases and require a divalent cation such as $Ca^{2+}$ for maximal activity (Hojima et al. 1989, Prockop et al. 1998). Propeptides are cleaved by two enzymes, procollagen N proteinase and procollagen C proteinase, which cleave N- and C-propeptides respectively (Prockop et al. 1998, Kivirikko 1995) Procollagen C proteinase activity (Hojima et al. 1985) is exerted by members of the tolloid family of zinc metalloproteinases [bone morphogenetic protein 1 (BMP-1), mammalian tolloid (mTLD) and tolloid like 1 (TLL-1); (Hartigan et al. 2003, Kessler et al. 1996, Li et al. 1996, Scott et al. 1999)] and N-proteinase activity is provided by members of the ADAMTS (metalloproteinase with thrombospondin motifs) family: ADAMTS-2, ADAMTS-3 and ADAMTS-14 (Colige et al. 1997, 2002, Fernandes et al. 2001).

**8. Fibril assembly:** Fibril formation starts after removal of the globular N and C propeptides from procollagen by the procollagen N- and C-proteinases (Leung et al. 1979) (Fig 1.5). Cleavage of the C-propeptides reduces the solubility of the protein and the collagen molecules self-assemble into fibrils (Kadler et al. 1987, Prockop & Hulmes 1994, Canty & Kadler 2005). After assembly of the collagen molecules into fibrils, covalent cross-links are formed and these provide the fibrils with their tensile strength and mechanical stability. The crosslinks are formed

from lysine and hydroxylysine derived aldehydes that are synthesized in a reaction catalyzed by lysyl oxidase (Prockop & Kivirikko 1995), which causes the oxidative deamination of the ε-amino group in certain lysine and hydroxylysine residues (Kagan & Trackman 1991). The crosslinks formed from a hydroxylysine derived aldehyde are more stable than those formed from a lysine derived aldehyde (Prockop & Kivirikko 1995). A characteristic feature of the fibril forming collagens is that they form highly ordered, quarter-staggered, 67 nm banded fibrils, i.e. adjacent molecules overlap by a distance of 67 nm or a multiple of this, with a 40 nm gap between the ends of the continuous non-overlapping molecules (Prockop & Kivirikko 1995, Bella et al. 1994). Collagen fibril diameters range from 20 to 500 nm, depending on the tissue and age.

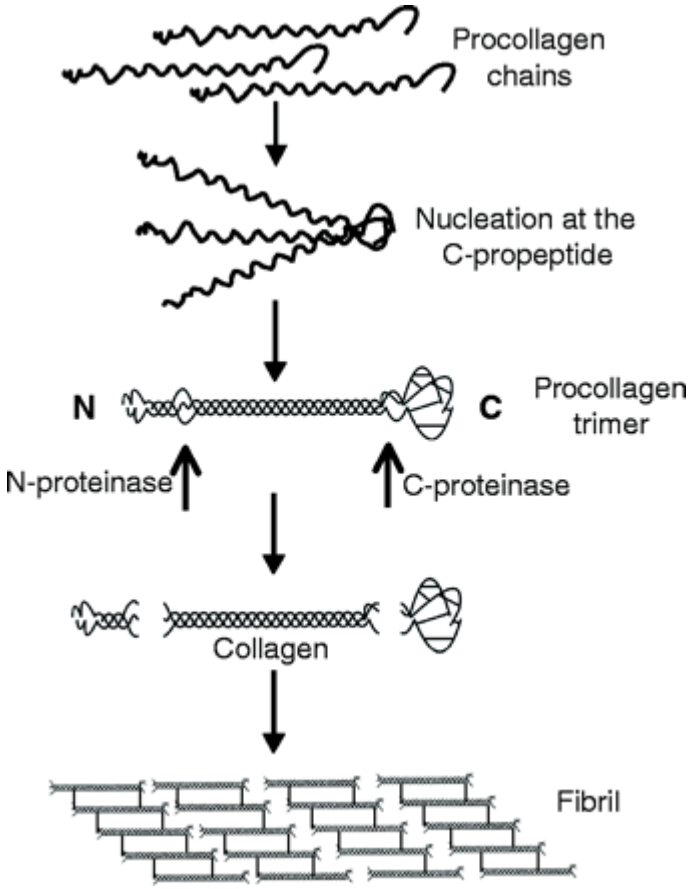


**Fig 1.5 Overview of the steps involved in the production of collagen fibrils by fibroblasts.** Procollagen chains are synthesized in the endoplasmic reticulum (ER), are brought together by interactions between the C-propeptides and fold to form a rod-like triple-helical domain flanked by globular N- and C-propeptides. Removal of the N- and C-propeptides from fully folded procollagen only occurs after transport of procollagen across the Golgi stacks and results in collagen molecules that are then able to assemble into fibrils. Covalent crosslinks occur within and between triple-helical collagen molecules in fibrils (Canty & Kadler 2005).

## 1.6 Collagen II

Collagen II is the major cartilage collagen, but is also expressed in the vitreous of the eye and is detected during early embryogenesis in the cranial mesenchyme and sclerotome of the somites. Collagen II transcripts are further present in the axial and appendicular skeleton, in nonchondrogenic tissues such as notochord, neural retina, cornea and conjunctival epithelia and in the sclera of the developing eye (Brewton & Mayne 1992). Unexpectedly, localisation of type II collagen mRNA was also reported in the proliferative ventricular cells of the forebrain and midbrain of embryos and in the cervical spinal cord (Cheah et al. 1991). The gene that codes for the polypeptide chains of homotrimeric collagen II consists of 54 exons and is about 31 kb in size (Ala-Kokko & Prockop 1990, Ala-Kokko et al. 1995). Its chromosomal location is 12q13.11-q13.12 (Takahashi et al. 1990). The sequence coding for the N-terminal propeptide contains an alternatively spliced exon that codes for a 69-amino acid cysteine-rich domain (Ryan & Sandell 1990). Collagen II molecules including (IIA) or excluding this domain (IIB) have distinct distributions in various stages of chondrogenesis, type IIA predominating in the prechondrogenic mesenchyme and differentiating chondrocytes, and type IIB in differentiated chondrocytes (Nah & Upholt 1991, Sandell et al. 1991). The collagen II molecules in the fibril overlap with each other by a distance of about a quarter of their length, thus forming a banded fibril. The molecules are covalently crosslinked between the triple helical domain and the N- or C-terminal telopeptides of adjacent collagen II molecules. Collagen II has a high capacity to withstand forces exerted. Indeed, it has a tensile strength comparable to steel and thereby strongly influences the biomechanical properties of cartilage (Vikkula et al. 1994). Due to the complexity of the folding and assembly, collagen II structure is susceptible to mutations leading to disease states that are collectively termed chondrodysplasias and possess a wide range of well-characterised clinical phenotypes (Vikkula et al. 1994, Mundlos & Olsen 1997, Myllyharju & Kivirikko 2001).

## 1.7 Chondrodysplasias

Most chondrodysplasias are due to point mutations in the collagen II gene. Most of these result in the substitution of glycine in the Gly-X-Y repeats and thereby the stability of the triple helix is decreased. Mutations in the X or Y position in the peptide chain are less frequent. In addition, nonsense mutations occur which lead to premature termination of the chain. Deletions have also been reported, leading to severe pathological conditions (Vikkula et al. 1994, Mundlos & Olsen

1997, Myllyharju & Kivirikko 2001). Table 1.1 summarizes the known mutations and the resulting phenotype. A brief description of each phenotype is given below.

**Table 1.1 Types of chondrodysplasia caused by point mutations and the resulting amino acid exchanges in collagen II.** Numbers indicate the position in the triple helical domain.

| Achondrogenesis II | Hypochondrogenesis | SEDC | Stickler syndrome |
|---|---|---|---|
| Gly253Asp | Gly313Ser | **Arg 75 Cys** | Arg9Term |
| Gly310Asp | Gly517Val | **Arg789Cys** | Gly 67Asp |
| Gly313Ser | Gly571Ala | Gly973Arg | Arg365Cys |
| Gly517Val | Gly604Ala | Gly997Ser | Leu467Phe |
| Gly571Ala | Gly691Arg | | Gly506Term |
| Gly571Asp | Gly805Ser | | Arg519Cys |
| Gly580Arg | Gly853Glu | | Arg585Term |
| Gly595Arg | Gly913Cys | | **Arg704Cys** |
| Gly694Glu | Gly988Arg | | Arg732Term |
| Gly748Asp | Thr1190Asn | | |
| Gly769Ser | | | |
| Gly781Ser | | | |
| Gly817Val | | | |

SEDC, Spondylo Epiphyseal Dysplasia Congenita; Term, Termination
Mutations in bold are used in the present study

## 1.8 Achondrogenesis type II and hypochondrogenesis (OMIM 200610)

Achondrogenesis type II and hypochondrogenesis were considered two separate disorders (Langer et al. 1969, Saldino 1971, Maroteaux et al. 1983). After individuals with an intermediate phenotype have been characterized they are defined as different degrees of severity of the same disorder (Borochowitz et al. 1986). Infants die perinatally or within the first weeks of life. The disorder is characterized by a short, barrel-shaped trunk, very short extremities, large head, soft cranium, flat face and hydropic appearance. Radiographically, the affected infants show varying degrees of underossification of the axial skeleton. Histological and electron microscopic examinations reveal hypercellular epiphyseal cartilage with poorly organized or absent growth plate and diminished extracellular matrix that contains thick, irregular collagen fibrils together with large chondrocytes with a dilated rough endoplasmic reticulum. Biochemical studies of

hyaline cartilage from these infants have indicated abnormal and diminished collagen II (Eyre et al. 1986, Godfrey et al. 1988, Godfrey & Hollister 1988, Spranger et al. 1994.). The genetic defect leading to this pathology was found in the COL2A1 gene (Vissing et al. 1989). Table 1.1 lists various mutations leading to achondrogenesis II. All of them involve the replacement of glycine by a bulkier amino acid in the triple helical region of the 1(II) chain.

## 1.9 Spondyloepiphyseal dysplasia and congenital spondyloepiphyseal dysplasia (SEDC, OMIM 183900)

Spondyloepiphyseal dysplasia is characterized by short trunk, short extremities, barrel-shaped chest, kyphosis, severe myopia, retinal detachments, cleft palate and clubfoot (Fig 1.6A). Radiographs (Fig 1.6B) show defects in ossification of the spine and primarily proximal extremities (Mortier et al. 2000).

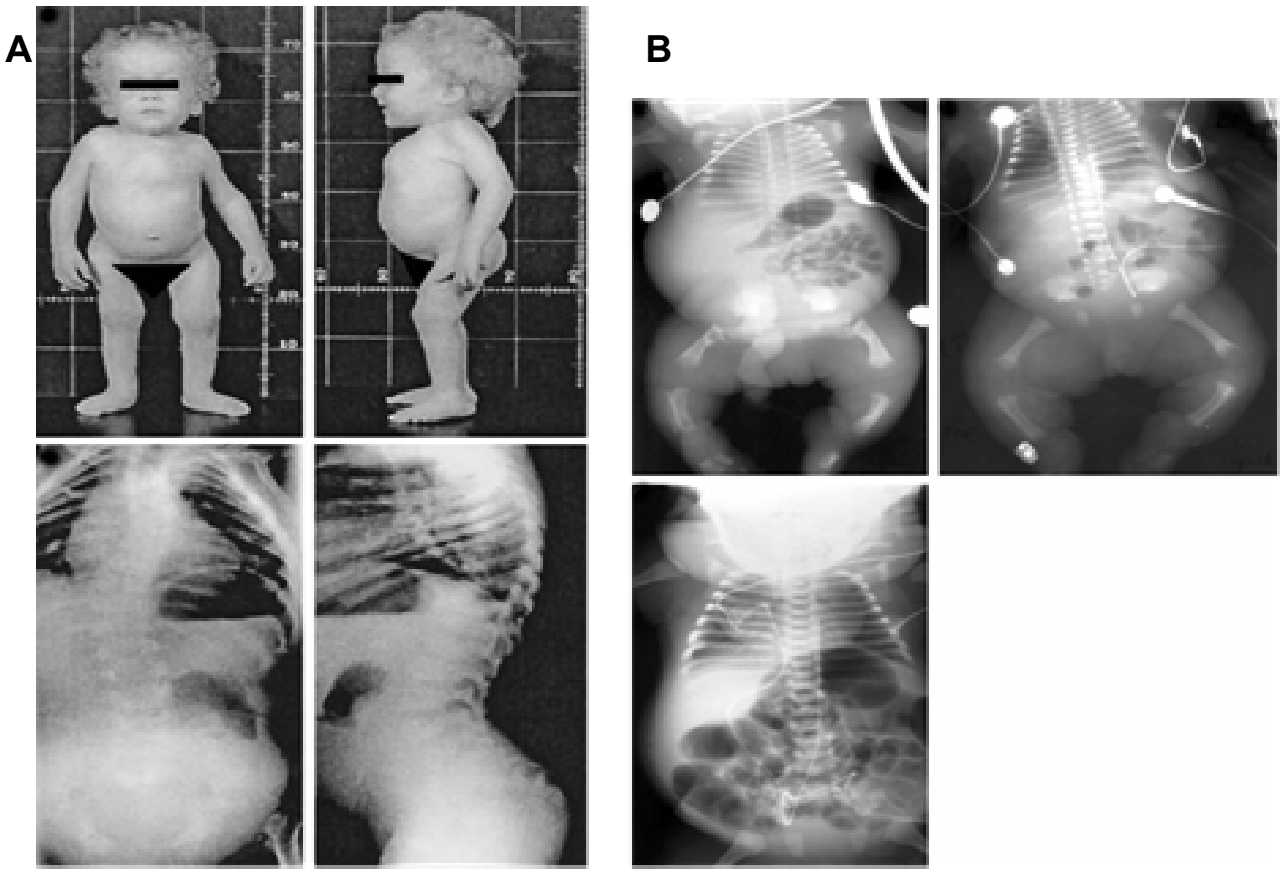


**Fig 1.6 Phenotype of SEDC patients. (**A**)** Disproportionate short stature with short trunk, enlarged joints, high forehead, frontal bossing, pectus carinatum deformity and radiological findings of anteroposterior and lateral spine radiographs shows: thoracic scoliosis, accentuated lumbar lordosis, flattened and dysplastic vertebral bodies (Sobetzko et al 2000). (B) Anteroposterior radiographs of patients representing a severe phenotype characterised by short tubular bones with metaphyseal irregularities, minimal ossification of vertebral bodies in the thoracic spine, hypoplastic iliac wings with flat acetabular roofs, short ribs along with shortened tubular bones, flattened but normally ossified vertebral bodies, normal acetabular roofs, and a less narrow chest (Mortier et al 2000).

Histological studies show defects resembling those in achondrogenesis type II-hypochondrogenesis and biochemical examination indicates abnormalities in collagen II (Murray et al. 1989). SED refers to a heterogeneous group of disorders that affect the spine and epiphyses, and as the symptoms resemble those of achondrogenesis type II-hypochondrogenesis, SED can be considered to represent the mild end of the spectrum comprising achondrogenesis type II-hypochondrogenesis and SED phenotypes (Spranger et al. 1994). Most cases of SED show autosomal dominant inheritance, but an autosomal recessive form of inheritance has also been suggested in some cases (Harrod et al. 1984). Lee et al. (1989) identified a deletion of an entire exon in the COL2A1 gene, leading to a lack of 36 amino acids in the triple helical region of the chain, and subsequently numerous mutations, including insertions, mutations causing aberrant RNA splicing, glycine mutations and a cysteine for arginine substitution, have been characterized (Vikkula et al. 1994, Kuivaniemi et al. 1997). However, some investigations have excluded the COL2A1 gene as a locus for SED in certain families, indicating that there must also be some other locus or loci for the disorder (Wordsworth et al. 1988, Anderson et al. 1990).
SEDC is of autosomal dominant inheritance and affects the vertebrae, juxtatruncal epiphyses and leads to a lack of ossification of pubis, distal femoral and proximal tibial epiphyses, talus and calcaneues, flattening of vertebral bodies, short limbs, cleft palate, mental retardation, myopia etc. (Gunthard et al. 1995). Previous studies indicate that the processing and assembly of collagen is disrupted but the exact mechanism of pathogenesis is yet to be unraveled.

## 1.10 Stickler syndrome (OMIM 108300)

Stickler syndrome, originally called hereditary progressive arthroophthalmopathy, is an autosomal dominantly inherited disorder and is caused by mutations in collagen II or collagen XI. The α 3 (XI) chain is the IIB splicing variant product of the COL2A1 gene (Wu & Eyre, 1995). Phenotypically, it is characterised by progressive myopia and vitreoretinal degeneration resulting in retinal detachment and blindness. Midfacial hypoplasia and micrognathia are often accompanied by cleft palate or a lesser degree of clefting, and a sensorineural hearing defect is common (Fig 1.7). The skeletal manifestations include juvenile progressive arthropathy, irregularity of the vertebral bodies and hypermobile joints. Skeletal growth is usually normal. Mild epiphyseal dysplasia and overtubulation of long bones are seen in radiological examination (Temple 1989, Snead & Yates 1999). Linkage studies have shown that the COL2A1 gene is a locus for Stickler syndrome (Francomano et al. 1987, Knowlton et al. 1989). Altogether nine mutations have been reported to date, all of them leading to premature termination of translation

(Ahmad et al. 1991, Brown et al. 1992, Ahmad et al. 1993, Ritvaniemi et al. 1993, Brown et al. 1995a, Ahmad et al. 1995, Williams et al. 1996). The COL2A1 gene has also been excluded in several families with Stickler syndrome (Knowlton et al. 1989, Vintiner et al. 1991, Bonaventure et al. 1992).

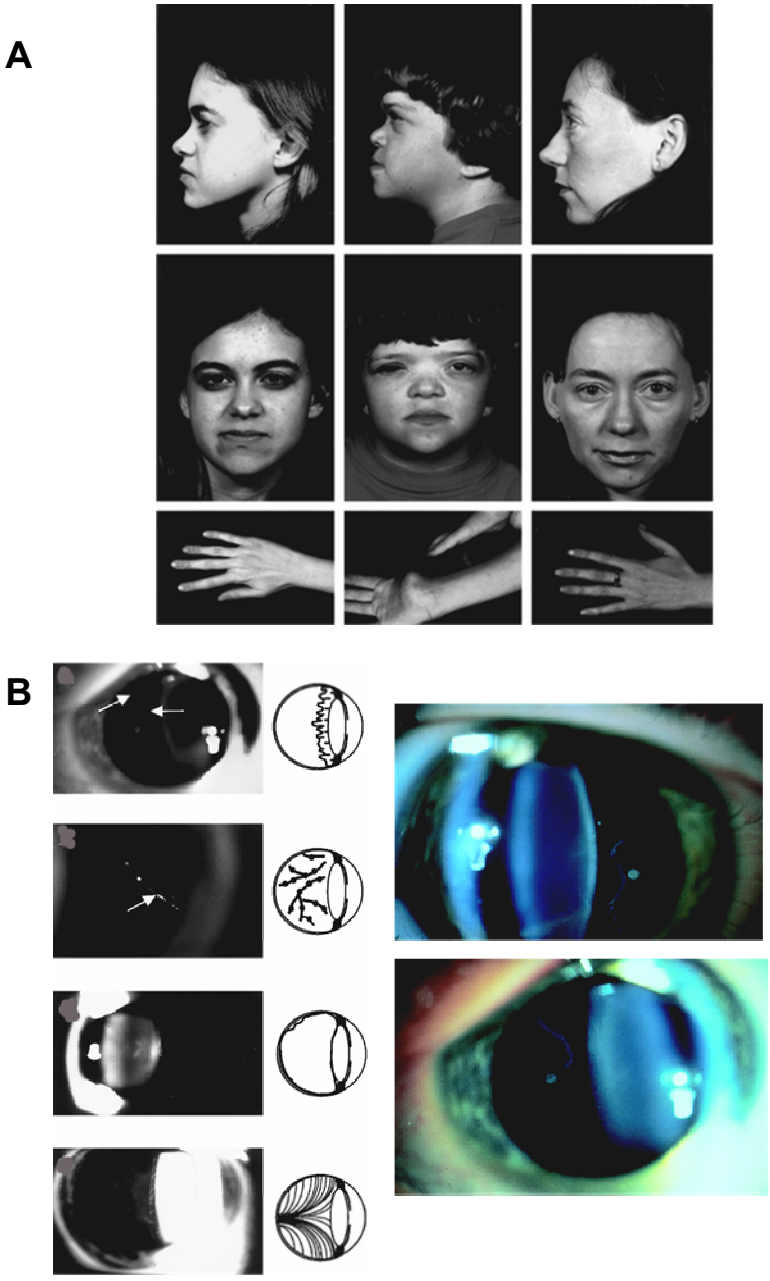


**Fig 1.7 Clinical phenotypes of patients with Stickler syndrome.** (A) Probands with mild nasal root hypolasia, midfacial hypoplasia and slender digits. (B) Vitreous phenotypes displaying membraneous congenital vitreous anomaly, beaded appearance of vitreous and a fibrillar vitreous anomaly in comparison with normal vitreous (Richards et al. 2000, 2002a, 2002b).

Stickler families with congenital vitreous anomaly (Stickler syndrome type 1) have linkage to the COL2A1 gene, while in those with congenitally defective vitreous gel architecture but no congenital vitreous anomaly (Stickler syndrome type 2) linkage is excluded (Snead et al.1994). Subsequently, Richards et al. (1996) demonstrated linkage to the COL11A1 gene in a family with Stickler syndrome type 2, and traced it to a missense mutation converting glycine to valine in the N-terminal part of the triple helix of the α 1(XI) chain. In addition, Wilkin et al. (1998) have suggested still another locus for Stickler syndrome apart from the genes coding for collagens II and XI. It has also been reported that mutations in collagen II in the X position cause haploinsufficiency and that leads to reduced amount of collagen, resulting in characteristic membranous congenital anomaly in the vitreous (Snead et al. 1994, Snead and Yates 1999, Richards et al. 2000).

## 1.11 Osteoarthritis (OMIM 165720)

Osteoarthritis (OA) is the most common articular disorder, characterized by joint pain and tenderness, stiffness, crepitus and limitation of motion. Radiographs of patients indicate joint-space narrowing, osteophytes, subchondral bone sclerosis and subchondral cyst formation. In severe cases, deformity of the bone ends is also seen. OA is classified into primary (idiopathic) or secondary subsets, in primary cases there is no known predisposing factor, while in secondary cases factors such as trauma, infection, some other joint disorder or a metabolic disease is indicated. Further classification divides OA into localized, in which only certain joints are affected, and generalized subsets. Even though OA is often considered an consequence of ageing, female sex, obesity, occupational load and heavy sports exertion are risk factors (Altman 1995, Creamer & Hochberg 1997). The suggestion of genetic factors underlying OA in certain families was made several decades ago (Stecher et al. 1953, Kellgren et al. 1963). Palotie et al (1989) demonstrated a linkage to the COL2A1 gene in a family with primary OA. Knowlton et al. (1990) made a similar finding in a family with precocious OA and associated mild chondrodysplasia and Ala-Kokko et al. (1990) subsequently demonstrated a mutation in this family. The mutation converted arginine to cysteine, an amino acid not normally found in the triple helical domains of collagens, in the middle of the triple helix of the collagen II molecule. Altogether five families with this same mutation have been reported to date, and three of them have been shown to have a common ancestor (Bleasel et al. 1995, Kuivaniemi et al. 1997, Bleasel et al. 1998). In addition, a mutation converting arginine to cysteine at position 75, near the N-terminus of the triple helix, has been reported in three families with OA associated with mild SED (Bleasel et al. 1995, Bleasel et al. 1996). Osteoarthritis is also associated with

mutations in other non collagenous proteins; e.g. a missense mutation in the epidermal growth factor-like domain in matrilin-3 leading to OA in several families. (Stefansson et al. 2003). Kizawa et al. (2005) identified polymorphism in asporin that suppressed TGF-β-mediated expression of the genes coding for aggrecan (AGC1) and type II collagen (COL2A1) accompanied with reduced proteoglycan accumulation leading to osteoarthritis. Similar forms of osteoarthritis were reported to be due to functional variants within the secreted frizzled-related protein 3 gene (Loughlin et al. 2004).

## 1.12 Animal models for chondrodysplasias due to collagen II mutations

Transgenic mice carrying a partially deleted human COL2A1 gene developed the phenotype of a chondrodysplasia with dwarfism, short and thick limbs, short snout, cranial bulge, cleft palate, and delayed mineralization of bone (Vandenberg et al. 1991). In cultured chondrocytes from transgenic mice, the minigene was expressed as shortened pro-α-1(II) chains that were disulfide-linked to normal mouse type II collagen chains. Therefore, the phenotype was probably explained by depletion of endogenous mouse type II procollagen through the phenomenon of procollagen suicide. Transgenic mice harboring a glycine-to-cysteine mutation at residue 85 of the triple helical domain of mouse type II collagen displayed severe chondrodysplasia with short limbs and trunk, craniofacial deformities, and cleft palate (Garofalo et al. 1991). Electron microscopic analysis showed a pronounced decrease in the number of typical thin cartilage collagen fibrils, distention of the rough endoplasmic reticulum of chondrocytes, and the presence of abnormally large banded collagen fibril bundles. Garofalo et al. (1991) postulated that the abnormally thick collagen bundles were related to a defect in crosslinking.

A transgenic mouse model of SEDC carrying collagen II transgene with an R789C mutation, in combination with a murine Col2a1 promoter directing the gene expression to cartilage, displayed overall short stature, had shorter limbs with disorganized growth plates, a short nose, cleft palate, and died at birth (Gaiser et al. 2002). The cellular organization of the cartilage architecture was profoundly disturbed in the transgenic growth plate, although the overall polarity was maintained. There were fewer stacks of flattened chondrocytes in the proliferative zone of the transgenic mouse growth plates. In addition to the abnormal cartilaginous tissue, the perichondrium was thicker and electron microscopy revealed a marked reduction of collagen fibrils in cartilage matrix from transgenic compared with that of wild-type mice. Increased distension of rough endoplasmic reticulum was observed in these chondrocytes. Using cell culture experiments and molecular modeling, Gaiser et al. (2002) suggested that this Y-position

mutation acts in a dominant-negative manner, resulting in destabilization of collagen molecules during assembly and reduction in the number of fibrils formed.

## 1.13 Arginine to cysteine mutations

The majority of chondrodysplasia-causing missense mutations in COL2A1 are substitutions of obligatory glycine residues in the triple helical domain. Only a few non-glycine missense mutations have been reported and among these, the arginine to cysteine substitutions predominate. In addition to the above R75C and R789C substitution (Table 1.1), recently two additional mutations (R365C and R1076C) were found in unrelated probands (Hoornaert et al. 2005).

Most of these chondrodysplasias are due to point mutations in the collagen II gene, altering the amino acid sequence in a way which destabilizes the triple helix. In addition, nonsense mutations have been detected which lead to premature termination of the chain. Deletions have also been reported leading to severe pathological conditions.

Point mutations leading to a change from arginine to cysteine are interesting since different mutations of this kind cause two well characterized clinical phenotypes, i.e Stickler syndrome and congenital spondyloepiphyseal dysplasia (SEDC) and osteoarthritis-associated SED. It is being speculated that Stickler syndrome can be caused by amino acid substitutions in the X position of Gly-X-Y repeats (Richards et al 2000) and substitutions in the Y position rather lead to SEDC. One of the notions derived from previous studies on mutations in collagen I is that the mutant molecules are intracellularly degraded leading to a reduction in the pool of extracellular collagen molecules available for the formation of correct collagen fibrils. Another hypothesis is that some of the mutant collagens are incorporated into collagen fibrils, thereby affecting their biomechanical characteristics (Royce & Steinmann 2002).

## 1.14 Aims of the thesis

The mechanisms by which point mutation in collagen II cause damage to the structure of connective tissues are not fully understood, and some of the proposed pathways leading from a mutant genotype to a phenotype are controversial.

The aims of this thesis were therefore:

1. To investigate the effects of arginine to cysteine mutations in collagen II and to determine how these may lead to different clinical phenotypes.

2. To analyze the biochemical consequences of arginine to cysteine substitutions found in patients with Stickler syndrome and SEDC.

3. To determine the importance of the position of point mutations within the triple helical domain (towards the N- or C-terminal of the triple helix) with respect to triple helical integrity and overall collagen structure.

4. To elucidate the effects of these point mutations on protein trafficking, secretion and cell survival.

# 2 Materials and methods

## 2.1 Materials

Standard chemicals and enzymes were if not otherwise mentioned purchased from Merck (Darmstadt), Sigma (Taufkirchen), Invitrogen (Karlsruhe), Biozym (Oldendorf), Roche (Mannheim) or New England Biolabs (Schwalbach).

### 2.1.1 Enzymes for molecular biology

Calf intestinal alkaline phosphatase (NEB)

M-MLV reverse transcriptase (Promega)

Restriction endonucleases (NEB)

Ribonuclease A (Sigma)

T4 DNA ligase (NEB)

Taq DNA polymerase (Roche)

### 2.1.2 Materials for cell culture

Dulbecco's Modified Eagle Medium (DMEM-F12), Invitrogen

Non-essential Amino Acids (NEA), Invitrogen

L-Glutamine (200 mM), Invitrogen

Penicillin/Steptomycin, Invitrogen

Fetal Bovine Serum (FBS), Biochrom KG, Berlin

### 2.1.3 Plasmid and adenoviral vectors

#### 2.1.3.1. pBluescript KS+ (pBS KS+) phagemid

The phagemid vector pBluescript KS+ (pBS KS+) was used for subcloning and was the basic vector used for site directed mutagenesis. Fig. 2.1 shows the map of pBS KS+ vector. The complete manual and sequence information for this vector are available from Stratagene at http://www.stratagene.com/manuals/212205.pdf.

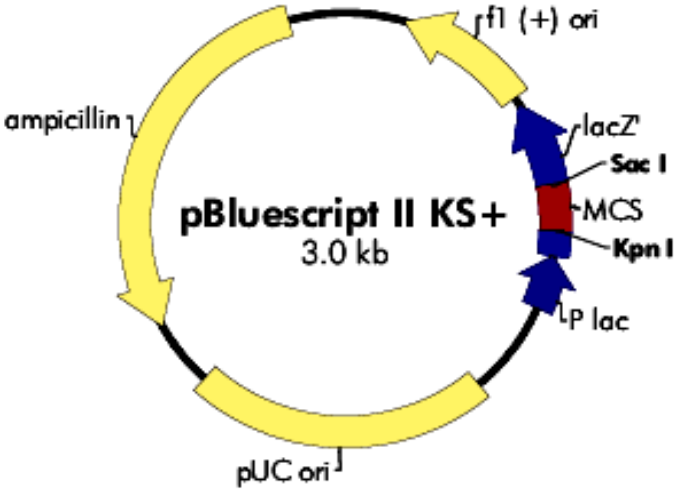


**Fig. 2.1 Map of the cloning vector pBS KS+.** It consists of an f1 and PUC origin of replication (pUC ori), an ampicillin resistance ORF and a multiple cloning site (MCS) flanked by restriction sites for Kpn I and Sac I.

## 2.1.3.2. pSL1180 superlinker phagemid

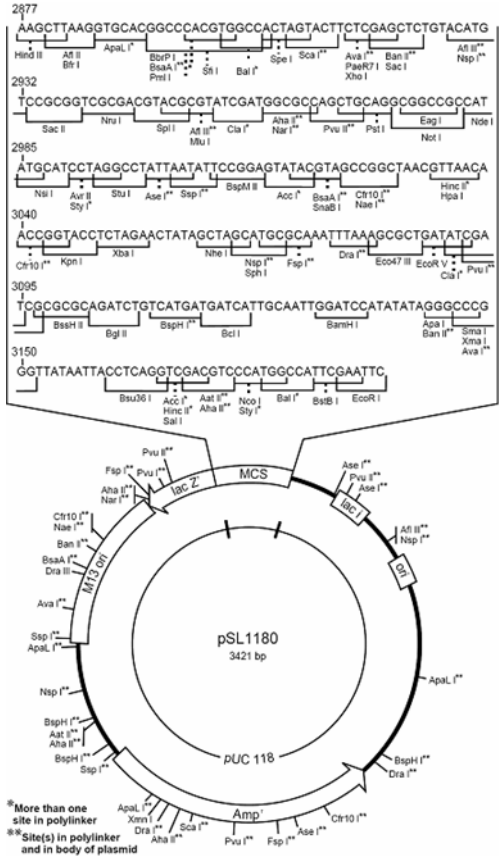


**Fig.2.2 Map of the superlinker phagemid pSL1180.** It consists of an ampicillin resistance ($Amp^r$) ORF, an M13 origin of replication and a multiple cloning site. The multiple cloning site and restriction sites are shown in detail.

The phagemid vector pSL1180 was used for subcloning of collagen II cDNA after site directed mutagenesis and to remove the endogenous signal peptide. Fig. 2.2 shows the map of pSL1180 vector with details of multiple cloning sites. A complete description and sequence informations from Amersham Biosciences can be found at http://www4.amershambiosciences.com/-aptrix/upp01077.nsf/Content/Products? OpenDocument&parentid=41175&moduleid=41179.

### 2.1.3.3 Eukaryotic expression vector pCEP-Pu

The mammalian expression vector pCEP-Pu is a derivative of the pCEP-3 vector (Invitrogen) now containing a BM-40 signal peptide and a puromycin resistance (Kohfeldt et al 1997). In addition an enterokinase cleavage site was introduced between the sequence coding for the $his_6$-myc tag and the multiple cloning site (Wuttke et al 2001). This vector was used for expression of all collagen II variants.

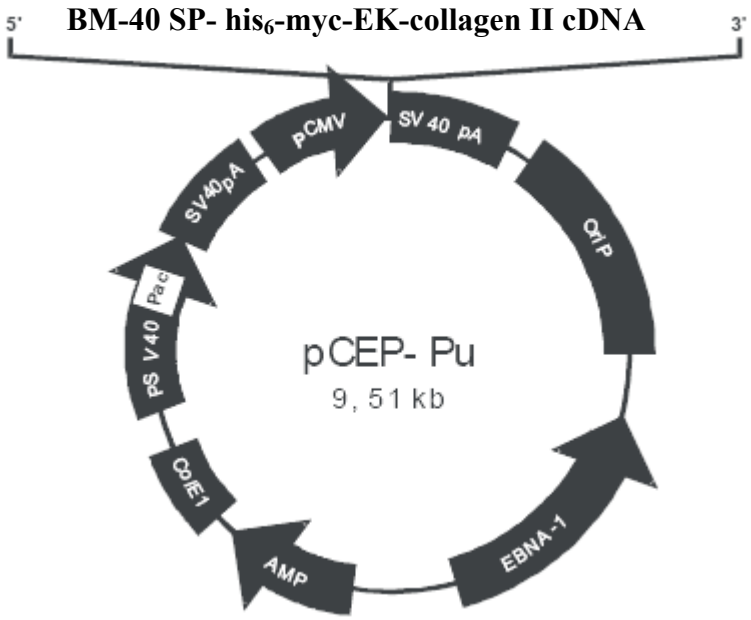


**Fig.2.3: Map of the eukaryotic expression vector pCEP-Pu (Wuttke et al 2001).** This vector contains the signal peptide from BM-40 (BM-40 SP), a $his_6$-myc-tag, the site for protease digestion by enterokinase (EK) and multiple cloning site between the CMV promoter and the SV 40 polyadenylation signal (pA).

### 2.1.3.4. Adenoviral vector pGS66 and shuttle vector pGS70

The adenoviral vector pGS66 and the shuttle vector pGS70 were kindly provided by G. Schiedner and S. Kochanek from the Center for Molecular Medicine, University of Cologne, Germany (Schiedner et al. 2000). The adenoviral vectors were used for transfection of primary chondrocytes.

Collagen II cDNA from pCEP-Pu and pGS70 vector was digested with Hind III and Bam HI. After separation by electrophoresis the collagen II cDNA was purified using a gel purification kit. The linearised vector pGS70 was dephosphorylated and purified. Equimolar amounts of vector and insert were used for ligation.

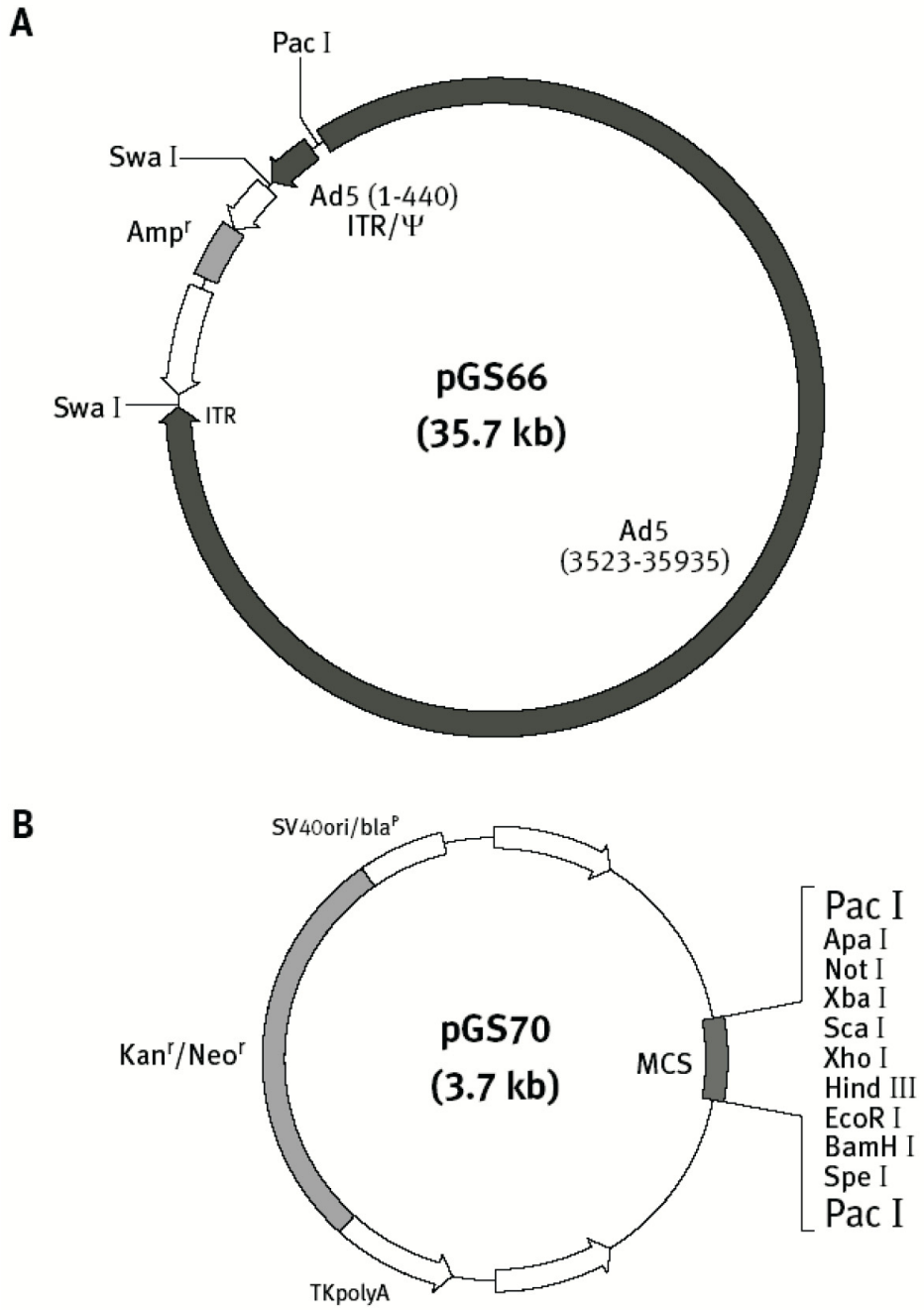


**Fig.2.4: Map of adenoviral vectors used for cloning collagen II cDNA**. (A) Map of the first generation adenoviral vector pGS66 with an ampicillin resistance (amp$^r$) ORF, an adenoviral internal terminal repeats and a Pac I site for cloning the insert. (B) Map of pGS70 vector with a multiple cloning site for use with different enzymes, a SV40 origin of replication and antibiotic resistance markers for kanamycin (Kan$^r$) and neomycin (Neo$^r$).

Collagen II cDNA was released from the vector pGS70 by digesting with the restriction enzyme Pac I. The vector pGS66 was also digested with Pac I and dephosphorylated to avoid self

ligation. Equimolar amounts of vector and insert were used for ligation. 10 µl ligation mixture was used for transformation and colonies obtained were screened for the insert by restriction digestion of the plasmid.

## 2.2 Molecular biology methods

### 2.2.1 Plasmid DNA isolation from E. coli

Plasmid DNA was prepared from small scale bacterial cultures. Bacteria were lysed by treatment with a solution containing 1% SDS and 0.5 M NaOH. SDS denatures bacterial proteins and NaOH denatures chromosomal and plasmid DNA. The mixture was neutralized with potassium acetate, causing the plasmid DNA to reanneal rapidly. Most of the chromosomal DNA and bacterial proteins remain in the precipitate, as does SDS forming a complex with the potassium, and are removed by centrifugation. The reannealed plasmid DNA from the supernatant was concentrated by ethanol precipitation.

### 2.2.2 Pure MIDI plasmid DNA isolation

Highly pure plasmid preparations were purified on either a small or a large scale using kits from Macherey-Nagel (Nucleobond AX kit for small scale plasmid preparations) or Qiagen (Qiagen Midi and Maxi Prep kit for large scale plasmid preparations). Briefly, overnight cultures of bacteria, grown in the presence of an antibiotic marker for the plasmid DNA, were pelleted and the cells were lysed by alkali. The supernantant containing plasmid DNA was passed over silica columns whereby the DNA was bound to the matrix and the impurities and digested RNA were washed off. The plasmid DNA was eluted and concentrated by ethanol precipitation. Highly pure plasmid DNA obtained was washed with 70% ethanol to remove salts and further reconstituted in TE buffer. This plasmid DNA was used for site directed mutagenesis, PCR, sequencing and transfection of eukaryotic cells. The protocols were as described by the manufacturer.

### 2.2.3 DNA agarose gel electrophoresis

| **10X DNA loading buffer** | **50X Tris acetate buffer (1000 ml)** |
|---|---|
| 40% w/v sucrose | 242.2 g Tris |
| 0.5% w/v SDS | 57.5 ml acetic acid |
| 0.25% w/v bromophenol blue | 100 ml of 0.5 M EDTA |
| in TE buffer pH 7.4 | pH 8.0, adjusted with NaOH |

**TE Buffer**

10 mM Tris-HCl

1 mM EDTA

pH 7.4

Agarose gel electrophoresis to resolve and purify DNA fragments was performed according to the method described by Sambrook et al. (1989). Electrophoresis was typically performed with 1 % (w/v) agarose gels submerged in a horizontal electrophoresis tank containing 1x TAE buffer at 1-5 V/cm. For resolving fragments less than 1,000 bp, 2.5 % (w/v) agarose gels in 1x TAE buffer were used. DNA size marker (Life Technologies) was always loaded along with the DNA samples in order to estimate the size of the resolved DNA fragments. The gel was run until the bromophenol blue dye present in the DNA loading buffer had migrated the appropriate distance. The gel was examined under UV light at 302 nm and was photographed using a gel documentation system.

### 2.2.4 Elution of DNA fragments from agarose gels

Elution of DNA fragments from agarose gels was performed using UltraClean$^{TM}$15 (MO BIO LABS Inc) DNA purification kit. DNA bands of interest were cut out of the gel along with a minimal amount of agarose and then solubilised in binding buffer at 50 °C. This solution containing DNA was incubated with a silica matrix suspension to bind the DNA which was subsequently centrifuged and washed to remove unbound material. Pure DNA was isolated by incubating the silica matrix DNA pellet with low salt alkaline elution buffer.

### 2.2.5 Measurement of DNA and RNA concentrations

Concentrations of DNA and RNA were estimated by determining the absorbance at 260 nm and 280 nm. A ratio of 1.8 indicates negligible protein contaminations.

### 2.2.6 Restriction digestion of DNA

Digestions were performed using restriction enzymes with the desired buffer and temperature as suggested by the manufacturers. Routinely 1-2 µg of plasmid DNA was digested for 1-2 h.

### 2.2.7 Dephosphorylation of 5´ends of linearized vectors

To avoid self-ligation of the vector having blunt ends or having been digested with a single restriction enzyme, the 5’ ends of the linearised plasmids were dephosphorylated by calf

intestinal alkaline phosphatase (CIP). Briefly, 1 µg DNA was suspended in the supplied NEB buffer before 1.0 U CIP was added and incubated for 60 min at 37 °C. DNA was purified by agarose gel electrophoresis.

## 2.2.8 Ligation reaction

The desired DNA fragment and the appropriate linearised plasmid were mixed in approximately equimolar amounts. T4 DNA ligase and ligation buffer containing ATP were added as indicated below and the ligation reaction left overnight at 10-12 ºC.

**Ligation reaction**

Linearised vector

DNA fragment

2 µl 10X ligation buffer

1.5 U T4 ligase

add $H_2O$ to make 20 µl

## 2.2.9 Polymerase chain reaction (PCR)

PCR can be used for in vitro amplification of DNA fragments (Saiki et al., 1985). A double stranded DNA (dsDNA) serving as a template, two oligonucleotides (primers) complementary to the template DNA, deoxyribonucleotides and heat resistant Taq polymerase are required for this reaction. Primers are designed as required (Table 2.1). The first step in PCR reactions involves the denaturing of dsDNA at 95 °C. Second, the reaction mix is incubated at different annealing temperatures, depending on the G/C content of the primers. The third step is the elongation step with a temperature of 68-72 °C that allows the elongation of the new strand of DNA by the Taq polymerase. A PCR machine (thermocycler) can be programmed to regulate these different cycles automatically. A “standard program” is presented in the table 2.2.

**Table 2.1 Primers and primer sequences used in the present study**

| primer name | sequence |
|---|---|
| XBP-1F | 5'-GGAGTTAAGACAGCGCTTGG-3' |
| XBP-2R | 5'-ACTGGGTCCAAGTTGTCCAG-3' |

**Table 2.2 Standard PCR programme used in the present study**

| step | cycles | temperature | time |
|---|---|---|---|
| Initial denaturation | 1 | 95 °C | 1 min |
| Denaturation | 25-35 | 95 °C | 50 s |
| Annealing | 25-35 | 55-68 °C | 50 s |
| Elongation | 25-35 | 68 °C | 1-15 min |
| Final elongation | 1 | 68 °C | 7 min |
| Cooling | 1 | 4 °C | |

## 2.2.10 Collagen II cDNA and patient derived mutations

Human collagen II cDNA of 4.5 kb including the poly adenylation signal was kindly provided by Fibrogen Europe. In the present work, the shorter isoform of collagen II was used (Nah & Upholt 1991, Sandell et al. 1991). Table 2.3 summarizes the patient based and artificial mutations which were selected for this study, all mutations selected lie in the triple helical domain of the protein.

**Table 2.3 Selected mutations used for the present study.** The number represents the amino acid position from the start of triple helical region.

| position in Gly-X-Y | X | | Y | |
|---|---|---|---|---|
| | position | amino acid exchange | position | amino acid exchange |
| | 134* | Arg → Cys | 75[a] | Arg → Cys |
| | 704[b] | Arg → Cys | 789[c] | Arg → Cys |
| | 740* | Arg → Cys | | |

* artificial mutation, [a] Williams et al 1993, [b] Ballo et al 1998, [c] Chan et al 1993

## 2.2.11 Subcloning of collagen II cDNA

Due to the size of the full length cDNA, collagen II cDNA was subcloned into smaller fragments which were ligated into pBluescript KS+ vector and used for site directed mutagenesis. After confirmation of the positive clones by restriction analysis and sequencing, the insert was removed and ligated back into the original vector to obtain the complete collagen II cDNA backbone with the desired mutation. The subcloning strategy is illustrated in figure 3.1.

## 2.2.12 Site directed mutagenesis

Site directed mutagenesis was carried out using the XL Quick Change mutagenesis kit from Stratagene following the manufacturers protocol. Fig 2.5 illustrates the procedure. In brief, mutagenesis was carried out by designing specific primers carrying single or double nucleotide changes in the middle of the primers leading to the mutation of interest. After sequencing of the wild type collagen II cDNA an unwanted mutation at codon 52 with a histidine in place of proline was identified. Since this mutation lies in the N-propeptide region, there was a need to correct it by site directed mutagenesis. Primers were designed (His Pro) possessing the desired change along with a silent mutation leading to a new restriction site Xma I, which facilitate identification of positive clones by restriction digestion. The primers used for respective site directed mutagenesis are listed in the table 2.4.

**Table 2.4 Primer pairs used for site directed mutagenesis**

| primer name | sequence |
|---|---|
| 75F | 5'-GGTCCTCAGGGTGCT**T**GTGGTTTCCCAGG-3' |
| 75R | 3'-CCTGGGAAACCAC**A**AGCACCCTGAGGACC-5' |
| 134F | 5'-CTGGTGAAAGAGGA**T**G**C**ACTGGCCCTGCTG-3' |
| 134R | 3'-CCAGCAGGGCCAGT**G**C**A**TCCTCTTTCACC-5' |
| 704F | 5'-GGAGCTGCTGGG**T**GCGTTGGACCCCC-3' |
| 704R | 3'-GGGGGTCCAACGC**A**CCCAGCAGCTCC-5' |
| 740F | 5'-CCCCCTGGC**T**G**C**GCTGGTGAACCCGG-3' |
| 740R | 3'-GGGTTCACCAGC**G**C**A**GCCAGGGGGG-5' |
| 789F | 5'-GGTCTGCCTGGGCAA**T**GTGGTGAGAGAGGATTCC-3' |
| 789R | 3'-GGAATCCTCTCTCACCAC**A**TTGCCCAGGCAGACC-5' |
| His ProF | 5'-GACCCAAAGGAC**C**TCC**C**GGGCCTCAGGG-3' |
| His ProR | 3'-CCCTGAGGCCC**G**GGA**G**GTCCTTTGGGTC-5' |

Numbers 75, 134, 704, 740 and 789 stand for the amino acid residue numbered from the start of the triple helix where the mutation of interest is located. His Pro stands for the primers used to correct the unwanted mutation histidine to proline (see details in the text). Nucleotides marked in bold indicate the nucleotide change to introduce the mutation of interest.

The tubes were gently mixed and a standard PCR was run as shown in Table 2.1 with 18 cycles using an annealing temperature of 60 °C for 50 s and an elongation step of 15 min at 68 °C in each cycle. A final elongation step at 68 °C for 7 min was followed. After the PCR reaction, the

enzyme Dpn I was added to each tube to digest the parental DNA and samples were incubated for 2 h at 37 °C.

**Table 2.5 Standard reaction mixture for site directed mutagenesis.**

| compound | amount |
|---|---|
| 10X reaction buffer | 5.0 µl |
| cDNA template | 2.0 µl (40 ng) |
| primer F | 1.25 µl (125 ng) |
| primer R | 1.25 µl (125 ng) |
| dNTP mix (10 mM) | 1.0 µl |
| Quick solution | 3.0 µl |
| dd $H_20$ | 36.5 µl |
| Pfu Turbo DNA polymerase | 1.0 µl |

10 µl of each sample was run on a 1% agarose gel. Another 10 µl was used for transformation of XL10 – Gold ultracompetent cells (Stratagene) and SURE cells. In brief, the ultra competent cells were thawed on ice. 2 µl of ß-mercaptoethanol was added to 25 µl of each cell suspension, swirled gently and incubated on ice for 10 min with occasional swirling. 10 µl of Dpn I digested DNA was added to the cells, swirled gently to mix and incubated for 30 min on ice, heat pulsed for 30 s at 42 °C and incubated again for 2 min on ice. Finally 0.5 ml of LB or SOC medium was added and incubated at 37 °C for 1h at 250 rpm shaking. 200 µl of the transformation mixture was plated on a LB-ampicillin agar plate and incubated at 37 °C for at least 16 h.

**SOC medium**

20 g Bacto-tryptone
5 g Bacto-yeast extract
0.5 g NaCl
2.5 ml 1 M KCl
add $H_2O$ to make 1000 ml
pH to 7.0 with 10 N NaOH
add 20 ml of sterile 1 M glucose before use after autoclaving

**Luria-Bertani (LB) Medium**

10 g Bacto-tryptone
5 g Bacto-yeast extract
10 g NaCl
add $H_2O$ to make 1000 ml
Autoclave

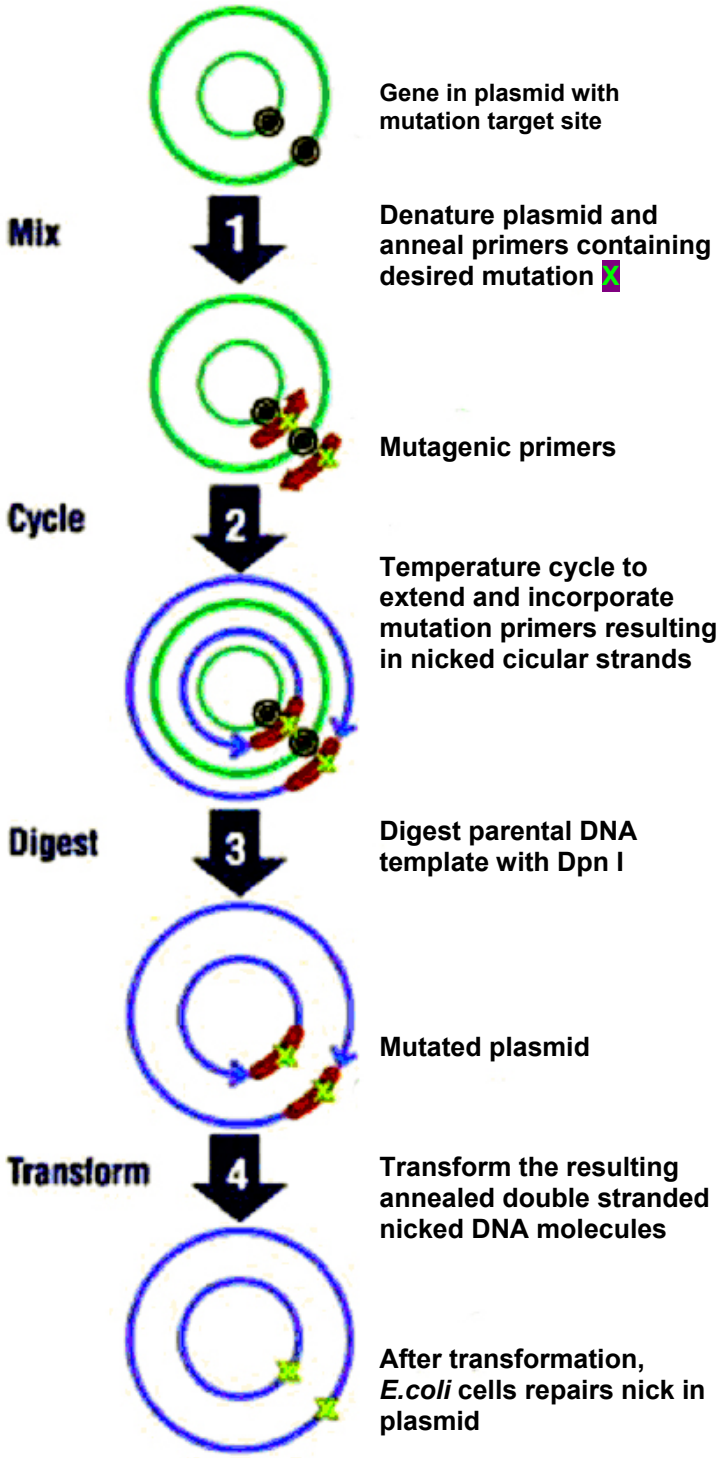


**Fig. 2.5: Schematic representation of XL10 quick change site directed mutagenesis.** A detailed description of the figure is explained in the text.

## 2.2.13 Removal of the collagen II signal peptide

Before cloning into the mammalian expression vector pCEP-Pu, the signal peptide of collagen II cDNA had to be removed since pCEP-Pu bears a BM-40 signal peptide. Primers (Table 2.6) were designed to have the Nhe I overhangs at the 5' end of the forward primer and start of collagen II cDNA without the signal peptide. The reverse primer was 1.4 kb downstream. The PCR product was cloned into pBS KS+ vector after digesting both the PCR product and the vector with Nhe I and Sac II. This insert after sequencing was excised from the vector and cloned into collagen II cDNA which was excised from pBS KS+ vector by Spe I and EcoR V and cloned into pSL 1180 vector.

**Table 2.6. Primers used for deletion of the signal peptide sequence from the collagen II cDNA**

| primer name | sequence |
|---|---|
| Nhe I SP-F | 5'-GCCGCTAGCCCAGGATGTCCGGC-3' |
| Sac II-R | 3'-GGAAACCGCGGTTGCCGGGAGC-5' |

## 2.2.14 Screening for recombinants and sequencing

Colonies obtained were individually picked and inoculated into sterile LB media containing 200 µg/µl of ampicillin and were allowed to grow for 12 h. Plasmid DNA was isolated by a slightly modified method based on a commercially available Qiagen kit as per the manufacturer's instructions. 2 µl was used to check for size and purity on a 1% agarose gel and further confirmation obtained by restriction digestion. 1 µl of the plasmid preparation was mixed with 1 µl of primer (125 ng) and sequenced at the service facility of the Centre for Molecular Medicine, University of Cologne by modified dideoxy nucleotide termination method using a 'Perkin Elmer ABI prism 377' DNA sequencer.

## 2.2.15 Eukaryotic expression vector

pCEP-Pu with an N-terminal $his_6$-myc tag was used as the basic mammalian expression system (Wuttke et al. 2001). Wild type and mutated collagen II cDNA was excised from pSL1180 using Nhe I and Bam HI and then ligated into the pCEP-Pu expression vector which was digested with the same enzyme and dephosphorylated. 10 µl ligation mixture was transformed into SURE cells and plated these on LB agar plates containing ampicillin and grown overnight. Colonies were picked and the plasmid isolated from 2 ml cultures. These clones were screened by restriction digestion and positive clones sequenced.

### 2.2.16 RNA isolation from cultured HT1080 cells

Total RNA was isolated from transfected HT1080 cells grown in 6 well plates using the TRIZOL reagent (Invitrogen) as per the manufacturer's instruction. Briefly, 3 x $10^5$ cells which were transfected in six well plates were rinsed with PBS and 1 ml of TRIZOL reagent was added to lyse the cells. This homogenate was transferred to an RNase free microfuge tube and incubated for 5 min at room temperature to permit the complete dissociation of nucleoprotein complexes. 0.2 ml of chloroform per 1 ml of TRIZOL reagent was added, mixed well by vortexing for 15 s and incubated for 2 to 3 min. Samples were centrifuged at 12,000 g for 15 min at 2 °C. During centrifugation the mixture separated into a lower red phenol-chloroform phase, an interphase, and a colorless upper aqueous phase. RNA remained exclusively in the aqueous phase which was carefully transferred to a fresh tube without transferring any of the interphase material. RNA was precipitated from the aqueous phase by mixing with 0.5 ml isopropyl alcohol and incubating the samples at room temperature for 10 min followed by centrifugation at 12,000 g for 10 min at 2 °C. The RNA precipitate obtained was washed with 70% ethanol to remove all traces of isopropanol and centrifuged again. The RNA pellet was reconstituted in 20 µl of DEPC water.

### 2.2.17 Reverse transcription

Isolated RNA was converted into cDNA using reverse transcriptase (Superscript II, Invitrogen) and an Oligo-dT-Primer. 2 µl of total RNA was mixed with 25 µM Oligo-dT-Primer and the volume adjusted to 12 µl with DEPC water. Secondary structures in the RNA were removed by incubating for 10 min at 70 °C and then cooling on ice. The reaction mixture containing 0.5 µl reverse transcriptase (200 U/µl), 4 µl enzyme buffer, 0.5 µl RNase-inhibitor (40 U/µl, Roche), 1 µl dNTP mix (10 mM mix of dATP, dGTP, dCTP, dTTP) und 2 µl 0.1 % DTT was added and incubated for 50 min at 42 °C. The reaction was stopped by denaturing the enzyme by heating to 72 °C. cDNA obtained was frozen at -20 °C and 1 µl of this cDNA was used for polymerase chain reaction.

## 2.3 Cell biology methods

### 2.3.1 Cell culture and maintenance of 293 EBNA and HT1080 cells

293 EBNA cells (Invitrogen) derived from human embryonic kidney and HT1080 (CCL-121 from ATTC) cells derived from human fibrosarcoma were cultured in DMEM-F12 containing 200 U/ml penicillin, 200 µg/ml streptomycin, 20 mM L-glutamine and 10 % FBS (Biochrom), from now on referred to as standard medium, at 37 °C in a humified incubator with a 5 % $CO_2$

atmosphere. Media were supplemented with 100 µg/ml ascorbate during expression of recombinant collagens since ascorbate acts as a cofactor for proline hydroxylation during collagen biosynthesis.

### 2.3.2 Transfection 293 EBNA and HT1080 cells with FuGene6

6 µl of FuGene6 (Roche) reagent was added to a microfuge tube containing 94 µl of serum free medium, after 5 min incubation this solution was added to another microfuge tube containing 1 µg of plasmid DNA. The reaction mixture was mixed well and incubated for 15 min at room temperature. The suspension was again mixed and added dropwise on to cells which were at 70 % confluency. After 24 h fresh medium was added, containing 1 µg/ml puromycin if the cells were to be selected.

### 2.3.3 Isolation and culture of bovine chondrocytes

Bovine humeroscapular joints of 18–24 month old animals were obtained from the local slaughterhouse. Aseptically, cartilage tissue was exposed, scraped and collected. Chondrocytes were isolated by sequential digestion with pronase (0.4 % w/v; 1.5 h at 37°C) and collagenase P (0.025% w/v; overnight at 37 °C) in DMEM containing 5% v/v FBS and 50 µg/ml gentamicin. The isolated chondrocytes were plated at a density of $3\times10^4$ cells/cm$^2$. After reaching confluency cells were subcultured twice a week and grown in standard medium containing ascorbate.

### 2.3.4 Passaging of cell lines

When the cells had grown to confluency cell culture media were aspirated and the cells washed once with sterile phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4$ $7H_2O$, 1.4 mM $KH_2PO_4$, pH 7.4) to remove all traces of serum that contains trypsin inhibitor. Cells were detached from the surface using a minimal volume of 0.05% trypsin and 0.02% EDTA in PBS, pH 7.4. Cells that were difficult to detach were placed at 37 °C to facilitate dispersal, fresh growth medium supplemented with serum added to neutralize the action of trypsin and cells aspirated by gentle pipetting. If required, cells were counted before aliquoting into new culture vessels or split 1:4 during routine cell culture. Cultures were incubated at 37 °C.

### 2.3.5 Freezing cells for storage

Cells were trypsinised, resuspended in growth medium supplemented with 20% FBS and 10% DMSO (Sigma-Aldrich) which acts as cryoprotectant. Cells were transferred into cryotubes

which were slowly cooled down to -80 °C in a cryobox containing isopropanol. 24 hours later they were placed in liquid nitrogen for longer conservation.

### 2.3.6 Thawing of frozen cells

Frozen cryotubes containing the cells were rapidly defrosted in a 37 °C water bath. The thawed cell suspension was diluted in standard medium and centrifuged at 1000 g for 5 mins to remove DMSO. The cells were then resuspended in fresh standard growth medium.

### 2.3.7 Cell counting

Cells were trypsinised, collected by centrifugation (1000 g, 5 min) and resuspended in a small volume of medium. A 10 µl aliquot was introduced in a Neubauer's chamber with a pipette, cells counted and the cell number per ml calculated.

### 2.3.8 Immunofluorescence staining

Sterile glass cover slips were washed with 1 N HCl, sterile water and 70% ethanol. After drying, they were placed in a 24 well culture plate and sterilized under UV light for 20 min. 25,000 cells were added into each well and grown overnight. Cells were transfected with Fugene6 as mentioned above and next day the media were replaced with fresh standard medium containing 100 µg/ml ascorbate. Depending on the experimental requirements, cells were fixed either two or three days after transfection at room temperature with a freshly prepared 4% paraformaldehyde solution in PBS for 15 min and subsequently permeabilised with 1% Triton-X100 in PBS for 4 min. Immediately, cells were washed three times with PBS followed by blocking with 1% normal goat serum (NGS) for 30 min to avoid unspecific binding of antibodies. Subsequently the fixed cells were incubated with the primary antibodies in PBS for 60 min and washed four times with PBS before applying the fluorochrome conjugated secondary antibodies in PBS for 60 min. Cells were washed three times with PBS to remove unbound secondary antibodies and incubated with bisbenzimide (Sigma, 0.1 µg/ml) for 5 min to stain the nuclei, followed by two additional washing steps to remove unbound bisbenzimide. Cover slips containing stained cells were carefully lifted from wells and mounted on histoslides with fluorescent mounting medium (DAKO). All the steps after adding secondary antibody were performed in dark. The list of primary and secondary antibodies used is shown in the table 2.7.

**Table 2.7 List of primary and secondary antibodies used in immunofluorescence**

| Antibody | Dilution | Company |
|---|---|---|
| **Primary antibodies** | | |
| Mouse monoclonal anti-c-myc | 1:1000 | Santa Cruz |
| Rabbit polyclonal anti-c-myc | 1:1000 | Santa Cruz |
| Mouse monoclonal anti-Golgi 58K protein | 1:1000 | Sigma |
| Mouse monoclonal anti-PDI (ER marker) | 1:1000 | Stress Gene |
| Rabbit polyclonal anti-Grp78 (BiP) | 1:1000 | Stress Gene |
| Mouse monoclonal FITC conjugated rabbit anti-active caspase-3 | 1:500 | BD Biosciences |
| **Secondary antibodies** | | |
| Alexa 488 goat anti-rabbit IgG | 1:1000 | Molecular Probes |
| Alexa 546 goat anti-rabbit IgG | 1:1000 | Molecular Probes |

## 2.3.9 Image analysis

Immunofluorescence stainings were observed with an inverted fluorescence microscope. Digital images obtained were acquired with the Metamorph software and processed using Adobe Photoshop.

## 2.3.10 Comet assay or single cell gel electrophoresis (SCGE)

Fully frosted slides were precoated with 1 % solution of normal agarose in PBS and allowed to dry, this layer supports the attachment of the cell containing agarose layer. Single cell suspensions obtained by trypsinisation were counted, diluted to give approximately 5 x $10^4$ cells/ml and reduced to 100 µl volume by centrifugation. 80 µl of each final suspension was added to 400 µl of 0.5% low melting agarose maintained at 37 °C and 90 µl of this suspension was pipetted onto a precoated slide. An additional 1% low melting agarose layer without cells was added after solidification of the above layer. After solidification, the slides were placed in alkaline lysis solution (2.5 M NaCl, 100 mM EDTA, 10 mM Tris-HCl, 1% Triton X-100, 10% DMSO, pH to 10.0) and the cells were lysed in the dark at 4 °C for 1 h. Slides were then rinsed for 30 min in three changes of neutralizing solution (400 mM Tris-HCl, pH 7.5) to remove all detergents and salts. After rinsing, slides were placed side by side in a horizontal electrophoresis chamber containing freshly made electrophoresis buffer (300 mM NaOH, 1 mM EDTA, pH>13)

and incubated in the alkaline buffer for 20 to 60 min to allow unwinding of the DNA. The slides were subjected to electrophoresis at 0.6 V/cm for 30 min. Following electrophoresis, the slides were rinsed with neutralisation buffer twice, stained with 100 µl ethidium bromide solution (10 mg/ml) and scored immediately or dried in cold 100 % ethanol before storage. Comets were visualized using an inverted fluorescence microscope and evaluated using TriTek CometScore™ software.

### 2.3.11 Nick labelling

Apoptosis was detected using the DeadEnd™ Fluorometric TUNEL System from Promega following a slightly modified manufacturer's protocol. Briefly, four days post transfection cells were fixed by immersion in 4% paraformaldehyde solution for 25 min, washed twice in PBS for 5 min and permeabilized by adding 0.2% Triton X-100 solution in PBS for 5 min. After washing in PBS cells were incubated in 100 µl of equilibration buffer at room temperature for 5–10 min. Cells were incubated with 50 µl of reaction buffer (rTdT incubation buffer) at 37 °C for 60 min inside a dark humidified chamber to allow for the tailing reaction, which was terminated by adding 2X SSC for 15 min. Cover slips were washed with PBS twice to remove unincorporated fluorescein-12-dUTP. Bisbenzimide (0.1 µg/ml in PBS) was added for 5 min to stain the nuclei and cells were washed twice to remove unbound bisbenzimide. Cover slips were washed once with deionized water for 5 min at room temperature and excess water was drained and blotted with tissue paper. Cover slips were mounted on histoslides with DAKO mounting medium.

**Equilibration buffer**
200 mM potassium cacodylate (pH 6.6)
25 mM Tris-HCl (pH 6.6)
0.2 mM DTT
0.25 mg/ml BSA
2.5 mM cobalt chloride

**Nucleotide mix**
50 µM fluorescein-12-dUTP
100 µM dATP
10 mM Tris-HCl (pH 7.6)
1 mM EDTA

**rTdT incubation buffer**
Combine the following:
90 µl equilibration buffer
10 µl nucleotide mix
2 µl rTdT enzyme

**20X SSC**
87.7 g NaCl
44.1 g sodium citrate
pH 7.2

### 2.3.12 Transduction of chondrocytes and transgene expression

First generation and high capacity gutless adenoviral particles were transduced by addition of 10 to 1000 moieties of infection (MOI). This was added with 0.75 ml of fresh culture medium to each well of a six-well plate containing 3 x $10^5$ adherent cells and incubated at 37 °C. Transgene expression in chondrocytes was monitored by the EGFP expression using fluorescence and phase contrast microscopy.

### 2.3.13 Cell toxicity

Trypsinised cells were collected and washed with PBS. After adding equal amounts of trypan blue (0.5% w/v in 0.9% NaCl) to the cell suspensions, cells were immediately counted in a Neubauer's chamber and simultaneously compared with the number of live cells expressing eGFP using fluorescent microscopy.

## 2.4 Biochemical methods

### 2.4.1 Harvesting of supernatants from 293 EBNA cells

Transfected and selected 293 EBNA cells were expanded in standard growth medium supplemented with 50 µg/ml β-aminopropionitrile to avoid crosslinking of expressed collagens. On reaching confluency the FBS was reduced to 2%. Cell culture supernatants were harvested every day and fresh media supplemented with fresh ascorbate and β-aminopropionitrile were added.

### 2.4.2 Isolation and purification of collagens from the supernatants

Immobilized-metal affinity chromatography (IMAC) was used for purification of recombinantly secreted collagens with N-terminal his-tags. Supernatants were thawed and the protease inhibitors PMSF (0.5 mM) and NEM (0.5 mM) added. Supernatants (pH adjusted to 8.0) were filtered through Whatman filter paper to remove suspended particles and the filtrate was loaded onto a column already packed with 2 ml Ni-NTA superflow matrix (Qiagen) at the rate of 1 ml/min at 4 °C. After complete loading of the supernatants the impurities and unspecifically bound proteins were washed off the column with 50 ml buffer A and the bound native proteins eluted with buffer B.

| **Buffer A/Wash buffer** | **Buffer B/Elution buffer** |
|---|---|
| 50 mM $NaH_2PO_4$ | 50 mM $NaH_2PO_4$ |
| 300 mM NaCl | 300 mM NaCl |
| 20 mM imidazole | 250 mM imidazole |
| 0.05% Tween 20 | Adjust pH to 8.0 using NaOH |
| Adjust pH to 8.0 using NaOH | |

## 2.4.3 Estimation of protein concentration

The concentration of purified collagen II variants was determined by the bicinchoninic acid assay (BC assay protein quantitation kit, Uptima) as per the manufacturer's instructions. Pepsin extracted collagen from bovine nasal septum (Sigma) was used as an additional standard along with BSA.

## 2.4.4 SDS polyacrylamide gel electrophoresis (SDS PAGE)

SDS polyacrylamide gel electrophoresis was performed using the buffer system of Laemmli (1970). Gels were prepared using glass plates of 10 cm x 7.5 cm and spacers of 0.5 cm thickness. A 15 well comb was generally used for formation of the wells in the stacking gel. The composition of 8% resolving and 4% stacking gels is given in Table 2.8.

Protein samples were mixed with an equal volume of 2X SDS sample buffer whereas cells were lysed and resuspended in 1X SDS sample buffer. Before loading, samples were reduced and denatured by adding 5% β-mercaptoethanol and heating at 95 ºC for 5 min. A molecular weight marker (NEB), run simultaneously on the same gel in an adjacent lane, was used as a standard to establish the apparent molecular weights of proteins resolved on SDS polyacrylamide gels.

**Table 2.8. Compostion of stacking and separating gels for SDS PAGE**

| component | stacking gel 4% | separating gel 8% |
|---|---|---|
| AA/Bis (30 %/0.8 %) | 4.1 ml | 7.2 ml |
| 1 M Tris pH 8.8 | - | 10.2 ml |
| 1 M Tris pH 6.8 | 3.7 ml | - |
| $H_2O$ | 22.3 ml | 9.2 ml |
| 10% SDS | 300 µl | 300 µl |
| 10% APS | 300 µl | 95 µl |
| TEMED | 41 µl | 24 µl |

The molecular weight markers were prepared according to the manufacturer's specifications. Pepsin extracted collagen II (Sigma) was used as an additional molecular standard marker. After loading the samples onto the gel, electrophoresis was performed in 1X running buffer at a constant voltage of 100-150 V until the bromophenol blue dye front had reached the bottom edge of the gel or had just run out of the gel. After electrophoresis, the resolved proteins in the gel were either stained with Coomassie blue, silver nitrate or transferred onto a nitrocellulose membrane for western blotting.

**2X SDS sample buffer**
100 mM Tris-HCl, pH 6.8
4 % v/v SDS
20 % v/v glycerine
0.2 % w/v bromophenol blue

**10X running buffer**:
1.9 M glycine
0.25 M Tris-HCl, pH 8.8
1% SDS

## 2.4.5 Coomassie blue staining of SDS-polyacrylamide gels

The resolved proteins were visualised by staining the gel with Coomassie blue staining solution at room temperature with gentle agitation for at least 60 min. Then, the staining solution was removed and destaining solution was added. The gel was destained at room temperature with gentle agitation with several changes of destaining solution until protein bands were clearly visible.

**Coomassie blue staining solution**
0.1% w/v Coomassie blue R250
50% v/v ethanol
10% v/v acetic acid
filter the solution before use

**Destaining solution**
7% v/v acetic acid
20% v/v ethanol

## 2.4.6 Silver staining of gels

Gels were carefully removed after separation of proteins and fixed by immersing in fixative solution for 15 min with two changes of fixative. To this 50 ml of sensitizer was added and incubated for 20 min. Gels were rinsed in demineralised water three times for 5 min each. 0.1 % silver nitrate solution was added and incubated for 15 min before the gel was quickly rinsed with demineralised water. 40 ml of developer was added and the reaction was stopped by adding the fixative solution as soon as bands appear.

| **Fixative** | **Sensitizer** |
|---|---|
| 250 ml methanol | 150 ml ethanol |
| 25 ml acetic acid | 27.2 g sodium acetate |
| to 500 ml with water | 1.5 ml acetic acid |
| | 10 ml glutaraldehyde (25 %) |
| | 0.5 g sodium thiosulfate pentahydrate |
| | to 500 ml with water |

| **Silver nitrate solution** | **Developer** |
|---|---|
| 0.5 g silver nitrate | 541 µl formalin |
| to 500 ml with water | 10 g sodium carbonate |
| 80 µl of formalin | to 500 ml with water |
| added freshly to every 50 ml solution | |

## 2.4.7 Transfer of proteins to nitrocellulose membranes

Proteins resolved on the gel were transferred electrophoretically to a nitrocellulose membrane in transfer buffer at a constant current of 100 mA overnight at 4°C. After transfer the membrane was stained in 10-15 ml of Ponceau S (Serva) solution for 2-5 min at room temperature to check the transfer of proteins. After staining, the membrane was removed from the Ponceau S solution and rinsed with water to destain until protein bands were visible and the background was clear. The positions of the constituent proteins of the molecular weight marker were marked and the membrane washed with TBS to completely remove the stain.

## 2.4.8 Immunodetection of membrane bound proteins

Nitrocellulose membranes with bound proteins were blocked with 5% low fat milk powder in TBS for 1 h. The membrane was incubated 1 to 2 h at room temperature with primary antibody diluted appropriately in 5% low fat milk powder in TBS followed by three washes of 5 min each in TBST buffer. Appropriately diluted secondary antibody coupled with horseradish peroxidase was added to the membrane and incubated at room temperature for one hour. Finally the membrane was rinsed in TBST buffer three times for 5 min each to remove unbound antibody and the reaction was detected using an chemiluminescence detection system. The light signals were captured by exposure to X-ray film (RX-Super, Fujifilm).

**Table 2.9. List of primary and secondary antibodies used in Western blot**

| Antibody | Dilution | Company |
|---|---|---|
| **Primary antibodies** | | |
| Mouse monoclonal anti-c-myc | 1:1000 | Santa Cruz |
| Rabbit polyclonal anti-c-myc | 1:1000 | Santa Cruz |
| Goat polyclonal anti-collagen II | 1:1000 | Chemicon |
| **Secondary antibodies** | | |
| Peroxidase conjugated anti-goat IgG | 1:1000 | Dako |
| Peroxidase conjugated anti-rabbit IgG | 1:1000 | Dako |
| Peroxidase conjugated anti-mouse IgG | 1:1000 | Dako |

| **Transfer buffer** | **ECL detection** |
|---|---|
| 50 mM boric acid, pH 8.5 | 100 mM Tris-HCl pH 8,5 |
| 10 % methanol | 1.25 mM luminol |
| | 225 nM coumarin |
| | 0.03 % hydrogen peroxide |

## 2.4.9 Trypsin digestion of purified collagen II proteins

10 µl of trypsin solution (100 µg in 10 mM EDTA) was added to 50 µl of eluted protein (50 µg/ml of protein) at pH 8.0 and incubated at 25 $^{o}$C for 2 min. The reaction was stopped by adding 5 µl of soyabean trypsin inhibitor and 50 µl 2X SDS PAGE sample. Samples were boiled at 95 $^{o}$C for 10 min, resolved on 8% SDS PAGE gels and subsequently the proteins detected by silver staining.

## 2.4.10 Measurement of circular dichroism spectra and melting curves

Purified collagen II proteins were dialyzed in 100 mM acetic acid and the concentration was adjusted to 60 µg/ml. 200 µl of this solution was used to measure the spectrum between 190 nm and 280 nm using a Jasco J-715 Polarimeter at 4 $^{o}$C. Melting curves were registered at 222 nm with in a temperature range of 10 °C to 55 °C at increments of 1 °C /min.

### 2.4.11 Negative staining of collagens for electron microscopy

Purified recombinant collagen II proteins and its variants were adsorbed to 400 mesh carbon-coated copper grids, which had been rendered hydrophilic by glow-discharge at low pressure in air. The grids were immediately blotted, washed with two drops of water, and stained with 0.75 % uranyl formate for 15 s. Samples were observed in a Jeol 1200 EX transmission electron microscope operated at 60 kV accelerating voltage and 300000X magnification. These experiments were performed at the University of Lund in cooperation with Dr. M. Mörgelin.

### 2.4.12 Inhibition of proteolytic processing of collagen II using specific inhibitors

Three days after transfection, HT1080 cells in six well plates were washed with medium and fresh medium, supplemented with 100 µg/ml of ascorbate and 50 µM proteasomal inhibitor MG132 (10 mM stock in DMSO) was added. Supernatants were collected 6 h, 12 h and 24 h respectively and resolved on an 8 % SDS-PAGE. Collagen II and fragments thereof were detected by western blot using an antibody directed against the myc epitope.

GM6001 is a universal inhibitor of the MMP pathway. Three days after transfection of HT1080 cells with collagen II constructs, supernatants were collected and the cells rinsed with fresh media. 25 µM GM6001 (1 mM stock in DMSO) was added to the growth medium supplemented with 100 µg/ml of ascorbate. Supernatants and cell lysates were harvested 48 h after addition of the inhibitor and resolved by SDS-PAGE on 8% gels. Collagen II and fragments thereof were detected by western blot using the anti-myc antibody.

### 2.4.13 Mass spectrometry of purified collagen II variants

Purified collagen II samples were reduced with 10 mM DTT at 37 $^{o}$C for 2 h, alkylated with 20 mM idodoacetamide for 1 h at 37 $^{\circ}$C and than reduced again. MALDI-TOF spectra were obtained at service facility of the Centre for Molecular Medicine, University of Cologne, using a Reflex IV Brucker Daltonics Mass spectrometer.

# 3 Results

## 3.1 Subcloning of collagen II cDNA

Collagen II cDNA in the pBS KS+ vector was subcloned into smaller fragments and these fragments were used for site directed mutagenesis to introduce desired point mutations. Attempts to introduce specific point mutations in full length collagen II cDNA were futile due to the frequent introduction of unwanted mutations. Further the sequence analysis was cumbersome due to the high GC content. Fig 3.1 represents a schematic outline of the cloning strategy.

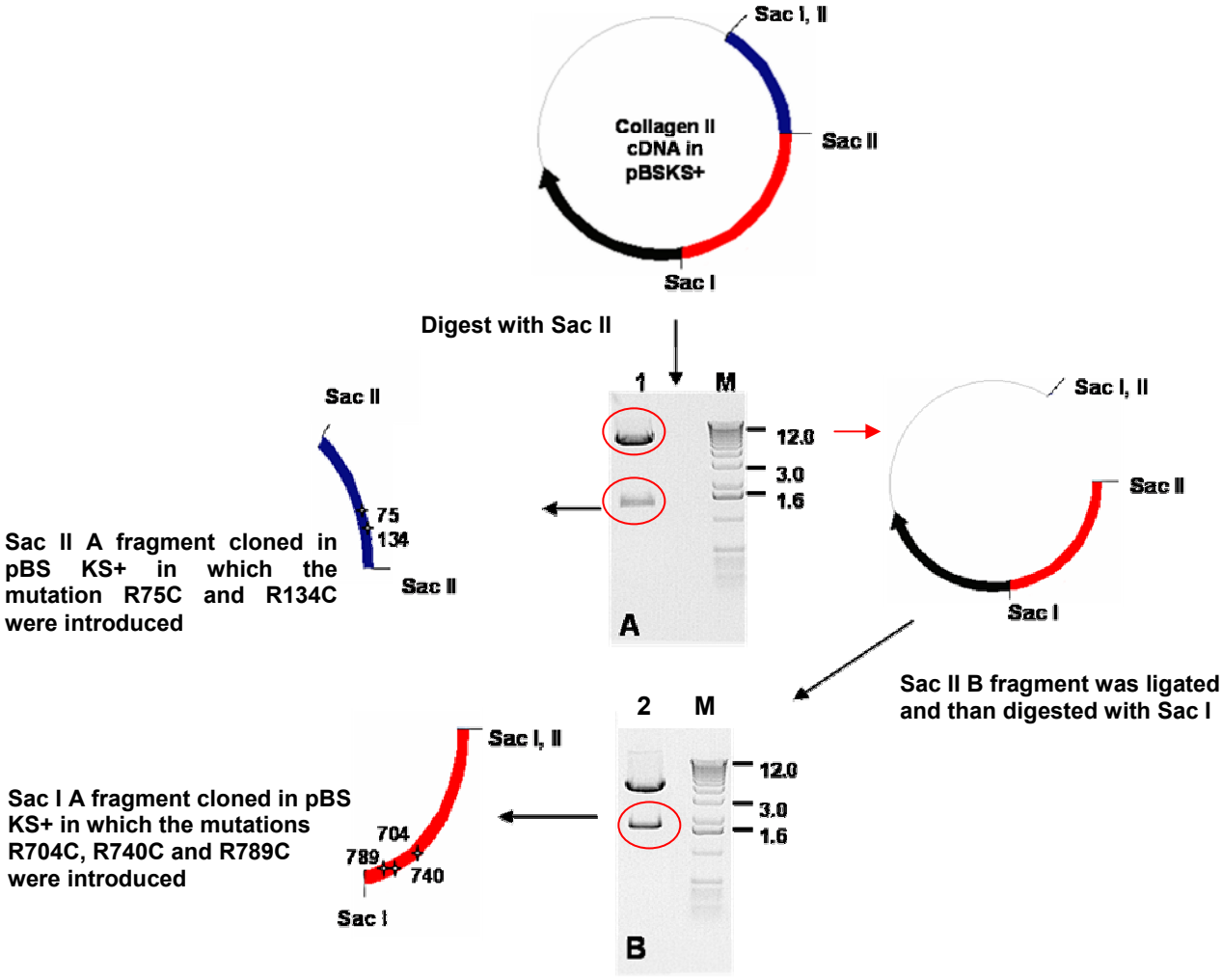


**Fig 3.1 The subcloning scheme for site directed mutagenesis of collagen II cDNA.** (A) represents a 1% agarose gel of DNA separated after digestion of collagen II cDNA in pBS KS+ with Sac II (Lane 1). Two fragments were obtained, the Sac II A fragment was cloned in pBS KS+ vector and used for introducing mutations leading to R75C and R134C. The fragment Sac II B was ligated back at the Sac II position and further digested with Sac I. (B) shows fragments after digestion of Sac II B with Sac I (lane 2). The Sac I A fragment was used for introducing mutations leading to R704C, R740C and R789C. After introducing the desired mutation the inserts was cloned back into the backbone vector in the same order. M, 1 kb DNA standard marker.

After subcloning of the smaller fragment, the point mutations R75C and R134C were introduced in the Sac II A fragment and the mutations R704C, R740C and R789C in the Sac I A fragment. After sequence analysis to confirm the mutation of interest the respective fragments were digested from their vectors and cloned back into the original backbone vector to give full length cDNA with the desired mutation. All the cDNAs were sequenced to rule out unwanted mutations.

## 3.2 Cloning of collagen II cDNA into pSL1180 vector

The endogenous signal peptide in the collagen II cDNA had to be removed since the mammalian expression vector pCEP-Pu already contains the BM 40 signal peptide. Hence it was necessary to clone the collagen II cDNA constructs from the pBS KS+ vector to the pSL1180 vector. After mutagenesis, the collagen II cDNA was excised from the pBS KS+ vector by digestion with Spe I and Eco RV and then ligated into pSL1180 which was digested with the same enzymes and dephosphorylated (Fig 3.2).

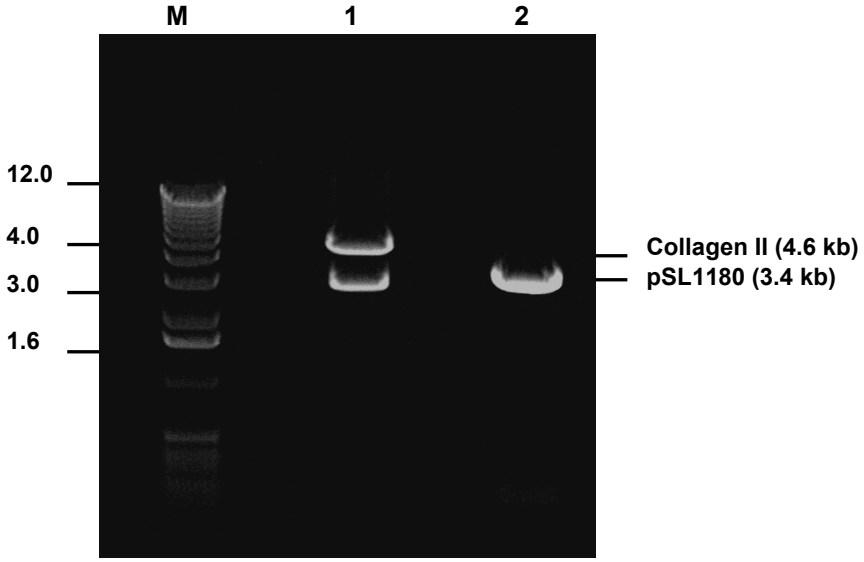


**Fig. 3.2 Cloning of collagen II cDNA from the pBS KS+ vector into the pSL1180 vector for deletion of the endogenous signal peptide.**A 1% agarose gel showing restriction digestion products after digestion with Spe I and Eco RV of collagen II cDNA in the pBS KS+ vector (lane 1) and the pSL1180 vector (lane 2). M represents the DNA standard marker.

## 3.3 Cloning of collagen II cDNA without signal peptide into the mammalian expression vector pCEP-Pu

Collagen II cDNA without endogenous signal peptide was excised from the pSL1180 vector at the Nhe I and Bam HI positions and ligated into the pCEP-Pu vector which had been digested with Nhe I and Bam HI. Fig 3.3 shows all the collagen II variants after digestion of the pCEP-Pu vector.

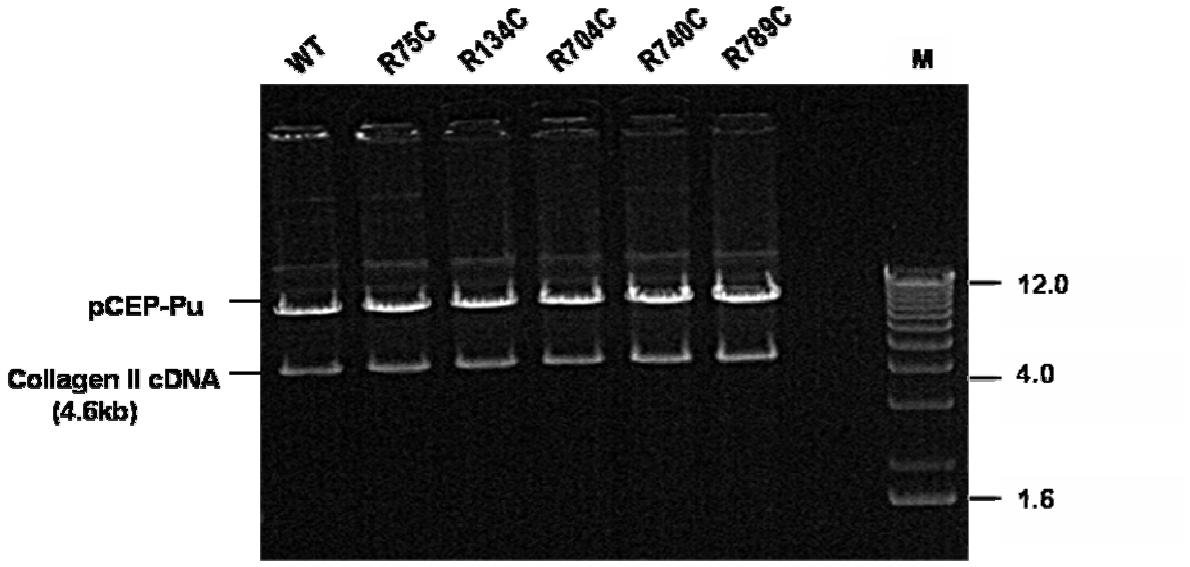


**Fig 3.3 Restriction digestion of all collagen II constructs from pCEP-Pu**.DNA was separated on a 1% agarose gel after restriction digestion with Nhe I and Bam HI of the collagen II cDNA cloned in the mammalian expression vector pCEP-Pu. M, DNA standard marker.

## 3.4 Transient expression of collagen II proteins in 293 EBNA cells

Supernatants and cell lysates from 293 EBNA cells transfected with collagen II constructs were reduced with mercaptoethanol and proteins separated on a 8% SDS PAGE gel (Fig 3.4). Recombinantly expressed collagen II was in all cases detected in the supernatants of transfected 293 EBNA cells. Significantly higher amounts of proteins were seen in the cell lysates of R740C and R789C transfected cells as compared with cells expressing other collagen II variants. In addition, the mobility of the R789C protein was altered as compared with other collagens. Collagens in the cell lysates migrate slightly higher than those in the supernatants.

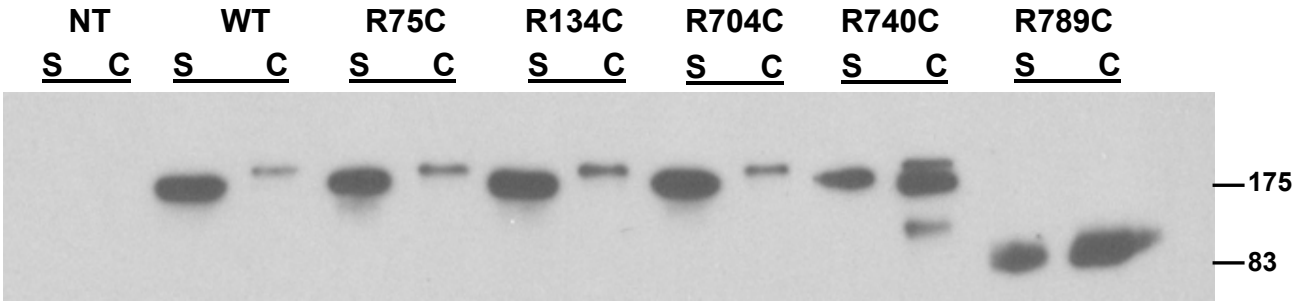


**Fig 3.4 Western blot analysis of supernatants and cell extracts of transfected 293 EBNA cells.** 293 EBNA cells transiently transfected with different collagen II constructs were analysed by western blot for expression of collagen II proteins. Supernatants (S) and cell lysates (C) harvested three days post transfection were separated by SDS PAGE (8% gel) under reducing conditions and the blot was developed with an antibody directed against the myc epitope. NT, non transfected; WT, wild type; R75C, R134C, R134, R704C, R740C and R789C are cells transfected with the respective mutated collagen II variants.

## 3.5 Purification of recombinantly expressed collagen II

To analyze the effect of these mutations by physicochemical and structural methods, proteins were purified from the media of 293 EBNA cells. Supernatants were harvested from 293 EBNA cells expressing $his_6$-myc-tagged recombinant collagen II constructs and these were subjected to affinity chromatography using a nickel-NTA column. The bound protein was eluted with increasing imidazole concentrations (up to 250 mM). Fractions containing collagen II variants were separated by SDS PAGE and the gels subjected to Coomassie blue staining. Fig 3.5 shows respresentative gels from the purification of the wildtype and R789C constructs.

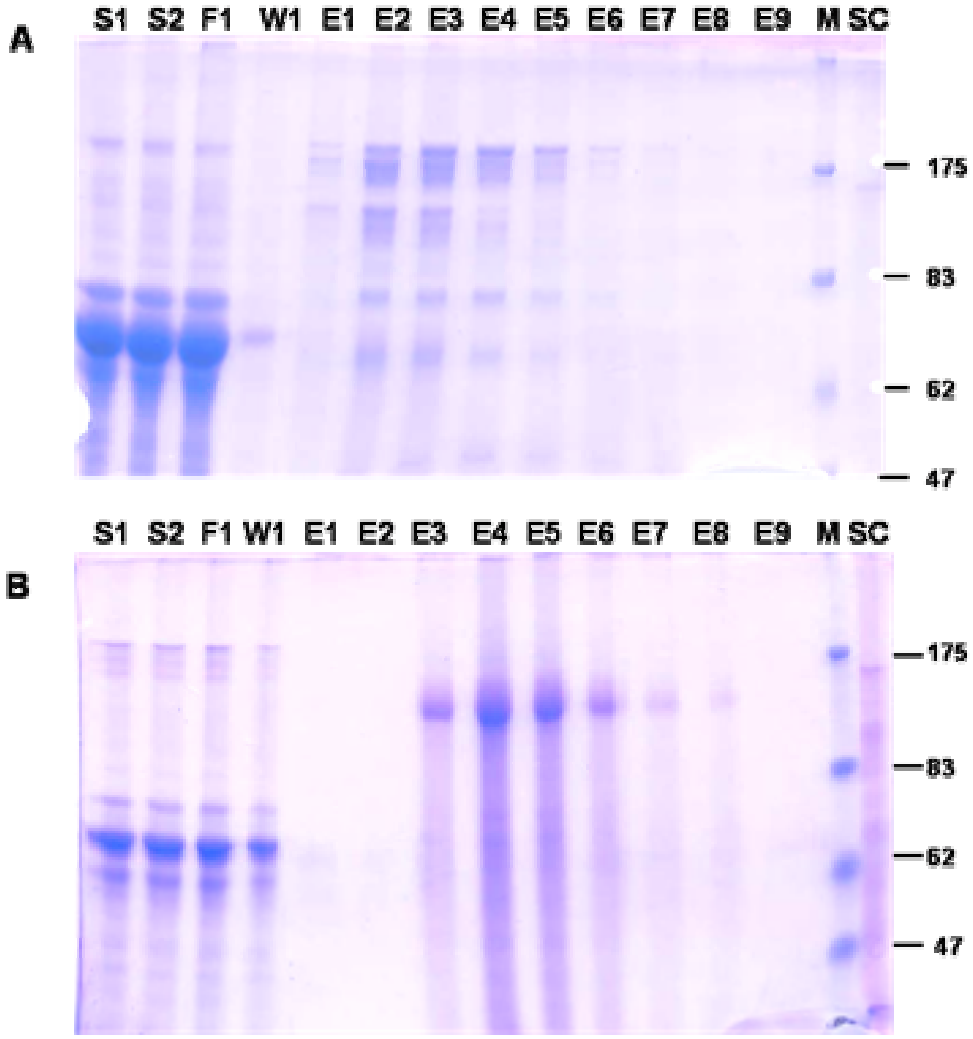


**Fig 3.5 Protein fractions from various stages of collagen purification resolved on 8% SDS PAGE gels stained with Coommassie blue.** Panel A depicts a Coomassie stained gel of fractions from wild type collagen II and panel B shows the corresponding gel for R789C mutant collagen.S1, S2, cell culture supernatants; F1, flow through; W1, wash; E1 to E9, elution fractions 1 to 9; M, marker; SC, pepsin extracted collagen II (Sigma)

## 3.6 Coomassie staining of purified collagen II variants

The concentrations of the purified collagen II proteins were measured and 1 µg of each variant was resolved on an 8% SDS PAGE gel and stained with Coomassie blue as shown in the Fig 3.6. Trimeric bands representing collagen II (with both the propeptides, with one propeptide and fully processed) as indicated by arrow marks were visible in samples of WT, R75C, R134C, R704C and R740C collagens. In the R789C sample a distinct single band migrating lower than all other variants, including the pepsin extracted collagen, was observed.

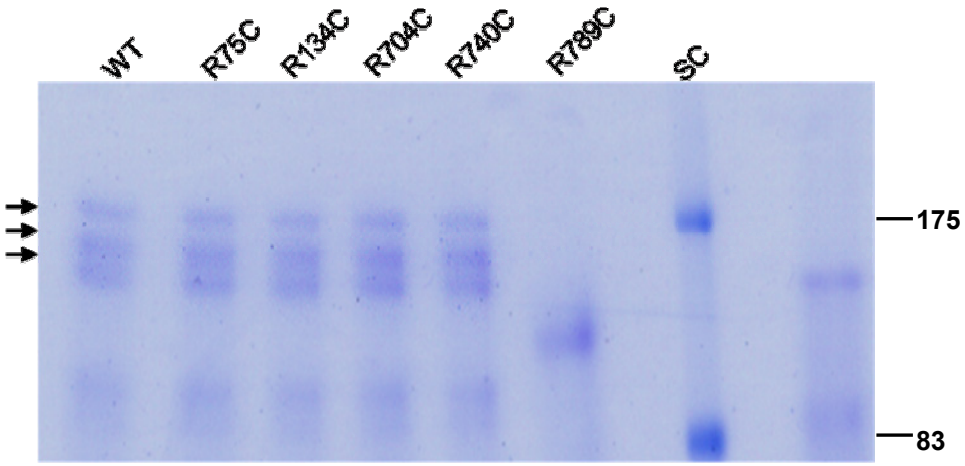


**Fig 3.6 Coommassie blue stained SDS-PAGE gel of all the purified collagen II variants.** Triple bands as indicated by arrows represent the WT, R75C, R134C, R704C and R740C collagen in contrast to a single band migrating lower for R789C. M, protein standard; SC, pepsin extracted collagen II (Sigma).

## 3.7 Immunoblot of purified collagen II variants

The purified collagen II proteins were analysed by western blot using antibodies specific for collagen II and the N-terminal tag, respectively. With both antibodies two bands were detected, with the upper one being more prominent when detected with antibody directed against the myc epitope (Fig 3.7 A) and the lower band more intense when detected with anti-collagen II (Fig 3.7 B). The anti-myc antibody detects the upper band which most probably corresponds to the full length procollagen II. Also the purified R789C migrated at a lower position than the other collagen II variants and the protein was intact at the N-terminal end as it could be detected with the anti-myc antibody.

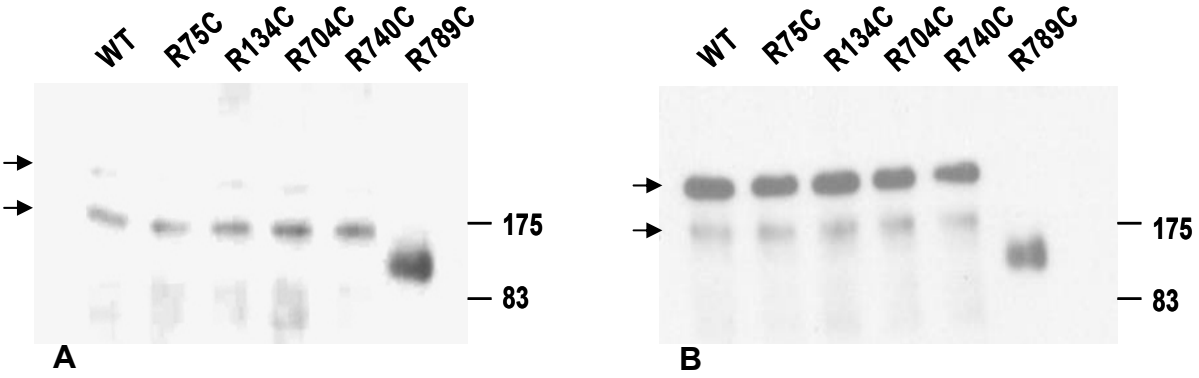


**Fig 3.7 Western blot analysis of all the purified collagen II variants.** Purified recombinant proteins were resolved on 8% SDS PAGE gels and the blots developed using antibodies against collagen II (A) and the myc epitope (B), respectively.

## 3.8 Trypsin digestion of different collagen II variants

To determine the triple helical stability of different collagen II variants, these were subjected to trypsin digestion. A single collagen II polypeptide has 135 potential trypsin cleavage sites. In the case of collagen II forming stable triple helical trimers, these sites are not exposed and therefore not accessible to trypsin. Only the N- and C-propeptides are cleaved. Fig 3.8 shows silver stained gels of all collagen II variants before and after trypsin digestion. WT, R75C, R134C and R704C collagens were trypsin resistant, indicating that the secreted protein has formed a stable triple helix. R740C and R789C collagens were susceptible to trypsin digestion and completely degraded. This implies that the secreted R740C and R789C proteins are not forming stable triple helical structures.

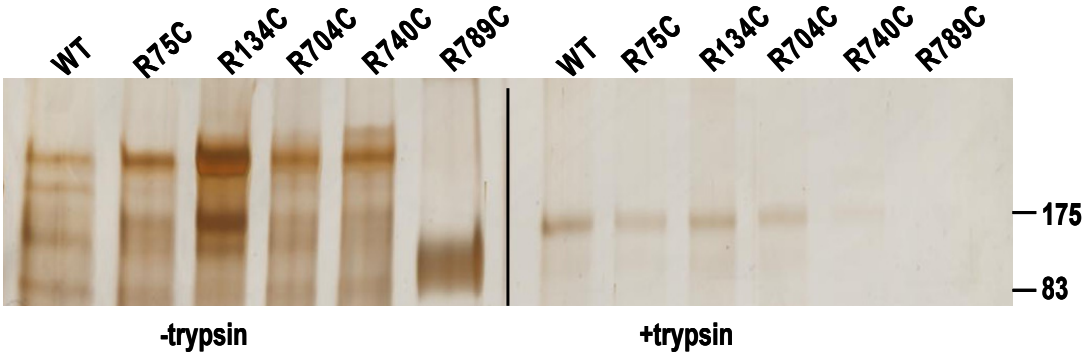


**Fig 3.8 Silver stained SDS-PAGE gel after trypsin digestion of the purified collagen II variants to analyse their triple helical integrity.** Collagen II variants were tested by digestion with trypsin at 25 °C for 2 minutes followed by SDS-PAGE. The collagens were visualized by silver staining. Left panel; collagen variants incubated without trypsin, and right panel; with trypsin. R740C and R789C collagens were extensively degraded, whereas for all the other constructs a single band of the size of the triple helical domain remained.

## 3.9 Circular dichroism of purified collagen II proteins

The folding of triple helical collagens can be analysed by CD spectroscopy. At 222 nm a typical collagen spectrum shows a positive ellipticity indicating a triple helical structure. In case of a unstable structure, e.g. as a result of low hydroxylation, the ellipticity shifts below zero at 222 nm. The CD spectra of the purified collagen II variants are shown in Fig 3.9. WT, R75C, R134C, and R704C proteins displayed normal collagen II spectra indicating a high triple helical content and a stable overall structure. The structure of the collagen II mutants R740C and R789C was clearly affected when compared to wild type collagen and the spectra indicate a decreased triple helical content in these proteins.

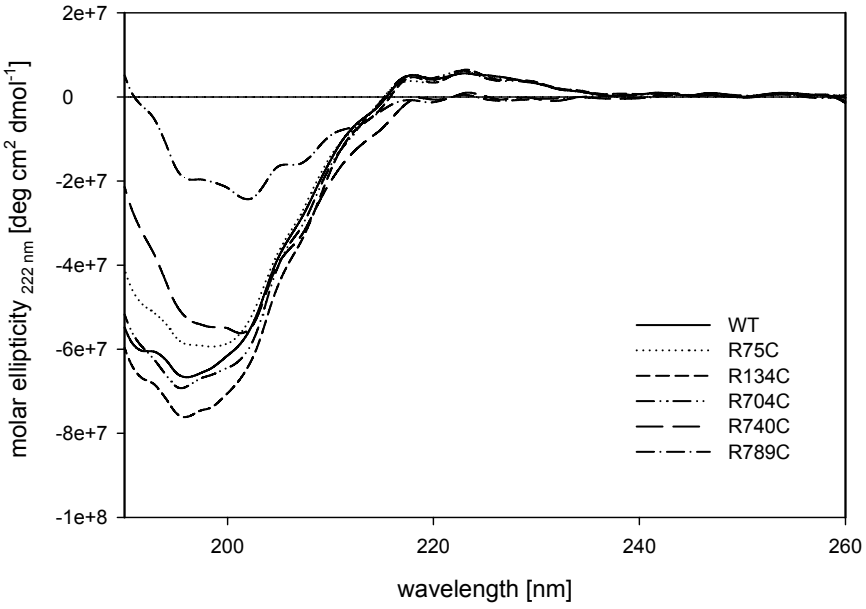


**Fig 3.9 CD spectra of purified collagen II variants.** Spectra were recorded using collagens at a concentration of 60 µg/ml after dialysis against 100 mM acetic acid. The structure of the mutant collagen II proteins R740C and R789C were altered as compared with other collagen II variants and the shape of the spectra indicate a decrease in triple helical structure.

## 3.10 Melting curves of purified recombinant collagen II proteins

Purified collagens were subjected to incremental heating with simultaneous measurement of molar ellipticity at 222 nm. The temperature where 50% of the triple helical structure is lost is defined as the melting temperature (Tm) of the respective collagen. Tm depends on the stability of collagen triple helix, the more stable the triple helix is the higher the Tm. The melting curves of all the purified collagen II proteins are shown in Fig 3.10A. The percentage of folded collagens as a function of temperature was determined by using a curve fit model assuming 100 % initially folded molecules (Fig 3.10B). The curves for WT, R75C, R134C and R704C

collagens show similar profiles while R740C and R789C proteins showed a thermal instability as compared to the other collagens. The Tm values are summarized in Table 3.1.

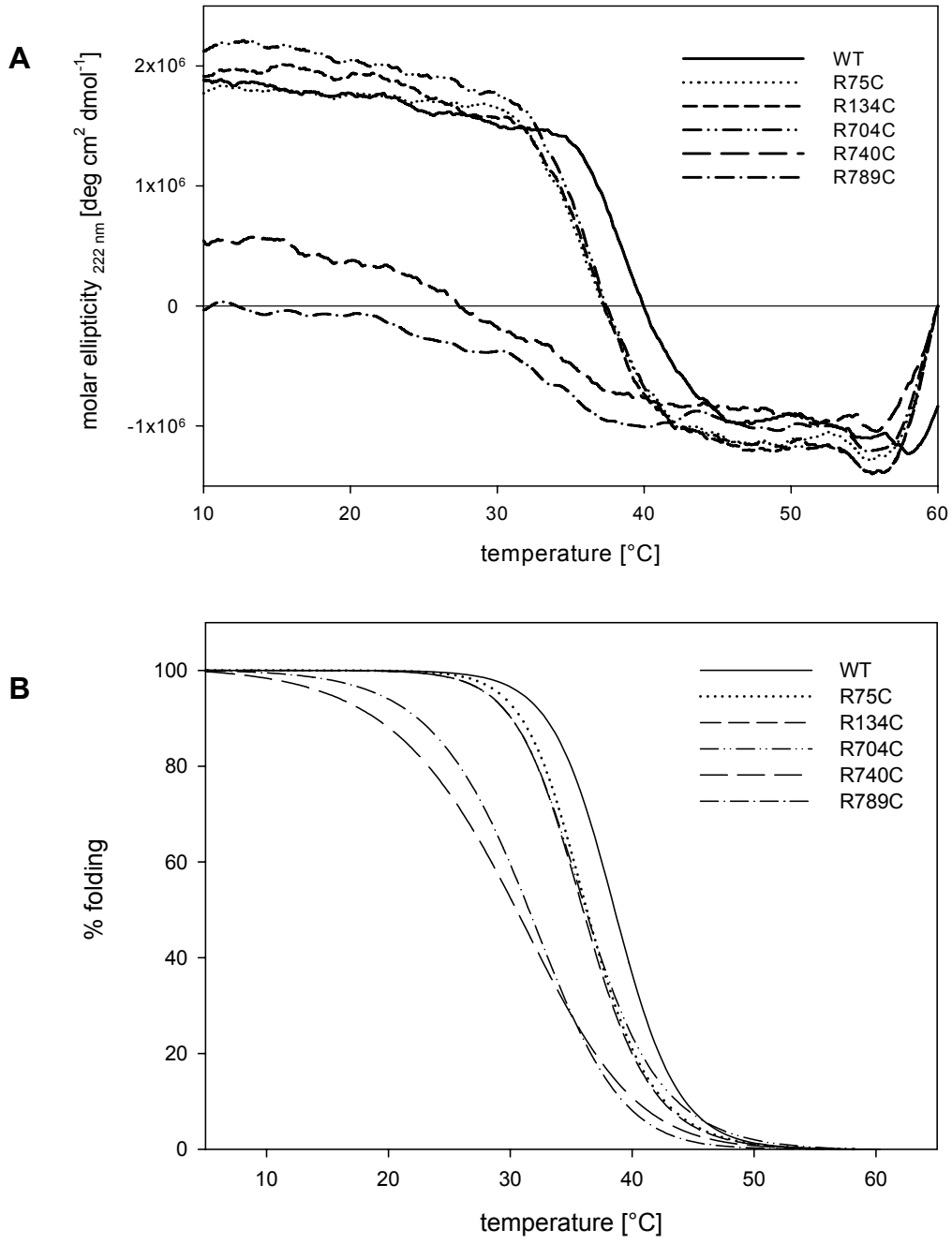


**Fig 3.10 Melting curves of purified collagen II variants.** (A) CD melting curves of collagen II proteins (60 µg/ml) recorded after dialysis into 100 mM acetic acid at 222 nm with a 1 °C/min temperature gradient from 10 to 55 °C. (B) The percentage folding using a curve fit model assuming that initially 100 % of the collagen molecules were folded. The R740C and R789C proteins displayed a decreased thermal stability when compared to all other collagen II variants.

**Table 3.1 Melting points (Tm) of the purified collagen II variants**

| sample | melting point (Tm) |
|---|---|
| WT | 38.6 °C |
| R75C | 36.2 °C |
| R134C | 36.1 °C |
| R704C | 36.1 °C |
| R740C | 30.2 °C |
| R789C | 31.5 °C |

## 3.11 Mass spectrometric analysis of purified collagen II proteins

Purified collagen II proteins were analysed by mass spectrometry to determine the masses of the intact protein. Fig 3.11 shows the combined mass spectra for the wildtype, R789C and pepsin extracted collagens (SC). Wild type collagen showed a mass of approximately 150.6 kD. This value was also obtained for R75C, R134C, R704C and R740C proteins (data not shown). In case of R789C collagen the mass of the processed protein was around 92.9 kD (Fig 3.11), lower than the mass obtained for pepsin extracted collagen which lacks propeptides (123.9 kD). Table 3.2 compares the theoretical with the measured mass for all the collagen II variants.

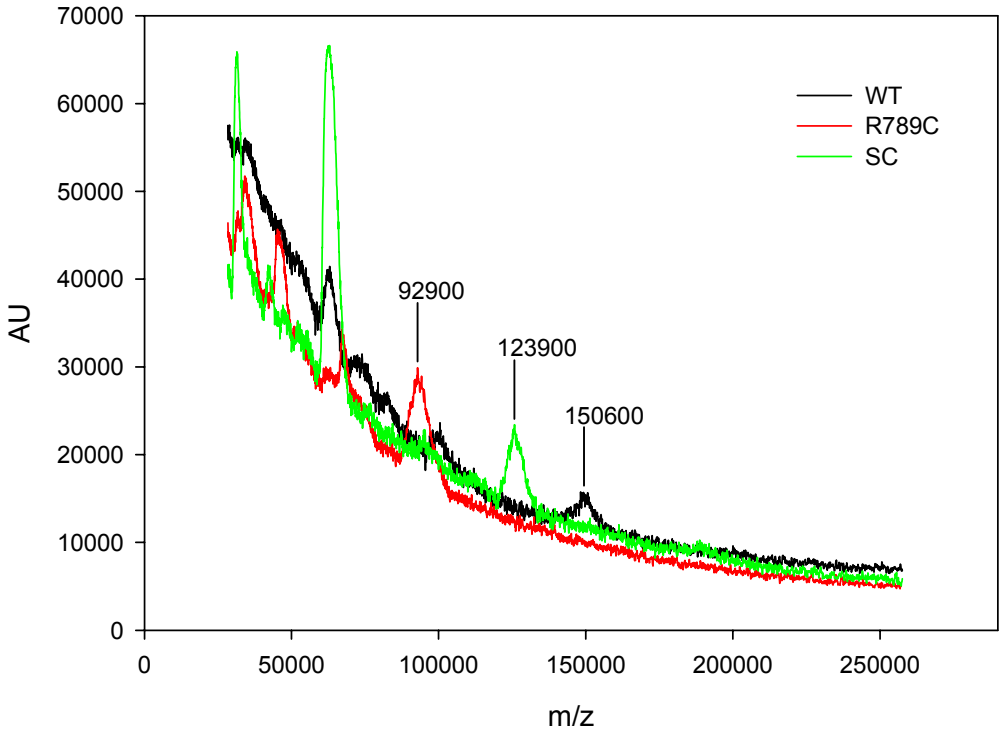


**Fig 3.11 MALDI-TOF spectra of purified collagen II proteins.** Wild type collagen II protein (black) gave a measured mass of 150.6 kD which includes the mass of the $his_6$ myc tag. The mass of the R789C protein (red) was

approximately 92.9 kD. Pepsin extracted collagen from the bovine nasal cartilage, which lacks both propeptides, displayed a mass of 123.9 kD (green).

**Table 3.2 Comparison of theoretical and measured mass of collagen II variants**

| sample | mass from MALDI-TOF | theoretical mass* |
|---|---|---|
| WT | 150.6 kD | 136.8 kD |
| R789C | 92.9 kD | 136.8 kD |
| SC[a] | 123.9 kD | 96 kD |

*Theoretical mass include the propeptides and tag; [a] pepsin extracted collagen (Sigma)

## 3.12 Negative electron microscopy of purified collagen II proteins

Purified collagen II proteins were analysed by electron microscopy after negative staining. Fig 3.12 shows representative electron micrographs of WT, R75C R740C and R789C collagens. Characteristic filaments were observed in the WT and R75C samples. R134C and R704C collagens displayed similar structures (data not shown). In the R740C and R789C samples the characteristic filaments were not seen and the short proteins detected were significantly thinner and more sparse. In addition, R75C collagen showed kinks (arrow mark) close to the end of the structure. Because of the low contrast in the electron micrographs, selected filaments were highlighted to illustrate representative structures.

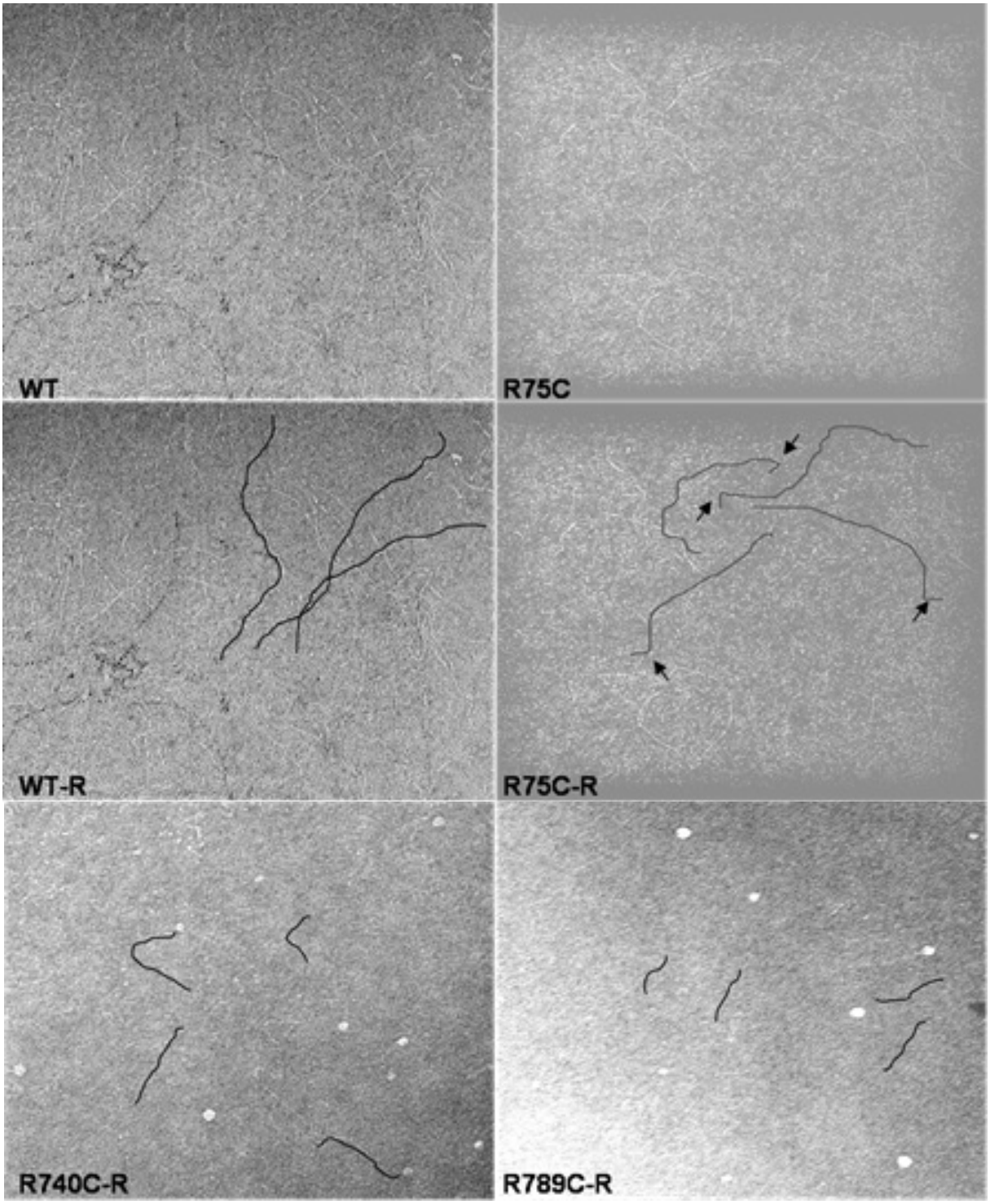


**Fig 3.12 Electron micrographs of purified and negatively stained collagen II variants.** In the wildtype and R75C samples characteristic collagen filaments are seen (redrawn for WT-R and R75C-R). In the R740C-R and R789C-R panels these filaments, if at all present, all are thinner and hardly detectable. In R75C collagen kinks were consistently seen (arrows).

## 3.13 HT 1080 cells transiently transfected with collagen II constructs

HT 1080 is a fibroblast derived cell line and was used to study the intracellular trafficking of wild type and mutant collagen II. Three days after transfection with collagen II constructs, supernatants and cell lysates were analysed. Fig 3.13 shows a western blot developed with antibodies directed against the myc epitope. Collagen II was detected exclusively in the supernatants of WT, R75C, R134C and R704C cultures. In the R740C culture, protein was found in cell lysate and in about equal amounts in the supernatant. In the case of the R789C mutation, the protein was almost exclusively detected in the cell lysate indicating significant intracellular retention. The R789C collagen mobility was shifted as compared to other collagen II proteins in a manner similar to that seen with protein from 293 EBNA cells.

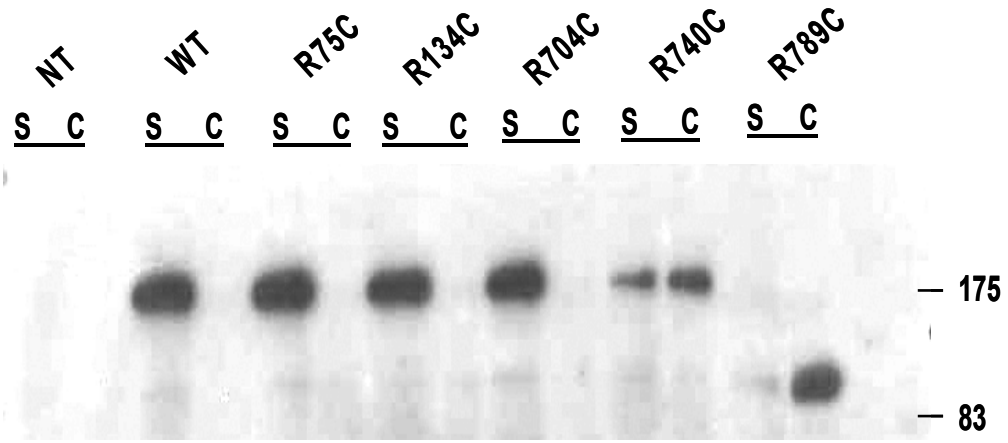


**Fig 3.13 Western blot analysis of collagen II variants expressed in HT1080 cells.** Supernatants (S) and cell lysates (C) from HT1080 cells were analysed three days post transfection with antibodies directed against the myc epitope. Significant amounts of R740C and R789C protein were detected in the cell lysates and only a minor portion was secreted into the supernatants.

## 3.14 Effect of the MMP inhibitor GM6001 on HT1080 cells expressing collagen II

The extracellular degradation of the fibrillar collagens type I, II, and III can occur both at nonhelical sites (Liu et al. 1995) and through a triple helix cleavage. This is achieved by collagenases, which belong to the family of zinc-dependent endopeptidases called matrix metalloproteinases (MMPs). Collagenase-1 or MMP-1 (interstitial collagenase), collagenase-2 or MMP-8 (neutrophil collagenase), and the more recently cloned and characterized collagenase-3 or MMP-13 (Freije et al. 1994 and Knauper et al. 1996) are the only mammalian enzymes known to be able to initiate the intrahelical cleavage of triple helical collagen at neutral pH. These three

collagenases have the ability to cleave the fibrillar collagens type I, II, and III at a single site (Gly 775–Leu/Ile 776) within each chain of the triple helical collagen molecule.

To investigate if the shift in R789C protein mobility is due to MMP cleavage, GM6001, an universal inhibitor for MMPs, was used. Cells expressing wildtype and R789C collagen were incubated for two days with the inhibitor. Supernatants and cell lysates from these cells were analysed by western blot and detected with an antibody directed against myc epitope (Fig 3.14). In the presence of GM6001 significant amounts of R789C proteins remained uncleaved, in contrast to what was seen in the absence of the inhibitor. This indicates that the mutation R789C is aiding MMPs to cleave collagen II and thereby causes the shift in the mobility when compared with wild type protein. In the presence of GM6001, R789C protein can also be detected in the supernatant showing that inhibition of cleavage results in increased secretion.

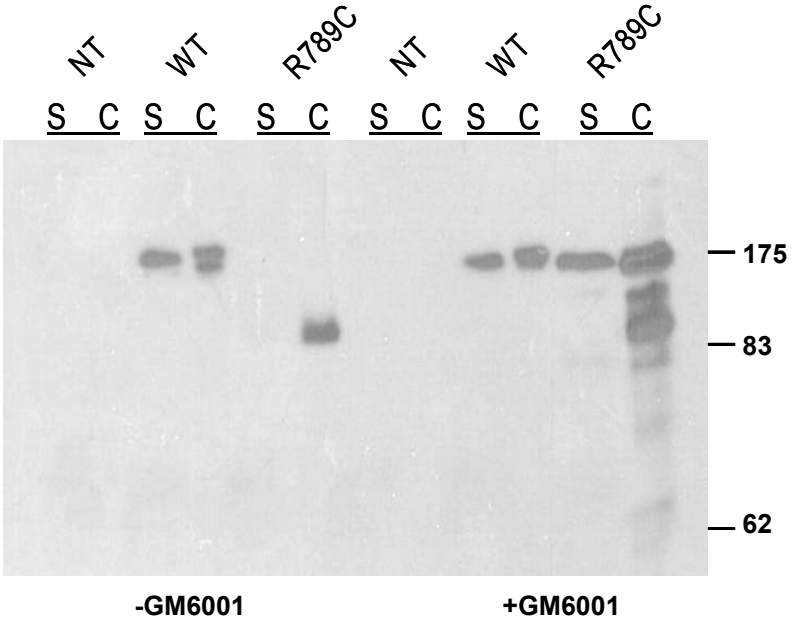


**Fig 3.14 Analysis of proteolytic processing in the absence (left) or presence (right) of the general MMP inhibitor GM6001 by western blot.** Non transfected cells (NT) and cells expressing wildtype (WT) and R789C collagens were treated for two days with GM6001 (25μM). Supernatants (s) and cell lysates (c) were harvested and analysed by western blot. After treatment with GM6001 the R789C collagen was partially protected from degradation, indicating that the reduced mass or shift in mobility is due to a proteolytic cleavage by MMPs.

## 3.15 Effect of proteasome inhibitor MG132 on processing of R789C

Endoplasmic reticulum (ER)-associated protein degradation (ERAD) eliminates misfolded or unassembled proteins from the ER. ERAD targets are selected by a quality control system within the ER lumen and are ultimately degraded by the cytoplasmic ubiquitin-proteasome system (UPS) (Meusser et al 2005).

MG132 is a universal inhibitor used to block the proteasomal degradation pathway. To determine if the cleavage collagen occurs at proteasomes. HT1080 cells expressing wildtype and R789C collagens were treated with MG132 for 6 h, 12 h and 24 h. Supernatants from these cells were analysed by western blotting and collagens were detected with an antibody directed against the myc epitope (Fig 3.15). Treatment with MG132 did not affect the mobility of R789C collagen when compared with non treated R789C samples. This indicates that the cleavage of the R789C mutant does not involve proteasomes.

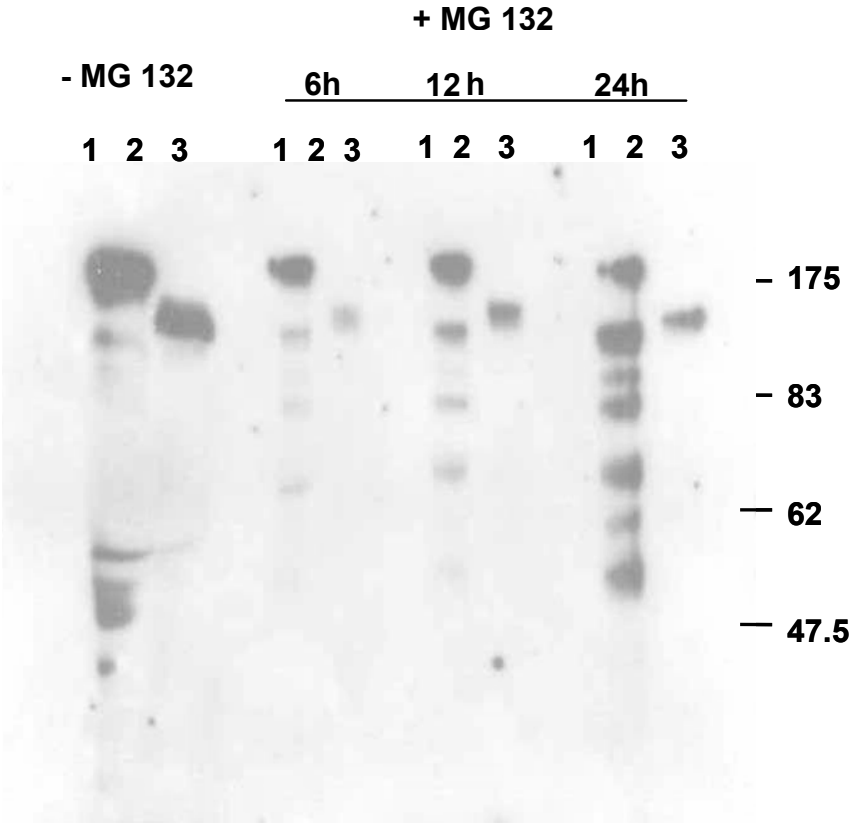


**Fig 3.15 Effect of the proteasomal inhibitor MG132 on R789C protein expression.** Supernatants from cells treated for 6 h, 12 h and 24 h with MG132 (50 µM) were analysed with an anti-myc antibody. Lane 1, non transfected (NT); lane 2, wildtype collagen II (WT) and lane 3, mutant collagen R789C (R789C). MG132 did not show any affect on the collagen II processing in HT1080 cells when compared to untreated control cells (-MG132).

## 3.17 ER stress response in collagen expressing HT1080 cells

XBP-1 is a chaperone involved in the endoplasmic reticulum stress response occurring when misfolded proteins are retained in the ER and is used as a marker for this response. It has been shown previously that XBP-1 is specifically spliced as a consequence of ER stress. Therefore RNA of transfected HT1080 cells was isolated and XBP-1 splicing was investigated by RT-PCR using specific primers giving rise to a 248 bp band in the case of unspliced XBP-1 and to a 222 bp band where XBP-1 is spliced. Fig 3.17 shows the PCR products for XBP-1. In non-

transfected cells as well as cells transfected with wildtype, R75C, R134C and R704C constructs only the unspliced form of XBP-1 was detected whereas in cells transfected with R740C and R789C the spliced variant (XBP-S) was also present. This indicates that ER stress response is activated in these cells due to the mutation.

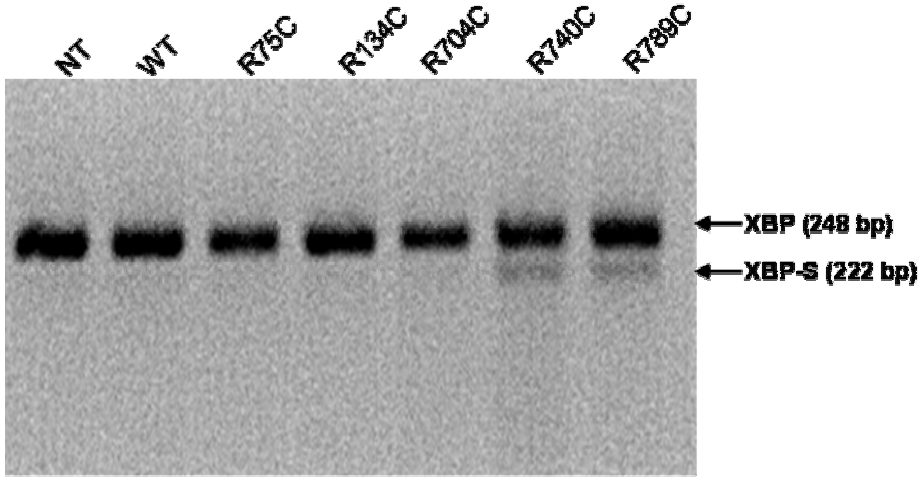


**Fig 3.16 Analysis of stress induced XBP-1 splicing by RT-PCR.** mRNA isolated from HT1080 cells was used for RT-PCR with primers specific for XBP-1. The PCR product was analysed by agarose gel electrophoresis on 2.5 % gels. XBP-1 mRNA could be detected in all cells. Frame switch splicing of XBP-1 (XBP-S) due to ER stress was detected in cells transfected with R740C and R789C constructs.

## 3.18 Immunofluorescence and light microscopy of HT1080 cells expressing collagen II variants

The morphology of transfected HT1080 cells was visualized by light microscopy (Fig 3.17). To analyze the intracellular localization of the transfected collagen II variants in HT1080 cells, immunofluorescence staining was performed. In addition to the staining for the myc tag, a costaining was performed using antibodies directed against compartment specific proteins, e.g 58K for the Golgi apparatus and PDI for the endoplasmic reticulum (ER). Using these antibodies the trafficking of collagen from endoplasmic reticulum to the Golgi apparatus can be followed and a potential retention or secretion defect revealed.

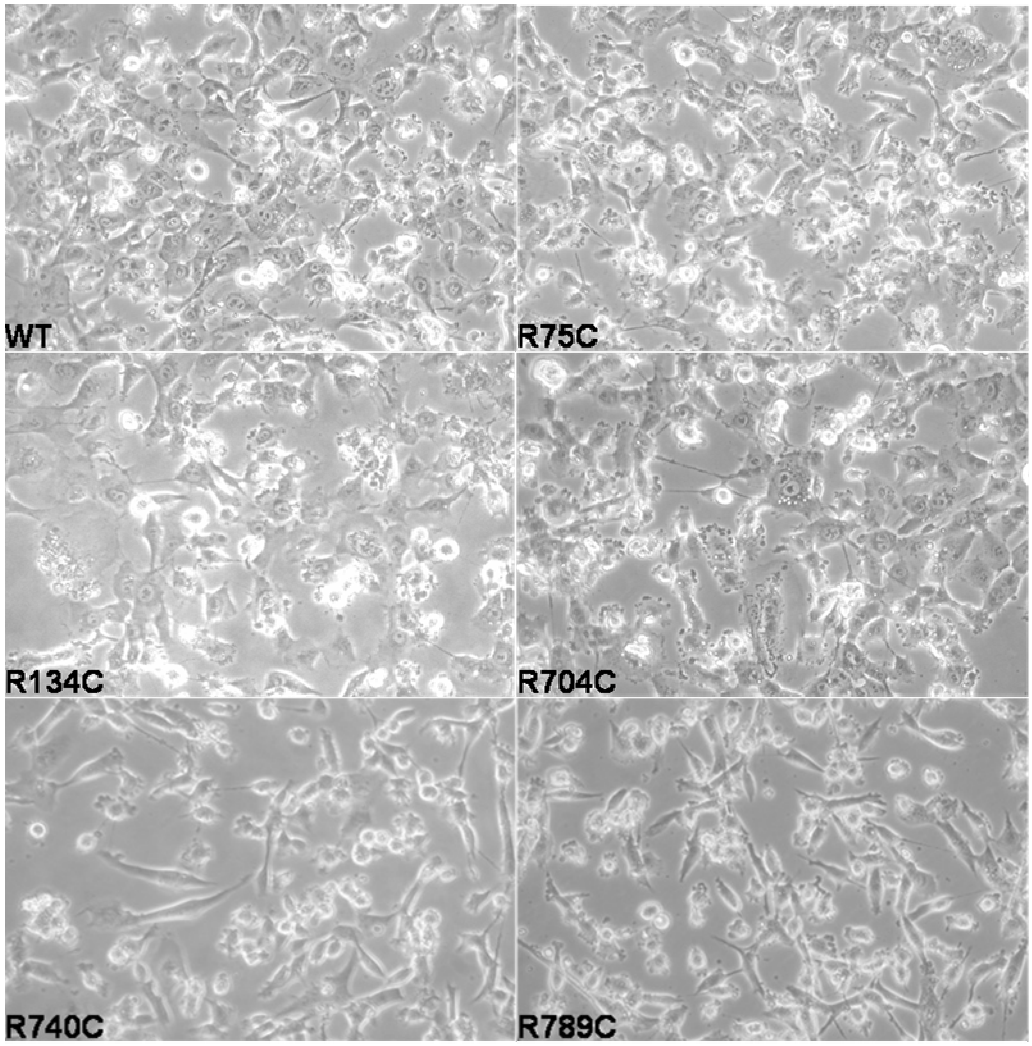


**Fig 3.17 Phase contrast images of HT1080 cells after transfection with indicated collagen II constructs.** Cells transfected with R740C and R789C constructs were significantly more rounded and spindle shaped as compared with other cells.

The non-transfected cells do not show any expression of collagen II, but show the characteristic halfmoon shaped perinuclear staining of the Golgi compartment. In cells transfected with wild type, R75C, R134C and R704C constructs, expression of collagen II (green) was observed and this was significantly colocalised with the signal for the Golgi apparatus, indicating that most of the protein has reached this compartment and is in transit for secretion. In cells transfected with the constructs R740C and R789C significant amounts of protein were detected in other intracellular compartments leading to a separate green and red signal (Fig 3.18). Less protein was seen colocalised with the Golgi marker, further the morphology of cells expressing R740C and R789C collagens was severely changed.

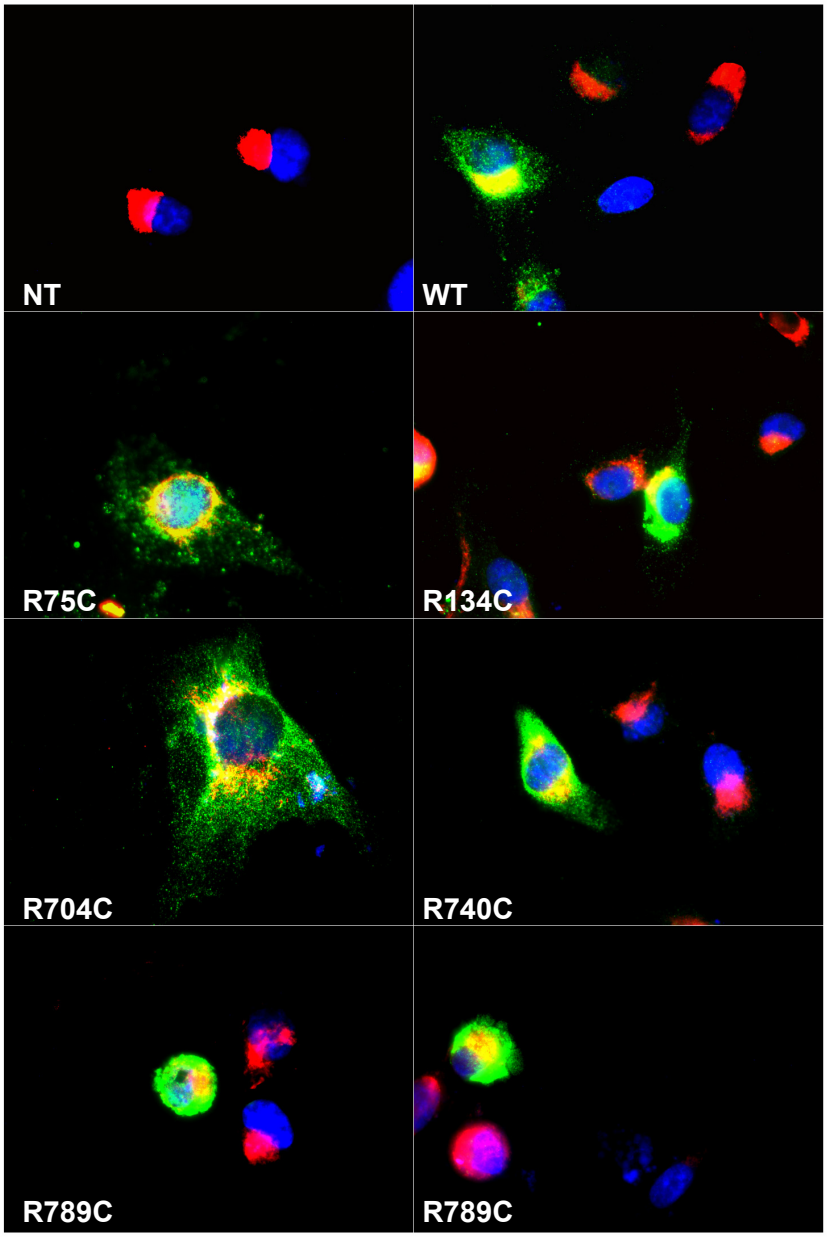


**Fig 3.18 Costaining of HT1080 cells with antibodies against collagen II and the 58K Golgi marker**. HT1080 cells transfected with collagen II constructs were analysed three days post transfection with antibodies directed against the Golgi apparatus (red) and collagen II (green). Nuclei (blue) were counterstained with bisbenzimide. Colocalisation of collagen with 58K in the Golgi compartment is seen in cells transfected with WT, R75C, R134C and R704C constructs. In cells expressing R740C and R789C collagens significant amounts of protein are detected outside the Golgi apparatus.

In addition, transfected HT1080 cells were analysed by immunofluorescence using an antibody directed against protein disulfide isomerase (PDI), an ER specific enzyme. In non-transfected HT1080 cells, the ER was distributed all over the cells. Fig 3.19 shows that WT, R75C, R134C, R704C collagen expressing cells display a clear difference between the two stainings and little colocalisation between the ER marker and collagen II proteins was seen. In contrast, in R740C

and R789C expressing cells a significant colocalisation of the ER marker and collagen II was observed, whereas a characteristic staining of the Golgi apparatus did not appear.

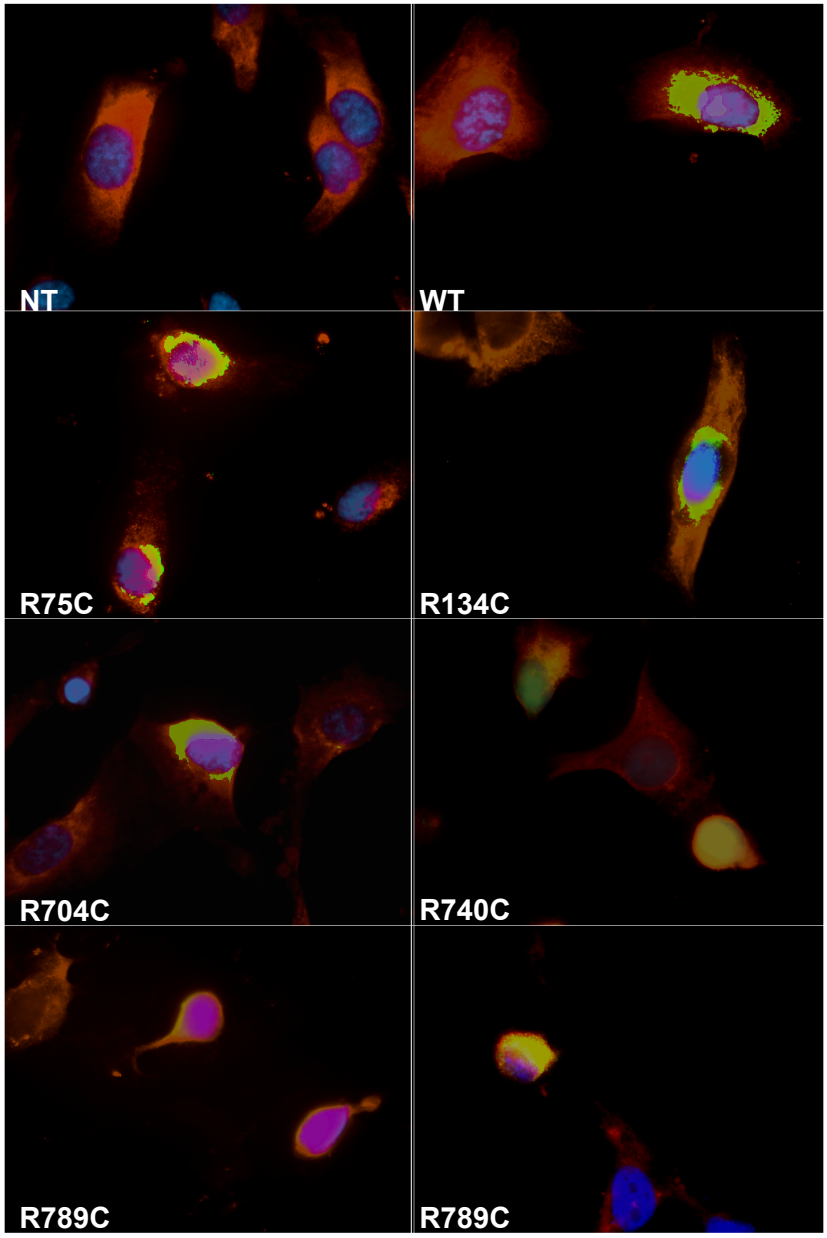


**Fig 3.19 Analysis of transfected HT1080 cells by costaining for PDI and collagen II**. Three days after transfection with collagen II variants, HT1080 cells were analysed by immunofluorescence staining using antibodies directed against collagen II (green) and PDI (red). Nuclei were counterstained with bisbenzimide (blue). Non-transfected cells show staining for PDI distributed across the cells. Cells transfected with WT, R75C, R134C, and R704C constructs displayed separate distinct staining for the ER and the expressed collagens with much less colocalisation. Significant colocalisation of PDI with collagen II staining was seen only in cells expressing R740C and R789C and these cells were also rounded in appearance.

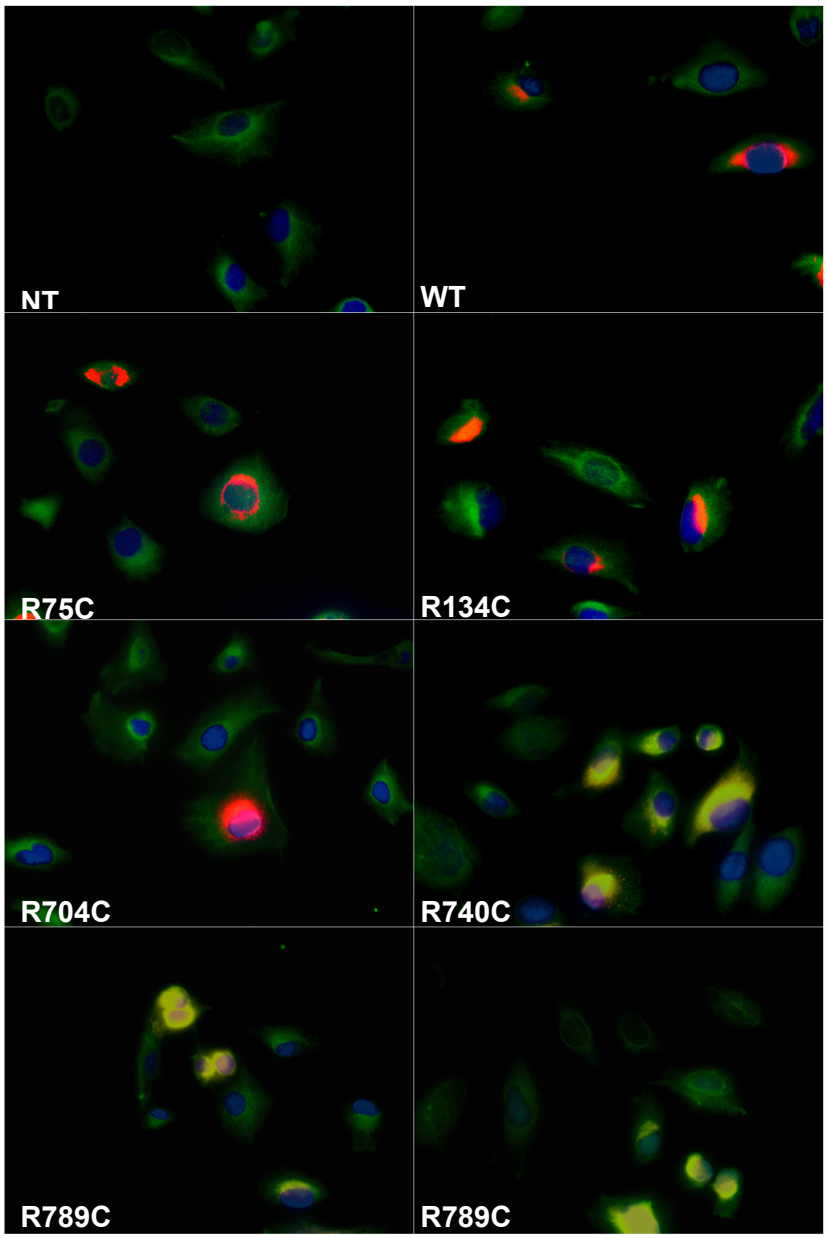


**Fig 3.20 Analysis of BiP expression by immunofluorescence microscopy of transfected HT1080 cells.** HT1080 cells transfected with collagen II constructs were analysed three days post transfection by staining with antibodies directed against collagen (red) and BiP (green). Nuclei were counterstained with bisbenzimide (blue). BiP expression was observed in all the cells. Collagen II (red) expression was detected in cells transfected with WT, R75C, R134C, and R704C constructs and there was no colocalisation with BiP. Colocalisation of BiP and collagen II was seen only in cells expressing R740C and R789C collagens and these cells were also rounded.

An accumulation of misfolded proteins in the ER leads to an induction of stress response related chaperones which attempt to refold the affected proteins. One example is BiP (Sitia and Braakman, 2003) which is an ER resident protein.

To determine if the intracellular protein accumulation is initiating an ER stress response in collagen II transfected HT1080 cells, these were analyzed by immunofluorescence microscopy with antibodies directed against BiP (Fig 3.20). BiP expression was observed in all the cells

(green fluorescence) and collagen II expression was seen around the Golgi compartment (red). In cells expressing R740C and R789C collagens colocalisation of BiP and collagen II proteins was significant and the intensity of BiP staining was also increased. Further, the cell morphology was rounded.

The rounding and detachment of the cells might be a sign of cell death. Accumulation of large amounts of misfolded proteins might interfere with cell viability. It was investigated if these transfected HT1080 cells were destined to apoptosis mediated by an ER stress response due intracellular retention of misfolded proteins. For this purpose an FITC conjugated antibody directed against active caspase-3 was used for staining of HT1080 cells. Caspase-3 belongs to the caspase family of cysteine proteases (Martin and Green, 1995), which exist as latent zymogens or procaspases. Apoptosis is a form of cell death that involves the concerted action of a number of intracellular signaling pathways including members of the caspase family. Oligomerization of procaspases is an important step leading to caspase activation, which in turn is amplified by the cleavage and activation of other downstream caspases (Earnshaw et al 1999 and Salvesen and Dixit, 1997). Activated caspase-8 and -9 activate executioner caspases, including caspase-3. Once active, executioner caspases cleave a number of cellular proteins, e.g. nuclear proteins, structural proteins, signaling molecules and cytoskeletal proteins (Stroh and Schulze 1998). Proteolysis of these proteins disrupts cellular homeostasis and terminates survival signals.

To analyse if R740C and R789C transfected HT1080 cells were positive for activated caspase-3, these along with cells expressing other collagen II variants and non transfected cells, were stained for active caspase-3 three days after transfection by using an FITC conjugated rabbit anti-active caspase-3 antibody (Fig 3.21). Active caspase-3 was not detected in non-transfected cells or the cells transfected with WT, R75C, R134C and R704C constructs, but significant expression was seen in R740C and R789C transfected cells. Since the active form of caspase-3 was present in cells expressing R740C and R789C collagens, these cells are undergoing apoptosis, presumably due to the irreversible accumulation of misfolded proteins in the ER which leads to the activation of the ER stress response and, finally, to cell death.

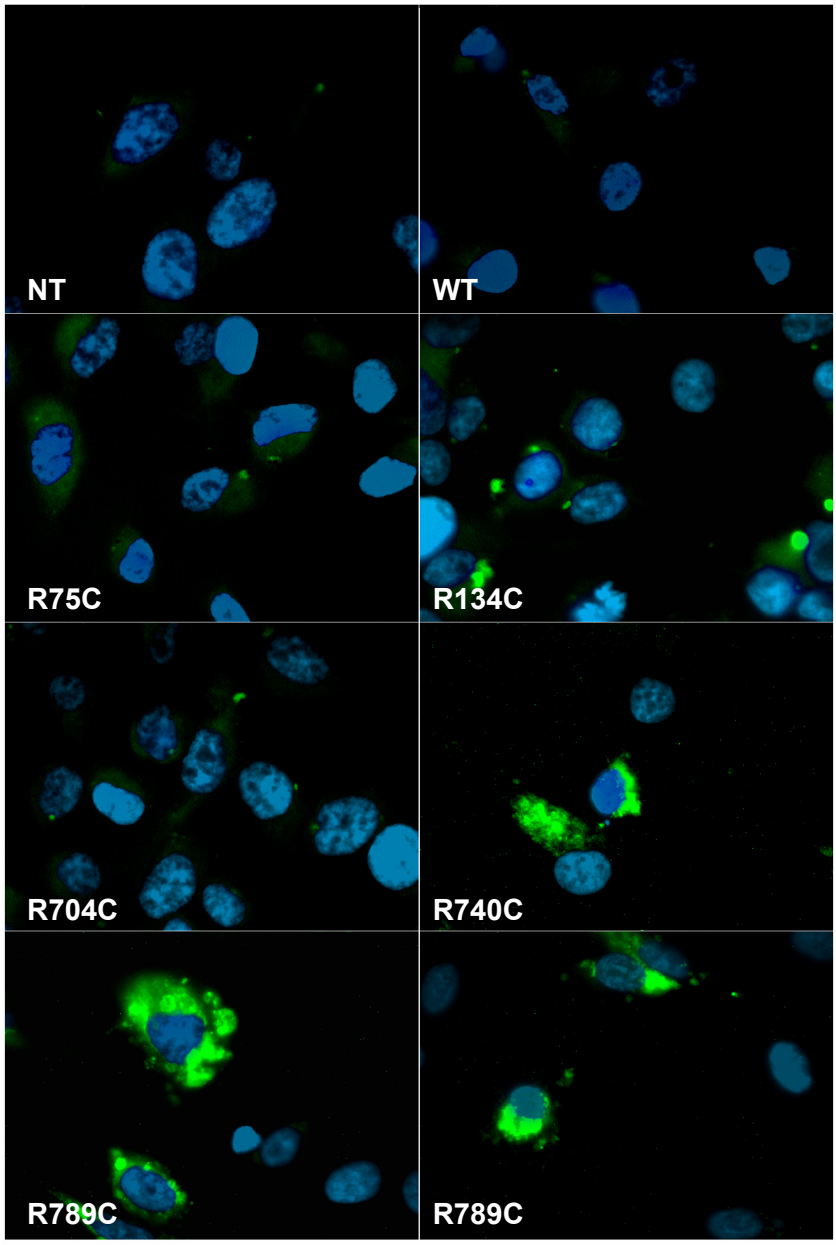


**Fig 3.21 Immunofluorescence staining of transfected HT1080 cells using antibodies directed against active caspase 3.** HT1080 cells transfected with collagen II constructs were analysed three days post transfection for the presence of active caspase-3 (green), a marker for apoptosis. No active caspase-3 was detected in non-transfected cells nor in cells transfected with WT, R75C, R134C, and R704C constructs, whereas active caspase-3 was present in cells expressing R740C and R789C collagens.

## 3.19 Detection of DNA cleaved due to apoptosis by nick labelling

Caspase activation is accompanied by nuclear fragmentation, another hallmark of apoptosis. Hence, it was necessary to analyze if DNA fragmentation occurs in the R740C and R789C collagen expressing HT1080 cells. Non transfected cells and cells transfected with collagen II

variants were labeled for nicked DNA four days after transfection (Fig 3.22). Nick labeled DNA was observed in cells expressing R740C and R789C.

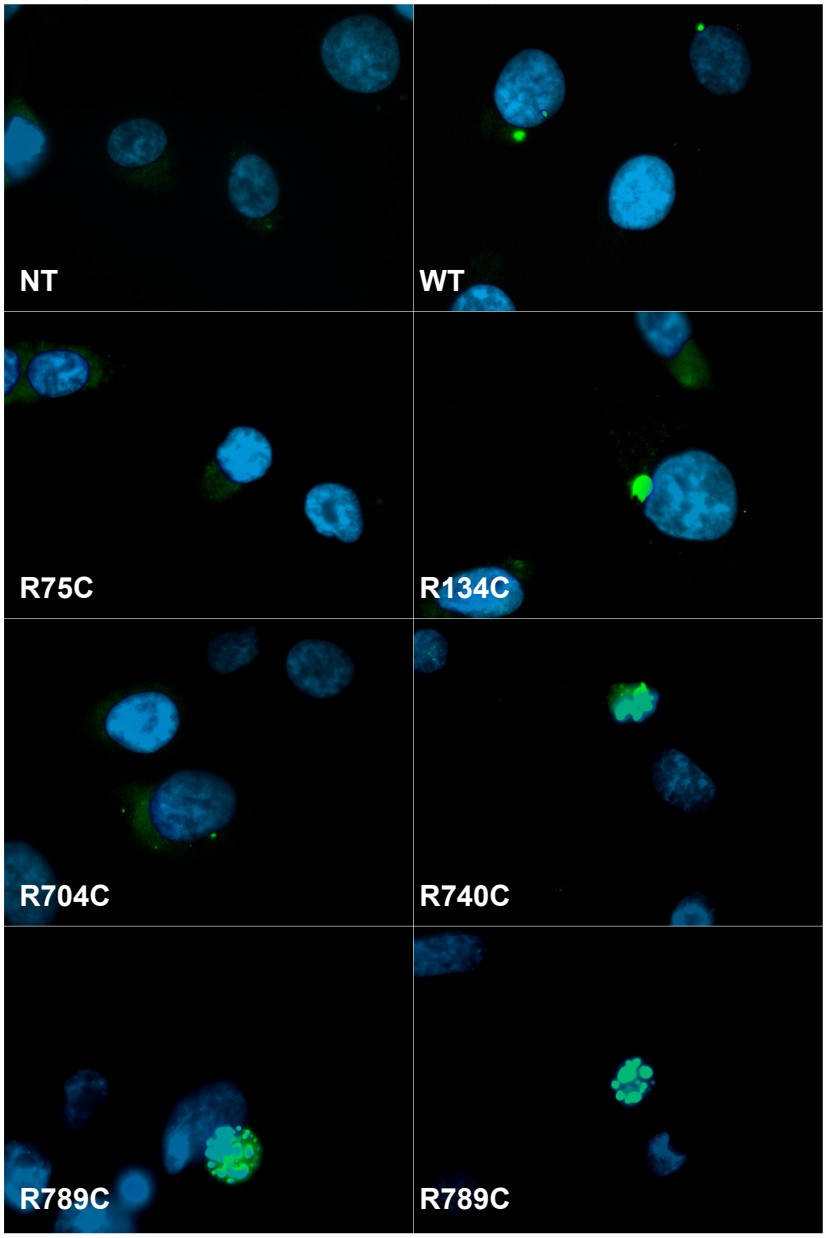


**Fig 3.22 Nick labeling of HT1080 cells transfected with collagen II constructs to detect cleaved DNA.** Four days after transfection with collagen II constructs, HT1080 cells transfected with R740C and R789C constructs showed distinct nick labeling (green). Nick labeling was not seen in non-transfected cells nor in cells transfected with constructs carrying other mutations.

## 3.20 Single cell gel electrophoresis (comet assay) of HT1080 cells transfected with collagen II variants

To ascertain whether activation of caspase-3 and nick labeling of DNA in cells transfected with R740C and R789C constructs leads to DNA fragmentation, HT1080 cells were analysed by

single cell gel electrophoresis four days post transfection. (Fig 3.23). DNA tailing leading to a comet shaped appearance was observed in cells transfected with R740C and R789C constructs (Fig 3.23A). Further, the mean tail length was evaluated for these comets (20 cells from each group) using the CometScore software (Fig 3.23B). Significant increase in tail length, corresponding to DNA fragmentation and apoptosis was observed in cells expressing R740C and R789C collagens.

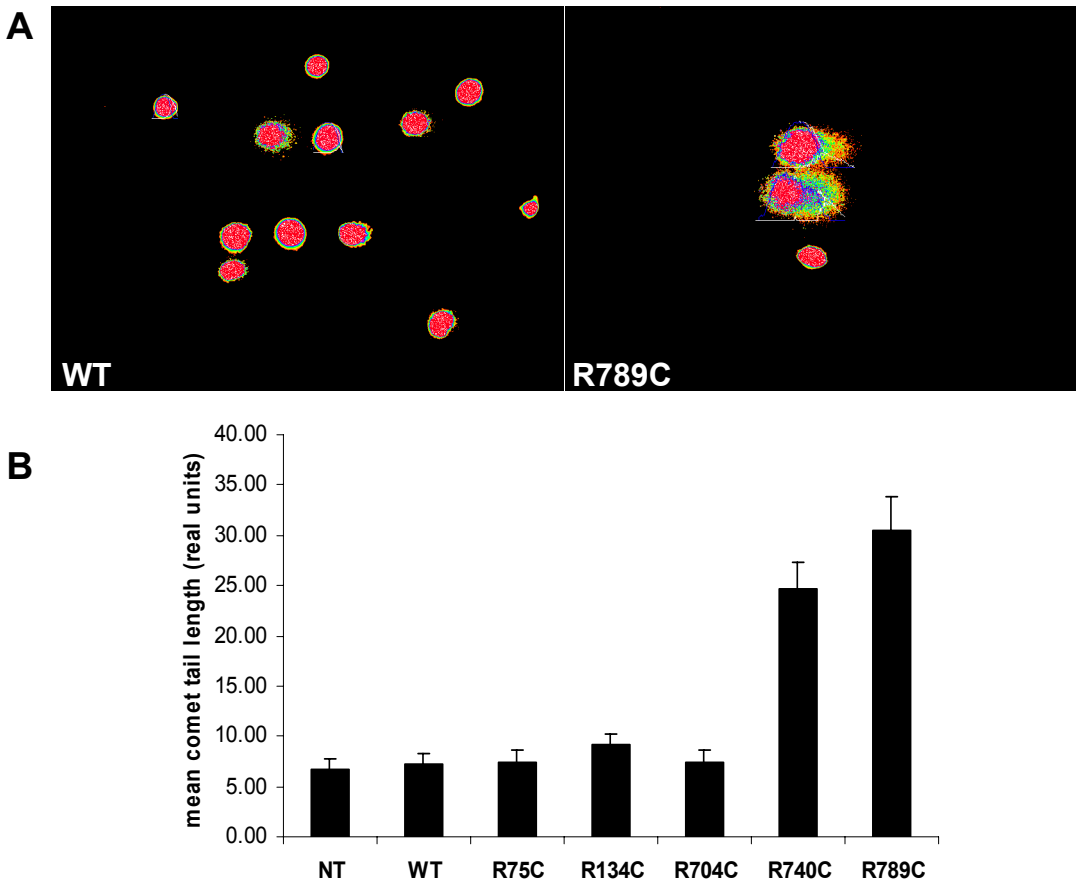


**Fig 3.23 Single cell gel electrophoresis of HT1080 cells transfected with collagen II variants**. Panel A shows representative pictures of wild type and R789C collagen transfected cells after comet assay. In panel B the mean tail lengths of the comets are shown. A significant increase in tail length was observed in cells expressing R740C and R789C collagens whereas cells transfected with WT, R75C, R134C, R704C constructs were similar to non-transfected controls.

## 3.21 Comparison of the efficiency and toxicity of first and third generation adenoviral vectors

Chondrocytes, the cells expressing collagen II in vivo, show a very low transfection efficiency when conventional methods are employed. Adenovirus based gene delivery yields a higher efficiency of gene transfer. Currently, two generations of adenoviral vectors are commonly used for transduction. First generation viruses are easier to produce and handle but only short DNAs

can be inserted. Further, those viruses show immunogenic potential and are toxic in animal models. Third generation viruses are cumbersome to produce and need a helper virus for packaging. However, this type of virus is non-immunogenic and less toxic in animal models and can accept 30-32kb foreign DNA as insert.

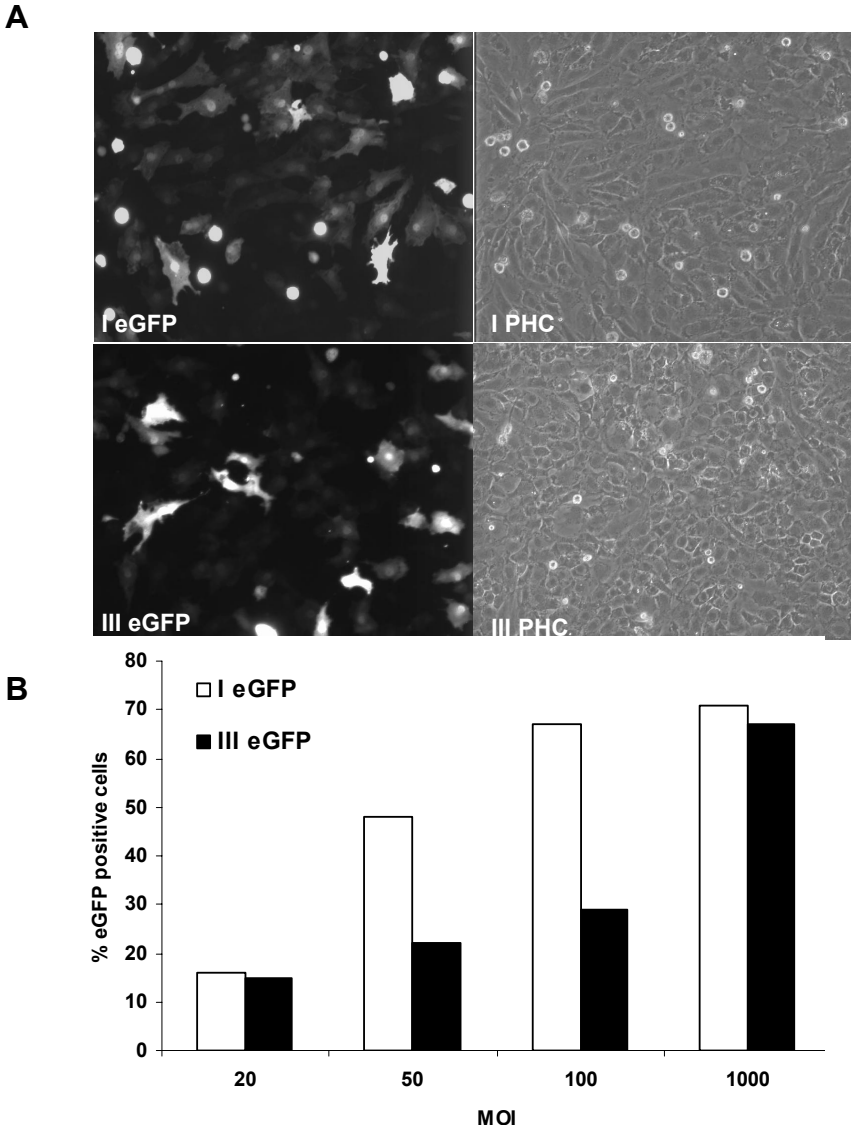


**Fig 3.24 Transduction of primary chondrocytes using first and third generation adenoviral vectors.** Chondrocytes were transduced with first (I) and third generation (III) adenoviral vectors expressing enhanced green fluorescent protein (eGFP) to determine efficiency of transduction and toxicity. A, fluorescent and phase contrast (PHC) images of chondrocytes three days after transduction with 100 MOI. B, viable cells expressing eGFP after transduction with 20, 50, 100 and 1000 MOI (moieties of infection). At 1000 MOI only few cells are viable and these were positive for eGFP.

eGFP expressing first generation and third generation adenoviruses were used for transduction of chondrocytes. The proportion of cells positive for eGFP expression was higher after transduction with first generation vectors as compared with third generation vectors (Fig 3.24A). Similarly the percentage of viable cells expressing eGFP was higher in chondrocytes transducted with first generation adenoviruses as compared to third generation vectors (Fig 3.24B).

# 4 Discussion

Point mutations in fibrillar collagens cause a number of abnormalities in connective tissues that manifest, as for example, brittle bone disease, osteoarthritis and osteochondrodysplasias. Earlier studies on collagen I mutations have revealed that the effects of these are remarkably dependent on their positions within the triple helix and that mutations localized at the N-terminus of collagens are less destabilizing than those occurring at the C-terminus (Byers & Cole 2002). In the present study, collagen II mutations leading to substitution of arginine to cysteine with different patient phenotypes (R75C, R789C (spondyloepiphyseal dysplasia congenital), R704C (Stickler syndrome)), as well as artificial mutations (R134C and R740C) were selected to understand the impact of the amino acid position within the Gly-X-Y repeats and the location either at the N- or the C-terminus of the triple helix.

## 4.1 Effect of point mutations on protein expression and secretion

Wild type and mutant collagen II proteins were expressed by transfection of 293 EBNA cells. Secretion of collagen II proteins into the cell culture media was detected for wild type, R75C, R134C and R704C expressing cells. The R740C protein was detected in equal amounts in both the cell lysate and supernatant, whereas R789C protein was detected almost exclusively in the cell lysate (Fig 3.4). Additionally, the R789C protein migrated lower on SDS gels in comparison to other collagen II variants. All recombinant collagen II variants could be detected with antibodies directed against the myc epitope as well as with collagen II antibodies, demonstrating that the proteins have an intact N-terminus. A shift in the SDS PAGE migration of R789C protein was observed for the first time and is in contrast to earlier reports (Steplewski et al. 2004). In these studies identical mutations (R75C and R789C) were introduced into recombinantly engineered collagen II cDNA (Steplewski et al. 2004, Arnold et al. 1997, 1998, Majsterek et al. 2003), but no differences in the migration behaviour in comparison to the wild type collagen II were observed. In addition, the occurrence of high molecular mass aggregates, due to the presence of newly formed disulfide bridges were described. There is no unpaired cysteine present in the triple helical domain of the wild type protein and a cysteine substitution might therefore lead to abnormal intra- and intermolecular disulfide bridges. In the present study, some high molecular aggregates were detected in case of the R75C protein (data not shown) but not with the other collagen II variants. Altered secretion and moderate intracellular retention of R789C collagen was also reported for transfected HT1080 and SW-1353 cells (Steplewski et al.

2004, Ito et al. 2005), whereas significant intracellular retention and very low rate of secretion was observed for HT1080 cells expressing R789C protein.

Analysis of pepsin extracted R789C protein expressed in rat chondrosarcoma cells showed no shift in the migration pattern upon SDS PAGE and no disulfide linked oligomers were detected (Gaiser et al. 2002). The impression that R789C is not degraded might have been due to the extraction procedure used. Pepsin digestion probably degraded all the mutant chains yielding only trimeric collagen II formed from endogenous wild type collagen II that is pepsin resistant.

The faster migration of R789C might either be due to a difference in the protein structure and folding or to enhanced proteolytic processing of the protein caused by the mutation. Using MALDI TOF mass spectroscopy (Table 3.1), a mass of 92.9 kD was determined for the R789C protein, whereas wild type collagen II and all other variants had a mass of around 150.6 kD. This mass was significantly higher than the theoretical mass, calculated to be 136.8 kD. This higher mass is most probably due to posttranslational modification (glycosylation, hydroxylation) of collagen II during biosynthesis. The extent of collagen II modification could be estimated from the mass determination of pepsin extracted collagen from bovine nasal cartilage. Even in the absence of N- and C-propeptides, which also carry posttranslational modifications, the difference between measured and calculated mass was around 30 kD.

It has been reported that hydroxylation and glycosylation varies markedly between collagen types and even within the same collagen type between tissues and species. In a given tissue the posttranslational modification depends on physiological and pathological states (Kivirikko & Myllyla 1979, Kivirikko 1995). The extent of these modifications plays a major role in collagen fibril formation and stability (Notbohm et al. 1999, Bann & Bachinger 2000, Dominguez at al. 2005). The hydroxylation and glycosylation of recombinantly expressed collagens by various cell culture systems is reviewed by Bulleid et al. (2000). From the results presented here, it might be speculated that the reduced mass for the R789C is due to proteolytic processing of the protein that results in the observed shift in SDS PAGE migration.

In addition to 293 EBNA cells, collagens were expressed in HT1080 cells, a fibroblastic cell line commonly used to study collagen synthesis (Fertala et al. 1994, Steplewski et al. 2005). Again, when expression was analyzed by SDS PAGE, wild type, R75C, R134C and R704C proteins were mainly detected in the supernatants. As in 293 EBNA cells the R740C protein was found in about equal proportions in supernatants and cell lysates. The R789C protein was exclusively found in the cell lysate and again migrated faster on SDS PAGE (Fig 3.13). This demonstrates a processing of the R789C protein irrespective of the cell line used for expression.

The mass spectrometry results strengthen the assumption that the protein is truncated. From the detection of the myc epitope in immunoblot it can be concluded that the R789C protein has an intact N-terminus. Therefore, the truncation of the R789C protein must occur at the C-terminal end of the collagen triple helix. The cleavage might be due to an increased accessibility of proteases to regions around the site of mutation. It is known that the degradation of fibrillar collagen types I, II, and III can occur in non-helical regions as well as through a triple helix cleavage (Liu et al. 1995) resulting in unfolding at physiological temperatures. The collagenases involved in these processes belong to the family of zinc-dependent endopeptidases called matrix metalloproteinases (MMPs). Collagenase-1 or MMP-1 (interstitial collagenase), collagenase-2 or MMP-8 (neutrophil collagenase), and the collagenase-3 or MMP-13 (Freije at al. 1994, Knauper et al. 1996) are the only mammalian enzymes so far identified which are able to initiate the intrahelical cleavage of triple helical collagen at neutral pH. These three collagenases cleave the fibrillar collagens type I, II, and III at a single site ($Gly_{775}$–$Leu/Ile_{776}$) within each α chain of the triple helical collagen molecule (Miller et al. 1976, Mitchell et al. 1996, Billinghurst et al. 1997) (Fig 4.1).

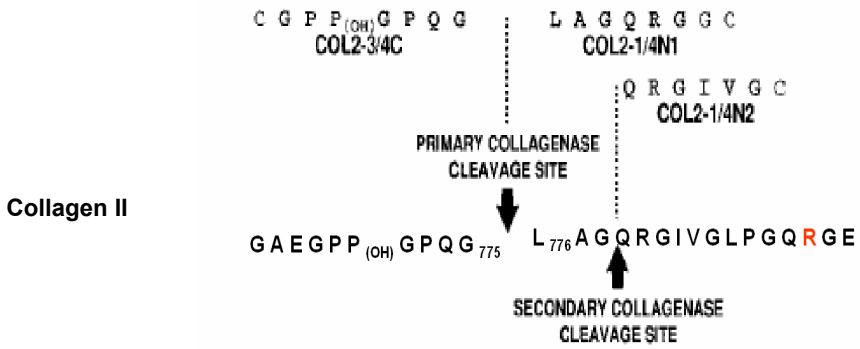


**Fig 4.1 Sequence around the potential collagenase cleavage sites in the triple helical domain of human collagen II. The position of arginine 789 is marked in red.** (modified from Billinghurst et al. 1997).

To investigate the involvement of MMPs in intrahelical cleavage of the R789C protein, GM6001, a universal inhibitor of MMPs which can block both intracellular and extracellular action of MMPs, was used on R789C expressing cells. The cleavage of the R789C protein was partially inhibited by pretreatment of the cells with GM6001 (Fig 3.14). Interestingly, the inhibition of cleavage leads to an increased secretion of the now fully intact protein into the supernatant. Inhibition of MMP cleavage of collagen by GM6001 was also reported earlier (Song & Windsor 2005, Naqvi et al. 2005). Billinghurst et al. (1997) demonstrated the presence of two potential collagenase cleavage sites at the amino acid position 775 and 777 in the triple

helical domain of collagen II (Fig 4.1). The exact cleavage position in the R789C protein is yet to be identified, but could potentially be determined by N-terminal sequencing of the smaller C-terminal fragment. Another possibility was the use of a neoepitope antibody directed against the C-terminus of the so-called COL2-3/4C fragment (Fig 4.1). This antibody would specifically detect the fragment if it is cleaved at the position 775. Unfortunately, no specific signal was detected by immunoblotting using this antibody (data not shown). This might indicate that after initial cleavage by MMPs, the R789C protein has undergone further cleavage, leading to the loss of the recognition site. Another possible explanation is that the mutation in immediate vicinity of the cleavage and recognition site alters the three dimensional structure so that the antibody does not detect the neoepitope even if generated.

The fact that the truncated R789C protein was detected in the cell lysate demonstrates that the cleavage takes place already inside the cells. This implies that MMPs are active within the cellular compartments where the protein is retained. The fact the R789C protein is secreted when not cleaved is in good agreement with earlier studies (Majsterek et al. 2003, Steplewski et al. 2004) where mutation of the MMP cleavage site gave increased secretion of R789C collagen (Arnold et al. 1997, Majsterek et al. 2003, Steplewski et al. 2004).

Since the mutation R789C is in close vicinity of the MMPs cleavage sites it can be speculated that the mutations increases the accessibility to MMPs and thus may cause truncation of the R789C protein. Analysis of patients with some forms of osteoarthritis and as well as transgenic mice with osteoarthritis have indeed shown collagen II cleavage at this position by MMPs (Wu et al. 2002, Salminen et al. 2002, Tchetina et al. 2005, Xu et al. 2005).

Treatment of HT1080 cells expressing recombinant R789C with a universal proteasome inhibitor, MG132, did not affect the mobility of R789C (Fig 3.15), indicating that the mutant protein is not cleaved by endoplasmic reticulum associated protein degradation (ERAD) using the cytoplasmic ubiquitin-proteasome system (Meusser et al. 2005).

## 4.2 Impact of arginine to cysteine mutations on the structure and stability of collagen II

To check the stability of recombinantly expressed collagen II and mutants containing arginine to cysteine mutations, proteins were analysed by protease digestion, CD spectroscopy and electron microscopy. Wild type, R75C, R134C and R704C proteins were protease resistant on treatment with trypsin at 25 °C as evidenced by a protease-resistant band corresponding to the collagenous domain (Fig 3.8). Similar treatment on R740C and R789C collagens lead to complete degradation indicating an unfolding of the triple helix. Such an instability was reported for

patient mutations in COL3A1 causing Ehlers-Danlos syndrome type IV (Narcisi et al. 1993), in COL1A2 mutations leading to type I osteogenesis imperfecta (Zhuang et al. 1993) and in collagens recombinantly expressed in vitro (Bruckner & Prockop 1981, Raghunath et al. 1994, Olague-Marchen et al. 2000, Galicka et al. 2003, Cabral at al. 2001). In contrast to the results presented here, other groups have shown that the R789C protein was resistant to digestion with trypsin and chymotrypsin (Steplewski et al. 2004). However, this trypsin resistance was observed only when the MMP cleavage site was mutated and intact full length R789C protein was secreted. As mentioned above, the presumed MMP processing reduces stability and increases the susceptibility for further degradation.

The impact of R740C and R789C on the stability and integrity of the triple helix was further confirmed by CD spectroscopy and by determination of melting temperatures (Tm). Wild type, R75C, R134C and R704C proteins displayed typical collagen CD spectra (Fig 3.9). In contrast, a shift of the spectrum to below zero at 222 nm for R740C and R789C proteins indicates that there is a decrease in triple helical structures in these mutants. Similarly glycine substitutions lead to destabilization in collagen XVII (Tasanen et al. 2000). The melting temperature of wild type collagen II was 38.6 °C, 2.4 °C lower than for collagen extracted from bovine nasal cartilage (Liang & Chakrabarti 1981, Amudeswari et al. 1987, Arnold et al. 1998). This difference in absolute melting temperature might be caused by inefficient hydroxylation by 293 EBNA cells (Wagner et al. 2000; Mizuno et al. 2003). The R75C, R134C and R704C proteins had a 2.5 °C lower melting temperature than the wild type protein and the Tm values of the R789C and R740C proteins were further decreased to 31.5 °C and 30.2 °C, respectively. The decreased Tm of the R789C protein is in agreement with earlier studies (Steplewski et al. 2004), although in the present study the Tm of the truncated protein was determined whereas others have analyzed the full length form.

The thermal instability is caused by unstable triple helical structures. An increased monomer content and lowered trimer population has been shown to result in dramatic decrease in triple helix content, melting temperature and, in turn, stability when using collagen like model peptides with glycine to alanine exchanges (Bhate et al. 2002). When using an algorithm to predict the structure and stability of collagen triple helices from amino acid sequences (Persikov et al. 2005) the R789C protein depicted a significant difference in the stability around the amino acid sequence 783 to 795. The other mutants studied did not show such pronounced local destabilization (Fig 4.2).

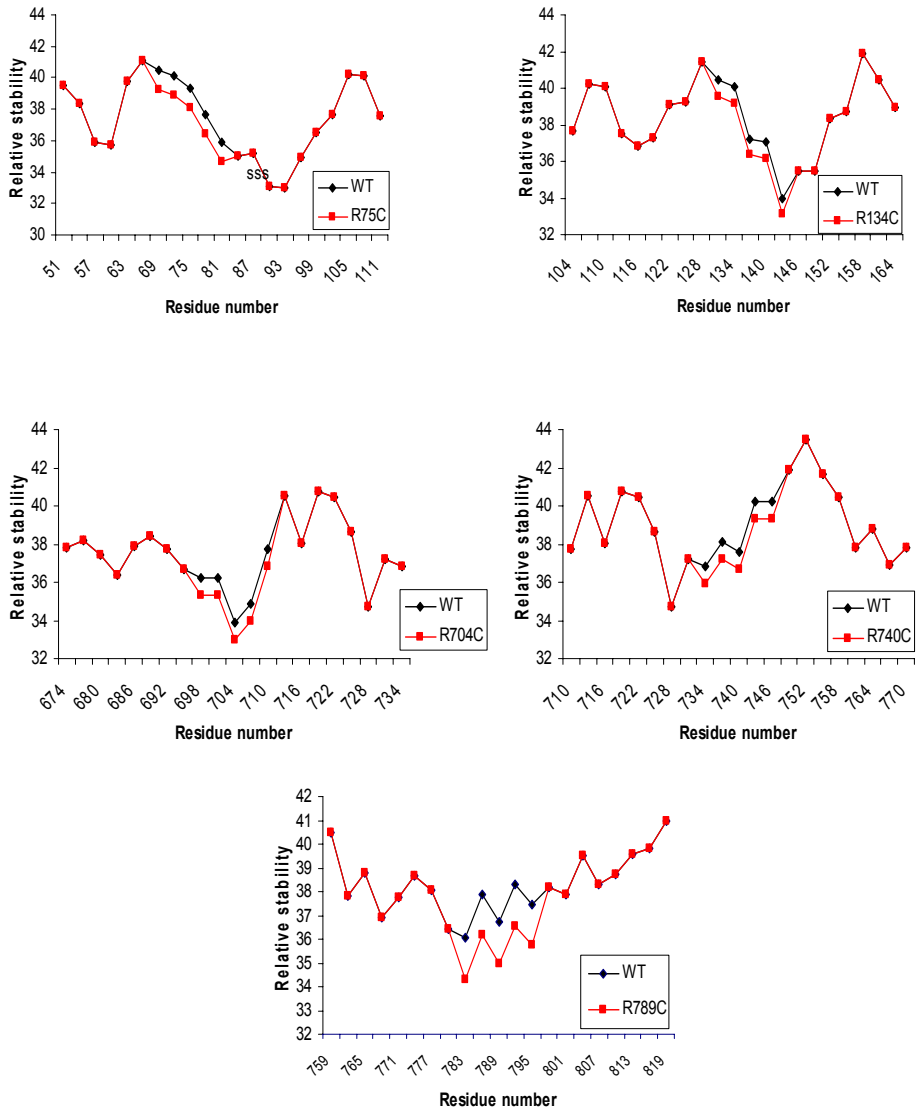


**Fig 4.2 Calculated relative stability profiles in the immediate vicinity of arginine to cysteine mutations in the collagen II compared to the wild type protein.** (http://semimajor.net/collagen_calculator/).

The structural instability of the R740C and R789C proteins was confirmed by negative staining electron microscopy (Fig 3.12). In contrast to the wild type collagen, which showed characteristic procollagen structures (Lightfoot et al. 1992), the R740C and R789C samples contain only short filaments. This again indicates that these molecules do not form stable triple helices and are therefore not visible by negative staining. Other mutant proteins, such as R75C, R134C and R704C, showed filamentous structures comparable with wild type. However, kinks were consistently seen in most of the R75C filaments. The kinks were at the end of the filaments which agrees with where the mutation is located. Similar kinked structures were previously reported for collagen I molecules harboring mutations leading to osteogenesis imperfecta (Vogel et al. 1988, Lightfoot et al. 1992).

All results from biochemical studies support the conclusion that the R740C and R789C proteins form unstable triple helices. In addition, the local instability around the R789C mutation leads to increased accessibility for MMP cleavage leading to a truncated protein which is then susceptible for digestion by less specific proteases.

## 4.3 Impact of collagen II mutations on protein trafficking

The fibroblastic cell line HT1080 is commonly used to study the protein expression, trafficking and secretion of collagen II (Fertala et al. 1994, Steplewski et al. 2005). Immunoblotting revealed that a considerable amount of R740C and R789C protein was present in the cell lysate, suggesting intracellular retention.

By immunofluorescence costaining with compartment specific antibodies, the WT, R75C, R134C and R704C proteins were seen to be mainly present in the Golgi compartment (Fig 3.18). This indicates that most of these proteins are being transported to the extracellular space. In contrast, R740C and R789C proteins were detected in other cellular compartments, and showed only weak colocalisation with the 58k Golgi marker. Moreover, colocalisation of R740C and R789C procollagens with PDI, a marker for the endoplasmic reticulum, was observed in these cells (Fig 3.19). PDI is an ER resident chaperone and functions in retaining unassembled or misfolded proteins within the ER. It binds to unfolded collagens preventing aggregation (John et al. 1993, Kivirikko & Myllyharju 1998) and acts as an ER resident quality control. This might suggest that the mutant R740C and R789C proteins are recognized as misfolded by the PDI which may be bound to these proteins, leading to colocalisation. Misfolded proteins are retained by the ER quality control system and subjected to either degradation or refolding (Bottomley et al. 2001). The intracellular retention of the misfolded R740C and R789C collagens in the ER is in agreement with the intracellular retention, due to point mutations, of other cartilage proteins such as COMP and collagen IX in patients with inherited skeletal dysplasias (Hecht et al. 2004, Vranka et al. 2001). Intracellular retention was also observed in mice harboring mutations in the collagen II C-propeptide (Fernandes et al. 2003). Intracellular retention has also been shown to result from point mutations in the triple helical domain of collagen II (Tiller at al. 1995). Interestingly, most of the retained collagens have substitutions of glycine in the Gly-X-Y repeats in the triple helical domain (Tiller et al. 1995, Vissing et al. 1989). Intracellular retention was reported to be due to abnormal disulfide bridges after the introduction of cysteine in the triple helix. Cysteine induced abnormal disulfide bridges were reported for COL1A1 mutations in cells cultured from patient skin biopsies, which also showed a dilated endoplasmic reticulum (Nuytinck et al. 2000). Along the same lines glycine to cysteine substitution in collagen II caused

reduced secretion of the protein and protein dimers were formed (Mundlos et al. 1996). Similarly, truncated protein due to exon deletion in the α3 chain of collagen IX caused intracellular retention in chondrocytes (Bonnemann et al. 2000). It appears that mutation-induced intracellular retention can be due to various factors, such as substitution by bulkier amino acid, presence of artificial disulfide bridges, misfolding as in R740C and truncation as in R789C.

The retention of misfolded protein in the ER leads to the activation of a complex signal transduction pathway called the misfolded or unfolded protein response (UPR) (Rutkowski & Kaufman 2004, Zhang & Kaufman 2004). The UPR has evolved to limit the potential impact of protein accumulation on cellular homeostasis and involves the activation of transcription factor-6 and XBP-1 (X-box DNA binding protein-1) (Lee et al. 2003, Yoshida et al. 2001). In cells undergoing ER stress, active XBP-1 is generated by excision of a 26 bp nucleotide sequence from the XBP-1 transcript by IRE1 endonuclease, one of three ER transmembrane "stress sensors" (Calfon et al. 2002). RT-PCR results from transfected HT1080 cells show the presence of XBP-1 spliced variant in addition to the unspliced form in cells expressing R740C and R789C (Fig 3.16). This result is in agreement with the results from Wilson et al. (2005) where misfolded protein accumulation due to a mutation in the NC1 domain of collagen X lead to splicing of XBP-1. XBP-1 splicing leads to the active up-regulation of ER-resident molecular chaperones like BiP and other components of the ER-associated protein degradation machinery (Yoshida et al. 2003, Oyadomari & Mori 2004). Upregulation and colocalisation of BiP (ER resident molecular chaperone) with R740C and R789C procollagen chains was observed by immunofluorescence (Fig 3.20). Similar results were also reported for upregulation and binding of BiP to collagen chains harboring mutations in type I collagen from patients with osteogenesis imperfecta (Chessler & Byers 1993). Splicing of XBP-1 leads to upregulation of BiP in response to accumulation of misfolded collagen X (Wilson et al. 2005). In case of the R740C and R789C collagen II mutations it appears that misfolded and truncated protein, accumulate in the ER and this triggers an ER stress response leading to the activation of ERAD. ER stress response also can downregulate protein synthesis and lead to growth arrest, PERK an ER transmembrane protein which plays a role in inhibition of cell cycle and PERK associated phosphorylation of translation elongation factor 2α (eIF2α) blocks translation (Brewer & Diehl 2000, Kaufman 1999, Sood et al. 2000). Inhibition of proteasomes using MG132 did not affect R789C collagen (Fig 3.14), which indicates that this collagen is not degraded by the ERAD machinery. This leads to accumulation of misfolded proteins in the ER of R740C and R789C expressing cells and this abnormal accumulation results in an altered cellular morphology (Fig 3.17) which correlates with cell death. Similar flattened chondrocytes as well as a reduced cell number was reported in mice

harboring R789C mutation. These chondrocytes also displayed a distended endoplasmic reticulum (Gaiser et al. 2002).

## 4.4 Fate of cells expressing R740C and R789C collagens

Accumulation of misfolded R740C and R789C procollagens in the ER does not result in degradation but continuous protein accumulation may cause programmed cell death (Breckenridge et al. 2003, Sitia & Braakman 2003, Rao et al. 2004). Indeed, the large amount of rounded and dead cells observed in cultures expressing R740C and R789C collagens (Fig 3.17) may indicate that the cells are undergoing apoptosis. This was confirmed by the detection of active caspase-3 in cells transfected with R740C and R789C constructs but not in cells transfected with other collagen II variants. Active caspase-3 is implicated in mutations leading to disruption of cellular homeostasis and apoptosis (Mulugeta et al. 2005). Presence of nicked DNA (Fig 3.22) and increased tail length of comets (Fig 3.23) (DNA fragmentation) in cells expressing R740C and R789C collagens confirms that these cells are undergoing apoptosis (Fig 4.3). This result agrees with other studies in which cell death was reported due to the intracellular accumulation of mutant COMP proteins (Hashimoto et al. 2003) and apoptosis in chondrocytes due to ER stress has been reported (Yang et al. 2005). Our results indicates that expression of R740C and R789C collagens leads to an ER storage disease, similar to that observed in pseudoachondroplasia (Dinser et al. 2002) and this may be the reason for the reduced number of chondrocytes observed in patients and in animal models (Hecht et al. 1998, 2005, Gaiser et al. 2002).

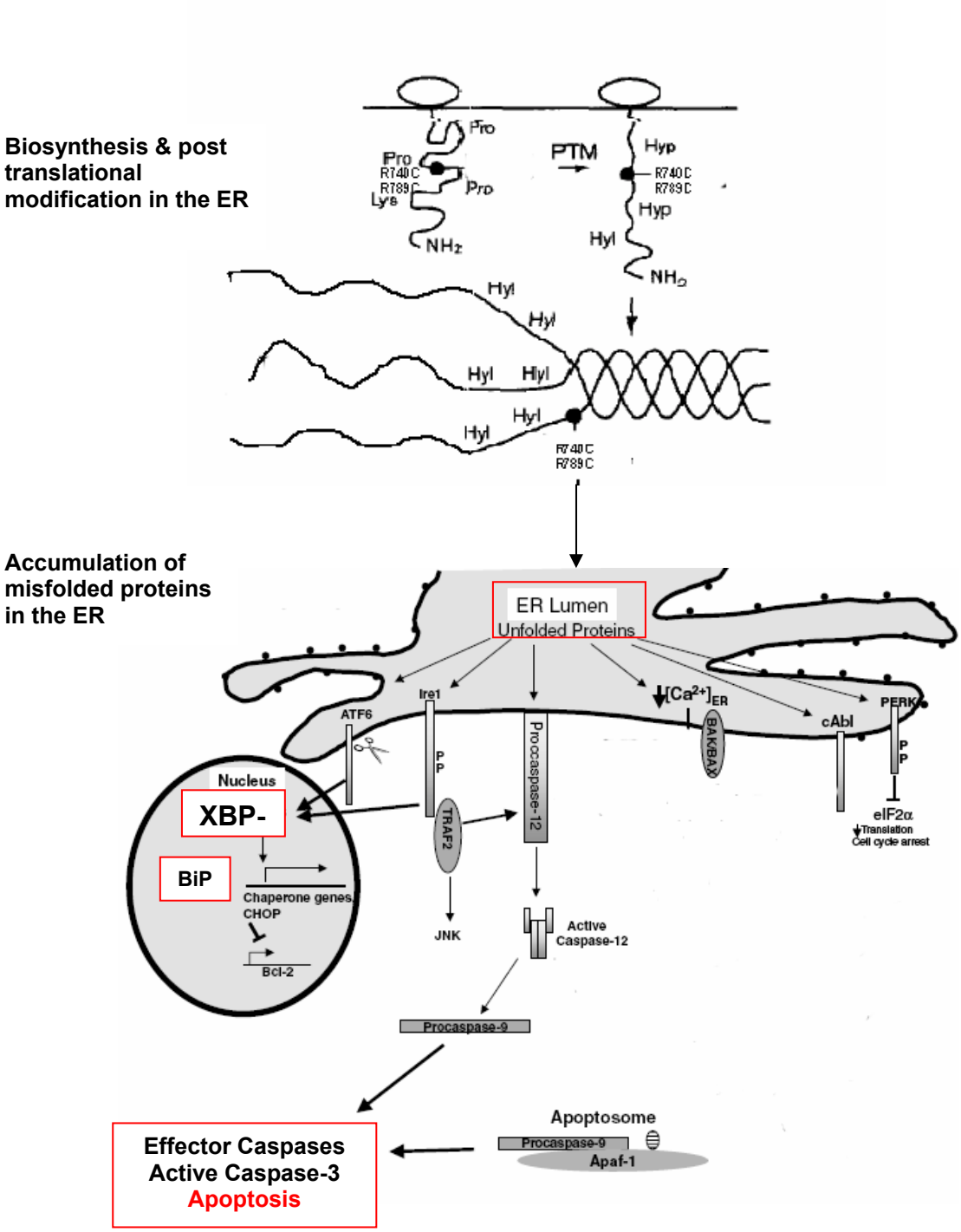


**Fig. 4.3 Schematic representation of intracellular events occurring in the bioynthesis of R740C and R789C collagens.** R740C and R789C mutant proteins do not form proper triple helical molecules. This leads to impaired trafficking and thereby accumulation in the ER lumen which in turn triggers an ER stress response leading to splicing of XBP-1 by IRE1. Finally as the proteins are not degraded, accumulation triggers apoptosis as seen by the activation of effector caspases. (modified from Baum & Brodsky 1999, Breckenridge et al. 2003).

## 4.5 Consequences of intracellular retention of collagen II for the assembly of extracellular matrix

The cartilage extracellular matrix is composed of intricate networks, formed by interactions between a large number of proteins and proteoglycans. Collagen II is the major collagenous component in cartilage and forms fibrils together with other collagens such as collagen IX and XI (Reginato & Olsen 2002) that directly or indirectly interact with noncollagenous components, e.g. aggrecan, fibronectin, decorin, fibromodulin, lumican, matrilins and COMP (Poole 2001, Mann et al. 2004). Thereby mutations in collagen II may have profound effects on the overall structure of the extracellular matrix.

Intracellular retention with delayed secretion of R789C collagen leads to abnormal binding to fibronectin, thereby altering the intracellular and extracellular processes in which fibronectin takes part (Ito et al. 2005). In addition, altered interactions of R789C collagen II with collagen IX may lead to a change in the architecture of the ECM in the patients harboring these mutations (Ito et al. 2005, Steplewski et al. 2005). Also using computer modeling, R789C displayed a decrease in electrostatic potential around the vicinity of the mutation (Steplewski at al. 2005) and this was correlated to the altered interaction with collagen IX.

In some chondrodysplasias secretion of mutant COMP is markedly delayed, and collagen IX is partially coretained in the ER. However, collagen II secretion is not affected and collagen II can be detected in the matrix (Dinser et al. 2002, Hecht et al. 2005). This indicates that the secretory pathways for COMP and collagen II are different and that the secretion of collagen IX is more closely associated with that of COMP.

Complete retention of collagen II could cause its complete absence in the matrix. This would represent a knockout situation as in collagen II deficient mouse models. These show highly disorganized chondrocytes in a cartilage which completely lacks extracellular fibrils (Liu et al. 1995). Similar observations were also made in transgenic mice carrying a partially deleted human COL2A1 gene which developed a chondrodysplasia phenotype. The extracellular matrix in these mice contained fewer collagen fibrils and the individual fibrils were thinner (Vandenberg et al. 1991). Similarly, transgenic mice harboring glycine to cysteine and arginine to cysteine mutations in the triple helical domain developed chondrodysplasia. The chondrocytes displayed a distended and dilated endoplasmic reticulum and the matrix consisted of thin collagen fibrils (Garofalo et al. 1991, Gaiser at al. 2002). This agrees well with the results with R740C and R789C collagen expressed by HT1080 cells where the intracellular accumulation of mutant proteins leads to cell death and this may contribute to altered extracellular matrix.

In case of R704C collagen, the mutation might lead to haploinsufficiency of collagens fibrils in the matrix. Haploinsufficiency was indicated in mutations in collagen II leading to Stickler syndrome, where mutated mRNA was subjected to nonsense-mediated mRNA decay (Richards et al. 2000) and may be the molecular basis of Stickler syndrome (Freddi et al. 2000). Although the HT1080 cells are a frequently used model, attempts should be made to confirm the result with real chondrocytes, which are the natural producers of collagen II. However, repeated attempts to transfect chondrocytes with plasmid vectors encoding collagen II and its variants were unsuccessful. Hence, adenoviral vectors containing collagen II cDNA with the desired mutations were cloned. Adenovirus based vectors have been successfully used to study the effects of COMP mutations in chondrocytes (Dinser et al 2002). The efficiency and toxicity of first and third generation adenoviral vectors in transducing chondrocytes were evaluated by using eGFP expressing viruses. First generation vectors were highly efficient with decreased toxicity in comparison with third generation vectors when used to transducer cultured bovine chondrocytes (Fig 3.24). Further analysis of transduced chondrocytes will shed light on the intricate role of different mutations in the cartilage.

## 4.6 Conclusion

The results obtained with point mutations resulting in substitution of arginine to cysteine within the triple helical domain of collagen II carry the following implications:

The R75C (Y position) and R134C (X position) mutations did not cause any biochemical alterations except for the presence of kinks in R75C collagen. Similar result were obtained for the R704C (X position) mutation. In case of R740C (X position) and R789C (Y position) the mutations had deleterious effects on protein structure irrespective of being in the X or Y position.

Cells expressing the R75C, R134C and R704C mutations were phenotypically normal and the protein trafficking was not affected, whereas cell death was observed for R740C and R789C mutants, caused by the accumulation of misfolded proteins in the endoplasmic reticulum.

With regard to the location, mutations at the N-terminus of the triple helix had less severe effects than such towards the C-terminus. The R740C and R789C mutation located towards the C-terminus displayed a deleterious effect on the protein structure and caused apoptosis of the expressing cells irrespective of being in the X or Y position. This result is in agreement with the finding that glycine substitutions towards the C-terminus of the collagen I chains are clinically more severe than those towards the N-terminus (Bateman et al. 1992) and may be due to the fact that mutations at the C-terminus lead to a disruptive effect during helix initiation and

propagation (Bonadio & Byers 1985, Byers 1990, Kuivaniemi et al 1991, Bateman et al. 1984, 1986, 1992). However R704C, another mutation towards the C terminus, did not show any lethal effect. The phenotype in these patients may be due to haploinsuficiency in collagen II (Richards et al. 2000).

In patients with R75C and R704C, the mutations might contribute to disease together with other factors such as genetic modifiers or the presence of additional mutations in other gene loci yet to be identified. Similar observations have been made in patients with MED (Jakkula et al. 2005).

These results suggests that a single amino acid alteration in collagen II could lead to skeletal abnormalities through multiple secondary effects on the synthesis and assembly of ECM components. Chondrodysplasia is clearly caused not just by the formation of abnormal matrix molecules but also by the effect of these mutations on protein trafficking and apoptosis possibly augmented by environmental factors and genetic modifiers. The alteration of ECM components may lead to a cascade of disruption of other gene activities which collectively contribute to the pathological changes and thereby contribute to the phenotype in patients.

# 5 References

Ahmad NN, Ala-Kokko L, Knowlton RG, Jimenez SA, Weaver EJ, Maguire JI, Tasman W, Prockop DJ. (1991). Stop codon in the procollagen II gene (COL2A1) in a family with the Stickler syndrome (arthro-ophthalmopathy). Proc Natl Acad Sci USA. (88), 6624-7.

Ahmad NN, Dimascio J, Knowlton RG, Tasman WS. (1995). Stickler syndrome. A mutation in the nonhelical 3' end of type II procollagen gene. Arch Ophthalmol. (113), 1454-7.

Ahmad NN, McDonald-McGinn DM, Zackai EH, Knowlton RG, LaRossa D, DiMascio J, Prockop DJ. (1993). A second mutation in the type II procollagen gene (COL2AI) causing stickler syndrome (arthro-ophthalmopathy) is also a premature termination codon. Am J Hum Genet. (52), 39-45.

Ala-Kokko L, Baldwin CT, Moskowitz RW, Prockop DJ. (1990). Single base mutation in the type II procollagen gene (COL2A1) as a cause of primary osteoarthritis associated with a mild chondrodysplasia. Proc Natl Acad Sci USA. (87), 6565-8.

Ala-Kokko L, Kvist AP, Metsaranta M, Kivirikko KI, de Crombrugghe B, Prockop DJ, Vuorio E. (1995). Conservation of the sizes of 53 introns and over 100 intronic sequences for the binding of common transcription factors in the human and mouse genes for type II procollagen (COL2A1). Biochem J. (308), 923-9.

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. (1994). Molecular biology of THE CELL (Third ed). Garland publishing Inc, New York, USA, 981-982.

Ala-Kokko L, Prockop DJ. (1990). Completion of the intron-exon structure of the gene for human type II procollagen (COL2A1): variations in the nucleotide sequences of the alleles from three chromosomes. Genomics. (8), 454-60.

Altman RD, Hochberg M, Murphy WA Jr, Wolfe F, Lequesne M. (1995). Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage. (3), 3-70.

Anderson IJ, Tsipouras P, Scher C, Ramesar RS, Martell RW, Beighton P. (1990). Spondyloepiphyseal dysplasia, mild autosomal dominant type is not due to primary defects of type II collagen. Am J Med Genet. (37), 272-6.

Arnold WV, Fertala A, Sieron AL, Hattori H, Mechling D, Bachinger HP, Prockop DJ. (1998). Recombinant procollagen II: Deletion of D period segments identifies sequences that are required for helix stabilization and generates a temperature-sensitive N-proteinase cleavage site. J Biol Chem. (273), 31822-8.

Arnold WV, Sieron AL, Fertala A, Bachinger HP, Mechling D, Prockop DJ. (1997). A cDNA cassette system for the synthesis of recombinant procollagens. Variants of procollagen II lacking a D-period are secreted as triple-helical monomers. Matrix Biol. (16), 105-16.

Bachinger HP, Fessler LI, Timpl R, Fessler JH. (1981). Chain assembly intermediate in the biosynthesis of type III procollagen in chick embryo blood vessels. J Biol Chem. (256), 13193-9.

Ballo R, Beighton PH, Ramesar RS. (1998). Stickler-like syndrome due to a dominant negative mutation in the COL2A1 gene. Am J Med Genet. (80), 6-11.

Bann JG, Bachinger HP. (2000). Glycosylation/Hydroxylation-induced stabilization of the collagen triple helix. 4-trans-hydroxyproline in the Xaa position can stabilize the triple helix. J Biol Chem. (75), 466-9.

Bateman JF, Chan D, Mascara T, Rogers JG, Cole WG. (1986). Collagen defects in lethal perinatal osteogenesis imperfecta. Biochem J. (240), 699-708.

Bateman JF, Mascara T, Chan D, Cole WG. (1984). Abnormal type I collagen metabolism by cultured fibroblasts in lethal perinatal osteogenesis imperfecta. Biochem J. (217), 103-15.

Bateman JF, Moeller I, Hannagan M, Chan D, Cole WG. (1992). Characterization of three osteogenesis imperfecta collagen alpha 1(I) glycine to serine mutations demonstrating a position-dependent gradient of phenotypic severity. Biochem J. (288), 131-5.

Baum J, Brodsky B. (1999). Folding of peptide models of collagen and misfolding in disease. Curr Opin Struct Biol. (9), 122-8.

Becker J, Craig EA. (1994). Heat-shock proteins as molecular chaperones. Eur J Biochem. (219), 11-23.

Bella J, Brodsky B, Berman HM. (1995). Hydration structure of a collagen peptide. Structure. (3). 893-906.

Bella J, Eaton M, Brodsky B, Berman HM. (1994). Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution. Science. (266), 75-81.

Bhate M, Wang X, Baum J, Brodsky B. (2002). Folding and conformational consequences of glycine to alanine replacements at different positions in a collagen model peptide. Biochemistry. (41), 6539-47.

Billinghurst RC, Dahlberg L, Ionescu M, Reiner A, Bourne R, Rorabeck C, Mitchell P, Hambor J, Diekmann O, Tschesche H, Chen J, Van Wart H, Poole AR. (1997). Enhanced cleavage of type II collagen by collagenases in osteoarthritic articular cartilage. J Clin Invest. (99), 1534-45.

Bleasel JF, Bisagni-Faure A, Holderbaum D, Vacher-Lavenu MC, Haqqi TM, Moskowitz RW, Menkes CJ. (1995). Type II procollagen gene (COL2A1) mutation in exon 11 associated with spondyloepiphyseal dysplasia, tall stature and precocious osteoarthritis. J Rheumatol. (22), 255-61.

Bleasel JF, Holderbaum D, Brancolini V, Moskowitz RW, Considine EL, Prockop DJ, Devoto M, Williams CJ. (1998). Five families with arginine 519-cysteine mutation in COL2A1: evidence for three distinct founders. Hum Mutat. (12), 172-6.

Bleasel JF, Holderbaum D, Mallock V, Haqqi TM, Williams HJ, Moskowitz RW. (1996). Hereditary osteoarthritis with mild spondyloepiphyseal dysplasia--are there "hot spots" on COL2A1?. J Rheumatol. (23), 1594-8.

Bonadio J, Byers PH. (1985). Subtle structural alterations in the chains of type I procollagen produce osteogenesis imperfecta type II. Nature. (316), 363-6.

Bonaventure J, Philippe C, Plessis G, Vigneron J, Lasselin C, Maroteaux P, Gilgenkrantz S. (1992). Linkage study in a large pedigree with Stickler syndrome: exclusion of COL2A1 as the mutant gene. Hum Genet. (90), 164-8.

Bonfanti L, Mironov AA Jr, Martinez-Menarguez JA, Martella O, Fusella A, Baldassarre M, Buccione R, Geuze HJ, Mironov AA, Luini A. (1998). Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. Cell. (95), 993-1003

Bonnemann CG, Cox GF, Shapiro F, Wu JJ, Feener CA, Thompson TG, Anthony DC, Eyre DR, Darras BT, Kunkel LM. (2000). A mutation in the alpha 3 chain of type IX collagen causes autosomal dominant multiple epiphyseal dysplasia with mild myopathy. Proc Natl Acad Sci USA. (97), 1212-7.

Borochowitz Z, Ornoy A, Lachman R, Rimoin DL. (1986) Achondrogenesis II-hypochondrogenesis: variability versus heterogeneity. Am. J. Med. Genet. (24), 273-288.

Bottomley MJ, Batten MR, Lumb RA, Bulleid NJ. (2001). Quality control in the endoplasmic reticulum: PDI mediates the ER retention of unassembled procollagen C-propeptides. Curr Biol. (11), 1114-8.

Breckenridge DG, Germain M, Mathai JP, Nguyen M, Shore GC. (2003). Regulation of apoptosis by endoplasmic reticulum pathways. Oncogene. (22), 8608-18.

Brewer JW, Diehl JA. (2000). PERK mediates cell-cycle exit during the mammalian unfolded protein response. Proc Natl Acad Sci USA. (97), 12625-30.

Brewton RG, Mayne R, (1992). Mammalian vitreous humor contains networks of hyaluronanmolecules: electron microscopic analysis using the hyaluronan-binding region (G1) of aggrecanand link protein. Exp Cell Res (198), 237-249.

Brown DM, Nichols BE, Weingeist TA, Sheffield VC, Kimura AE, Stone EM. (1992). Procollagen II gene mutation in Stickler syndrome. Arch Ophthalmol. (110), 1589-93.

Brown DM, Vandenburgh K, Kimura AE, Weingeist TA, Sheffield VC, Stone EM. (1995). Novel frameshift mutations in the procollagen 2 gene (COL2A1) associated with Stickler syndrome (hereditary arthro-ophthalmopathy). Hum Mol Genet. (4), 141-2.

Bruckner P, Prockop DJ. (1981). Proteolytic enzymes as probes for the triple-helical conformation of procollagen. Anal Biochem. (110), 360-8.

Bulleid NJ, John DC, Kadler KE. (2000). Recombinant expression systems for the production of collagen. Biochem Soc Trans. (28), 350-3.

Bulleid NJ, Wilson R, Lees JF. (1996). Type-III procollagen assembly in semi-intact cells: chain association, nucleation and triple-helix folding do not require formation of inter-chain disulphide bonds but triple-helix nucleation does require hydroxylation. Biochem J. (317), 195-202.

Byers PH, Cole WG. (2002). Osteogenesis imperfecta. In Connective Tissue and Its Heritable Disorders. Molecular, Genetic, and Medical Aspects, 2nd edn, (Royce, P.M. and Steinmann, B. ed), Wiley-Liss, 385-430.

Byers PH. (1990), Brittle bone-fragile molecules: disorders of collagen gene structure and expression. Trends Genet. (6), 293-300.

Cabral WA, Chernoff EJ, Marini JC. (2001). G76E substitution in type I collagen is the first nonlethal glutamic acid substitution in the alpha1(I) chain and alters folding of the N-terminal end of the helix. Mol Genet Metab. (72), 326-35.

Calfon M, Zeng H, Urano F, Till JH, Hubbard SR, Harding HP, Clark SG, Ron D. (2002). IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. Nature. (415), 92-6.

Canty EG, Kadler KE. (2005). Procollagen trafficking, processing and fibrillogenesis. J Cell Sci. (118), 1341-53.

Chan D, Taylor TK, Cole WG. (1993). Characterization of an arginine 789 to cysteine substitution in alpha 1 (II) collagen chains of a patient with spondyloepiphyseal dysplasia. J Biol Chem. (268), 15238-45.

Cheah KS, Lau ET, Au PK, Tam PP. (1991). Expression of the mouse alpha 1(II) collagen gene is not restricted to cartilage during development. Development. (111), 945-53.

Chessler SD, Byers PH. (1992). Defective folding and stable association with protein disulfide isomerase/prolyl hydroxylase of type I procollagen with a deletion in the pro alpha 2(I) chain that preserves the Gly-X-Y repeat pattern. J Biol Chem. (267), 7751-7.

Chessler SD, Byers PH. (1993). BiP binds type I procollagen pro alpha chains with mutations in the carboxyl-terminal propeptide synthesized by cells from patients with osteogenesis imperfecta. J Biol Chem. (268), 18226-33.

Colige A, Li SW, Sieron AL, Nusgens BV, Prockop DJ, Lapiere CM. (1997). cDNA cloning and expression of bovine procollagen I N-proteinase: a new member of the superfamily of zinc-metalloproteinases with binding sites for cells and other matrix components. Proc Natl Acad Sci. (94), 2374-9.

Colige A, Vandenberghe I, Thiry M, Lambert CA, Van Beeumen J, Li SW, Prockop DJ, Lapiere CM, Nusgens BV. (2002). Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying high homology with ADAMTS-2 and ADAMTS-3. J Biol Chem. (277), 5756-66.

Cooper GM, Hausman RE. (2004). The Cell: A Molecular Approach. Palgrave Macmillan, (third edn).

Creamer P, Hochberg MC. (1997). Osteoarthritis. Lancet. (350), 503-8.

Dalgleish R. (1997). The human type I collagen mutation database. Nucleic Acids Res. (25), 181-7.

Dalgleish R. (1998). The Human Collagen Mutation Database 1998. Nucleic Acids Res. (26), 253-5.

Dinser R, Zaucke F, Kreppel F, Hultenby K, Kochanek S, Paulsson M, Maurer P. (2002). Pseudoachondroplasia is caused through both intra- and extracellular pathogenic pathways. J Clin Invest. (110), 505-13.

Doege KJ, Fessler JH. (1986). Folding of carboxyl domain and assembly of procollagen I. J Biol Chem. (261), 8924-35.

Dominguez LJ, Barbagallo M, Moro L. (2005). Collagen overglycosylation: a biochemical feature that may contribute to bone quality. Biochem Biophys Res Commun. (330), 1-4.

Earnshaw WC. (1999). Apoptosis. A cellular poison cupboard. Nature. (397), 387, 389.

Engel J, Prockop DJ. (1991). The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. Annu Rev Biophys Chem. (20), 137-52.

Eyre DR, Upton MP, Shapiro FD, Wilkinson RH, Vawter GF. (1986). Nonexpression of cartilage type II collagen in a case of Langer-Saldino achondrogenesis. Am J Hum Genet. (39), 52-67.

Fernandes RJ, Hirohata S, Engle JM, Colige A, Cohn DH, Eyre DR, Apte SS. (2001). Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis. J Biol Chem. (276), 31502-9.

Fernandes RJ, Seegmiller RE, Nelson WR, Eyre DR. (2003). Protein consequences of the Col2a1 C-propeptide mutation in the chondrodysplastic Dmm mouse. Matrix Biol. (22), 449-53.

Fertala A, Sieron AL, Ganguly A, Li SW, Ala-Kokko L, Anumula KR, Prockop DJ. (1994). Synthesis of recombinant human procollagen II in a stably transfected tumour cell line (HT1080). Biochem J. (298), 31-7.

Forster SJ, Freedman RB. (1984). Catalysis by protein disulphide-isomerase of the assembly of trimeric procollagen from procollagen polypeptide chains. Biosci Rep. (4), 223-9. Francomano CA, Liberfarb RM, Hirose T, Maumenee IH, Streeten EA, Meyers DA, Pyeritz RE. (1987). The Stickler syndrome: evidence for close linkage to the structural gene for type II collagen. Genomics. (1), 293-6.

Freddi S, Savarirayan R, Bateman JF. (2000). Molecular diagnosis of Stickler syndrome: a COL2A1 stop codon mutation screening strategy that is not compromised by mutant mRNA instability. Am J Med Genet. (90), 398-406.

Freije JM, Diez-Itza I, Balbin M, Sanchez LM, Blasco R, Tolivia J, Lopez-Otin C. (1994). Molecular cloning and expression of collagenase-3, a novel human matrix metalloproteinase produced by breast carcinomas.J Biol Chem. (269), 16766-73.

Gaiser KG, Maddox BK, Bann JG, Boswell BA, Keene DR, Garofalo S, Horton WA. (2002). Y-position collagen II mutation disrupts cartilage formation and skeletal development in a transgenic mouse model of spondyloepiphyseal dysplasia. J Bone Miner Res. (17), 39-47.

Galicka A, Wolczynski S, Gindzienski A. (2003). Studies on type I collagen in skin fibroblasts cultured from twins with lethal osteogenesis imperfecta. Acta Biochim Pol. (50), 481-8.

Garofalo S, Vuorio E, Metsaranta M, Rosati R, Toman D, Vaughan J, Lozano G, Mayne R, Ellard J, Horton W, et al. (1991). Reduced amounts of cartilage collagen fibrils and growth plate anomalies in transgenic mice harboring a glycine-to-cysteine mutation in the mouse type II procollagen alpha 1-chain gene. Proc Natl Acad Sci USA. (88), 9648-52.

Gething MJ, Sambrook J. (1992). Protein folding in the cell. Nature. (355), 33-45.

Godfrey M, Hollister DW. (1988). Type II achondrogenesis-hypochondrogenesis: identification of abnormal type II collagen. Am J Hum Genet. (43), 904-13.

Godfrey M, Keene DR, Blank E, Hori H, Sakai LY, Sherwin LA, Hollister DW. (1988). Type II achondrogenesis-hypochondrogenesis: morphologic and immunohistopathologic studies. Am J Hum Genet. (43), 894-903.

Gunthard J, Fliegel C, Ohnacker H, Rutishauser M, Buhler E. (1995). Lung hypoplasia and severe pulmonary hypertension in an infant with double heterozygosity for spondyloepiphyseal dysplasia congenita and achondroplasia. Clin Genet. (48), 35-40.

Haas IG, Wabl M. (1983). Immunoglobulin heavy chain binding protein. Nature. (306), 387-9.

Harrod MJ, Friedman JM, Currarino G, Pauli RM, Langer LO Jr. (1984). Genetic heterogeneity in spondyloepiphyseal dysplasia congenita. Am J Med Genet. (18), 311-20.

Hartigan N, Garrigue-Antar L, Kadler KE. (2003). Bone morphogenetic protein-1 (BMP-1). Identification of the minimal domain structure for procollagen C-proteinase activity. J Biol Chem. (278), 18045-9.

Hartl FU, Hlodan R, Langer T. (1994). Molecular chaperones in protein folding: the art of avoiding sticky situations. Trends Biochem Sci. (19), 20-5.

Hashimoto Y, Tomiyama T, Yamano Y, Mori H. (2003). Mutation (D472Y) in the type 3 repeat domain of cartilage oligomeric matrix protein affects its early vesicle trafficking in endoplasmic reticulum and induces apoptosis. Am J Pathol. (163), 101-10.

Hecht JT, Hayes E, Haynes R, Cole WG. (2005). COMP mutations, chondrocyte function and cartilage matrix. Matrix Biol. (23), 525-33

Hecht JT, Makitie O, Hayes E, Haynes R, Susic M, Montufar-Solis D, Duke PJ, Cole WG. (2004). Chondrocyte cell death and intracellular distribution of COMP and type IX collagen in the pseudoachondroplasia growth plate. J Orthop Res. (22), 759-67.

Hecht JT, Montufar-Solis D, Decker G, Lawler J, Daniels K, Duke PJ. (1998). Retention of cartilage oligomeric matrix protein (COMP) and cell death in redifferentiated pseudoachondroplasia chondrocytes. Matrix Biol. (17), 625-33.

Hendershot LM, Bulleid NJ. (2000). Protein-specific chaperones: the role of hsp47 begins to gel. Curr Biol. (10), 912-5.

Hojima Y, McKenzie JA, van der Rest M, Prockop DJ. (1989). Type I procollagen N-proteinase from chick embryo tendons. Purification of a new 500-kDa form of the enzyme and identification of the catalytically active polypeptides. J Biol Chem. (264), 11336-45.

Hojima Y, van der Rest M, Prockop DJ. (1985). Type I procollagen carboxyl-terminal proteinase from chick embryo tendons. Purification and characterization. J Biol Chem. (260), 15996-6003.

Hoornaert PK, Dewinter C, Vereecke I, Beemer FA, Courtens W, Fryer A, Fryssira H, Lees M, Müllner-Eidenböck A, Rimoin D, Siderius L, Superti-Furga A, Temple K, Willems P, Zankl A, Zweier C, De Paepe A, Coucke P, Mortier GR. (2005). The phenotypic spectrum in patients with arginine to cysteine mutations in the COL2A1 gene. J Med Genet. In press.

Hulmes DJ. (1992). The collagen superfamily--diverse structures and assemblies. Essays Biochem. (27), 49-67.

Ito H, Rucker E, Steplewski A, McAdams E, Brittingham RJ, Alabyeva T, Fertala A. (2005). Guilty by association: some collagen II mutants alter the formation of ECM as a result of atypical interaction with fibronectin. J Mol Biol. (352), 382-95.

Jakkula E, Makitie O, Czarny-Ratacjzak M, Jackson GC, Damignani R, Susic M, Briggs MD, Cole WG, Ala-Kokko L. (2005). Mutations in the known genes are not the major cause of MED; distinctive phenotypic entities among patients with no identified mutations. Eur J Hum Genet. (13), 292-301.

John DC, Grant ME, Bulleid NJ. (1993). Cell-free synthesis and assembly of prolyl 4-hydroxylase: the role of the beta-subunit (PDI) in preventing misfolding and aggregation of the alpha-subunit. EMBO J. (12), 1587-95.

Kadler KE, Hojima Y, Prockop DJ. (1987). Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process. J Biol Chem. (262), 15696-701.

Kagan HM, Trackman PC. (1991). Properties and function of lysyl oxidase. Am J Respir Cell Mol Biol. (5), 206-10.

Kaufman RJ. (1999). Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. Genes Dev. (13), 1211-33. Kellgren JH, Lawrence JS, Bier F. (1963). Genetic factors in generalized osteoarthrosis. Ann Rheum Dis. (22), 237-55.

Kessler E, Takahara K, Biniaminov L, Brusel M, Greenspan DS. (1996). Bone morphogenetic protein-1: the type I procollagen C-proteinase. Science. (271), 360-2.

Kivirikko KI, Myllyharju J. (1998). Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. Matrix Biol. (16), 357-68.

Kivirikko KI. (1993). Collagens and their abnormalities in a wide spectrum of diseases. Ann Med. (25), 113-26.

.

Kivirikko KI & Myllylä R. (1979). Collagen glycosyltransferases. Int Rev Connect Tissue Res 8: 23-72.

Kivirikko KI. (1995). Posttranslational processing of collagens. In: Bittar EE & Bittar N (ed) Principles of Medical Biology. JAI Press, London, England, p 233-254.

Kizawa H, Kou I, Iida A, Sudo A, Miyamoto Y, Fukuda A, Mabuchi A, Kotani A, Kawakami A, Yamamoto S, Uchida A, Nakamura K, Notoya K, Nakamura Y, Ikegawa S. (2005). An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. Nat Genet. (37), 138-44

Knauper V, Lopez-Otin C, Smith B, Knight G, Murphy G. (1996). Biochemical characterization of human collagenase-3. J Biol Chem. (271), 1544-50.

Knowlton RG, Katzenstein PL, Moskowitz RW, Weaver EJ, Malemud CJ, Pathria MN, Jimenez SA, Prockop DJ. (1990). Genetic linkage of a polymorphism in the type II procollagen gene (COL2A1) to primary osteoarthritis associated with mild chondrodysplasia. N Engl J Med. (322), 526-30.

Knowlton RG, Weaver EJ, Struyk AF, Knobloch WH, King RA, Norris K, Shamban A, Uitto J, Jimenez SA, Prockop DJ. (1989). Genetic linkage analysis of hereditary arthro-ophthalmopathy (Stickler syndrome) and the type II procollagen gene. Am J Hum Genet. (45), 681-8.

Kohfeldt E, Maurer P, Vannahme C, Timpl R. (1997). Properties of the extracellular calcium binding module of the proteoglycan testican. FEBS Lett. (414), 557-61.

Kuivaniemi H, Tromp G, Prockop DJ. (1991). Mutations in collagen genes: causes of rare and some common diseases in humans. FASEB J. (5), 2052-60.

Kuivaniemi H, Tromp G, Prockop DJ. (1997). Mutations in fibrillar collagens (types I, II, III, and XI), fibril-associated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels. Hum Mutat. (9), 300-15.

Langer LO Jr, Spranger JW, Greinacher I, Herdman RC. (1969). Thanatophoric dwarfism. A condition confused with achondroplasia in the neonate, with brief comments on achondrogenesis and homozygous achondroplasia. Radiology. (92), 285-94.

Lee AH, Iwakoshi NN, Anderson KC, Glimcher LH. (2003). Proteasome inhibitors disrupt the unfolded protein response in myeloma cells. Proc Natl Acad Sci. (100), 9946-51.

Lee B, Vissing H, Ramirez F, Rogers D, Rimoin D. (1989). Identification of the molecular defect in a family with spondyloepiphyseal dysplasia. Science. (244), 978-80.

Leung MK, Fessler LI, Greenberg DB, Fessler JH. (1979). Separate amino and carboxyl procollagen peptidases in chick embryo tendon. J Biol Chem. (254), 224-32.

Li SW, Sieron AL, Fertala A, Hojima Y, Arnold WV, Prockop DJ. (1996). The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenic protein-1. Proc Natl Acad Sci USA. (93), 5127-30.

Lightfoot SJ, Holmes DF, Brass A, Grant ME, Byers PH, Kadler KE. (1992). Type I procollagens containing substitutions of aspartate, arginine, and cysteine for glycine in the pro alpha 1 (I) chain are cleaved slowly by N-proteinase, but only the cysteine substitution introduces a kink in the molecule. J Biol Chem. (267), 25521-8.

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), 227-37.

Loughlin J, Dowling B, Chapman K, Marcelline L, Mustafa Z, Southam L, Ferreira A, Ciesielski C, Carson DA, Corr M. (2004). Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females. Proc Natl Acad Sci. (101), 9757-62.

Majsterek I, McAdams E, Adachi E, Dhume ST, Fertala A. (2003). Prospects and limitations of the rational engineering of fibrillar collagens. Protein Sci. (12), 2063-72.

Mann HH, Ozbek S, Engel J, Paulsson M, Wagener R. (2004). Interactions between the cartilage oligomeric matrix protein and matrilins. Implications for matrix assembly and the pathogenesis of chondrodysplasias. J Biol Chem. (279), 25294-8.

Maroteaux P, Stanescu V, Stanescu R. (1983). Hypochondrogenesis. Eur J Pediatr. (141), 14-22.

Martin SJ, Green DR. (1995). Protease activation during apoptosis: death by a thousand cuts? Cell. (82), 349-52.

Meusser B, Hirsch C, Jarosch E, Sommer T. (2005). ERAD: the long road to destruction. Nat Cell Biol. (7), 766-72.

Miller EJ, Harris ED Jr, Chung E, Finch JE Jr, McCroskery PA, Butler WT. (1976). Cleavage of Type II and III collagens with mammalian collagenase: site of cleavage and primary structure at the NH2-terminal portion of the smaller fragment released from both collagens. Biochemistry. (15), 787-92.

Mironov AA, Mironov AA Jr, Beznoussenko GV, Trucco A, Lupetti P, Smith JD, Geerts WJ, Koster AJ, Burger KN, Martone ME, Deerinck TJ, Ellisman MH, Luini A. (2003). ER-to-Golgi carriers arise through direct en bloc protrusion and multistage maturation of specialized ER exit domains. Dev Cell. (5), 583-94.

Mironov AA, Weidman P, Luini A. (1997). Variations on the intracellular transport theme: maturing cisternae and trafficking tubules. J Cell Biol. (138), 481-4.

Mitchell PG, Magna HA, Reeves LM, Lopresti-Morrow LL, Yocum SA, Rosner PJ, Geoghegan KF, Hambor JE. (1996). Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage. J Clin Invest. (97), 761-8.

Mizuno K, Hayashi T, Bachinger HP. (2003). Hydroxylation-induced stabilization of the collagen triple helix. Further characterization of peptides with 4(R)-hydroxyproline in the Xaa position. J Biol Chem. (278), 32373-9.

Mortier GR, Weis M, Nuytinck L, King LM, Wilkin DJ, De Paepe A, Lachman RS, Rimoin DL, Eyre DR, Cohn DH. (2000). Report of five novel and one recurrent COL2A1 mutations with analysis of genotype-phenotype correlation in patients with a lethal type II collagen disorder. J Med Genet. (37), 263-71.

Mulugeta S, Nguyen V, Russo SJ, Muniswamy M, Beers MF. (2005). A surfactant protein C precursor protein BRICHOS domain mutation causes endoplasmic reticulum stress, proteasome dysfunction, and caspase 3 activation. Am J Respir Cell Mol Biol. (32), 521-30.

Mundlos S, Chan D, McGill J, Bateman JF. (1996). An alpha 1(II) Gly913 to Cys substitution prevents the matrix incorporation of type II collagen which is replaced with type I and III collagens in cartilage from a patient with hypochondrogenesis. Am J Med Genet. (63), 129-36.

Mundlos S, Olsen BR. (1997). Heritable diseases of the skeleton. Part II: Molecular insights into skeletal development-matrix components and their homeostasis. FASEB J. (11), 227-33.

Murray LW, Bautista J, James PL, Rimoin DL. (1989). Type II collagen defects in the chondrodysplasias. I. Spondyloepiphyseal dysplasias. Am J Hum Genet. (45), 5-15.

Myllyharju J, Kivirikko KI. (2001). Collagens and collagen-related diseases. Ann Med. (33), 7-21.

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

Nagata K, Hosokawa N. (1996). Regulation and function of collagen-specific molecular chaperone, HSP47. Cell Struct Funct. (21), 425-30.

Nagata K. (2003). HSP47 as a collagen-specific molecular chaperone: function and expression in normal mouse development. Semin Cell Dev Biol. (14), 275-82.

Nah HD, Upholt WB. (1991). Type II collagen mRNA containing an alternatively spliced exon predominates in the chick limb prior to chondrogenesis. J Biol Chem. (266), 23446-52.

Naqvi T, Duong TT, Hashem G, Shiga M, Zhang Q, Kapila S. (2005). Relaxin's induction of metalloproteinases is associated with the loss of collagen and glycosaminoglycans in synovial joint fibrocartilaginous explants. Arthritis Res Ther. (7), 1-11.

Narcisi P, Wu Y, Tromp G, Earley JJ, Richards AJ, Pope FM, Kuivaniemi H. (1993). Single base mutation that substitutes glutamic acid for glycine 1021 in the COL3A1 gene and causes Ehlers-Danlos syndrome type IV. Am J Med Genet. (46), 278-83.

Nelson DL, Cox MM. (2005). Lehninger Principles of Biochemistry. WH Freeman, New York, USA, (Fourth edn).

Noiva R, Lennarz WJ. (1992). Protein disulfide isomerase. A multifunctional protein resident in the lumen of the endoplasmic reticulum. J Biol Chem. (267), 3553-6.

Notbohm H, Nokelainen M, Myllyharju J, Fietzek PP, Muller PK, Kivirikko KI. (1999). Recombinant human type II collagens with low and high levels of hydroxylysine and its glycosylated forms show marked differences in fibrillogenesis in vitro. J Biol Chem. (274), 8988-92.

Nuytinck L, Freund M, Lagae L, Pierard GE, Hermanns-Le T, De Paepe A. (2000). Classical Ehlers-Danlos syndrome caused by a mutation in type I collagen. Am J Hum Genet. (66), 1398-402.

Olague-Marchan M, Twining SS, Hacker MK, McGrath JA, Diaz LA, Giudice GJ. (2000). A disease-associated glycine substitution in BP180 (type XVII collagen) leads to a local destabilization of the major collagen triple helix. Matrix Biol. (19), 223-33.

Olsen BR, Hoffmann H, Prockop DJ. (1976). Interchain disulfide bonds at the COOH-terminal end of procollagen synthesized by matrix-free cells from chick embryonic tendon and cartilage. Arch Biochem Biophys. (175), 341-50.

Oyadomari S, Mori M. (2004). Roles of CHOP/GADD153 in endoplasmic reticulum stress. Cell Death Differ. (11), 381-9.

Palotie A, Vaisanen P, Ott J, Ryhanen L, Elima K, Vikkula M, Cheah K, Vuorio E, Peltonen L. (1989). Predisposition to familial osteoarthrosis linked to type II collagen gene. Lancet. (1), 924-7.

Persikov AV, Ramshaw JA, Brodsky B. (2005). Prediction of collagen stability from amino acid sequence. J Biol Chem. (280), 19343-9.

Pihlajaniemi T, Rehn M. (1995). Two new collagen subgroups: membrane-associated collagens and types XV and XVII. Prog Nucleic Acid Res Mol Biol. (50), 225-62.

Poole AR. (2001). Cartilage in health and disease. In Arthritis and Allied Conditions. A Textbook of Rheumatology, 14th edn. Edited by Koopman W. New York: Lippincott Williams and Wilkins. 2260-2284.

Prockop D, Hulmes DJS. (1994). Assemply of collagen fibrils de Novo from soluble precursors: polymerization and copolymerization of procollagen, pN-Collagen, and mutated collagens. In: Yurchenco PD, Birk DE, & Mecham RP (ed.) Extracellular Matrix Assemply and Structure. Academic Press, New York, 47-91.

Prockop DJ, Kivirikko KI. (1995). Collagens: molecular biology, diseases, and potentials for therapy. Annu Rev Biochem. (64), 403-34.

Raghunath M, Bruckner P, Steinmann B. (1994). Delayed triple helix formation of mutant collagen from patients with osteogenesis imperfecta. J Mol Biol. (236), 940-9.

Rao RV, Ellerby HM, Bredesen DE. (2004). Coupling endoplasmic reticulum stress to the cell death program. Cell Death Differ. (11), 372-80.

Reginato AM, Olsen BR. (2002). The role of structural genes in the pathogenesis of osteoarthritic disorders. Arthritis Res. (4), 337-45.

Richards AJ, Baguley DM, Yates JR, Lane C, Nicol M, Harper PS, Scott JD, Snead MP. (2000a). Variation in the vitreous phenotype of Stickler syndrome can be caused by different amino acid substitutions in the X position of the type II collagen Gly-X-Y triple helix. Am J Hum Genet. (67), 1083-94.

Richards AJ, Martin S, Yates JR, Scott JD, Baguley DM, Pope FM, Snead MP. (2000b). COL2A1 exon 2 mutations: relevance to the Stickler and Wagner syndromes. Br J Ophthalmol. (84), 364-71.

Richards AJ, Scott JD, Snead MP. (2002). Molecular genetics of rhegmatogenous retinal detachment. Eye. (16), 388-92.

Richards AJ, Yates JR, Williams R, Payne SJ, Pope FM, Scott JD, Snead MP. (1996). A family with Stickler syndrome type 2 has a mutation in the COL11A1 gene resulting in the substitution of glycine 97 by valine in alpha 1 (XI) collagen. Hum Mol Genet. (5), 1339-43.

Ritvaniemi P, Hyland J, Ignatius J, Kivirikko KI, Prockop DJ, Ala-Kokko L. (1993). A fourth example suggests that premature termination codons in the COL2A1 gene are a common cause of the Stickler syndrome: analysis of the COL2A1 gene by denaturing gradient gel electrophoresis. Genomics. (17), 218-21.

Royce PM, Steinmann B. (2002). Connective Tissue and Its Heritable Disorders: Molecular, Genetic, and Medical Aspects (2nd ed.), Wiley.

Rutkowski DT, Kaufman RJ. (2004). A trip to the ER: coping with stress. Trends Cell Biol. (14), 20-8.

Ryan MC, Sandell LJ. (1990). Differential expression of a cysteine-rich domain in the amino-terminal propeptide of type II (cartilage) procollagen by alternative splicing of mRNA. J Biol Chem. (265), 10334-9.

Saga S, Nagata K, Chen WT, Yamada KM. (1987). pH-dependent function, purification, and intracellular location of a major collagen-binding glycoprotein. J Cell Biol. (105), 517-27.

Saldino RM. (1971). Lethal short-limbed dwarfism: achondrogenesis and thanatophoric dwarfism. Am. J. Roentgen. (112), 185-197.

Salminen HJ, Saamanen AM, Vankemmelbeke MN, Auho PK, Perala MP, Vuorio EI. (2002). Differential expression patterns of matrix metalloproteinases and their inhibitors during development of osteoarthritis in a transgenic mouse model. Ann Rheum Dis. (61), 591-7.
Salvesen GS, Dixit VM. (1997). Caspases: intracellular signaling by proteolysis. Cell. (91), 443-6.

Sandell LJ, Morris N, Robbins JR, Goldring MB. (1991). Alternatively spliced type II procollagen mRNAs define distinct populations of cells during vertebral development: differential expression of the amino-propeptide. J Cell Biol. (114), 1307-19.

Satoh M, Hirayoshi K, Yokota S, Hosokawa N, Nagata K. (1996). Intracellular interaction of collagen-specific stress protein HSP47 with newly synthesized procollagen. J Cell Biol. (133), 469-83.

Schiedner G, Hertel S, Kochanek S. (2000). Efficient transformation of primary human amniocytes by E1 functions of Ad5: generation of new cell lines for adenoviral vector production. Hum Gene Ther. (11), 2105-16.

Scott IC, Blitz IL, Pappano WN, Imamura Y, Clark TG, Steiglitz BM, Thomas CL, Maas SA, Takahara K, Cho KW, Greenspan DS. (1999). Mammalian BMP-1/Tolloid-related metalloproteinases, including novel family member mammalian Tolloid-like 2, have differential enzymatic activities and distributions of expression relevant to patterning and skeletogenesis. Dev Biol. (213), 283-300.

Sitia R, Braakman I. (2003). Quality control in the endoplasmic reticulum protein factory. Nature. (426), 891-4.

Snead MP, Payne SJ, Barton DE, Yates JR, al-Imara L, Pope FM, Scott JD. (1994). Stickler syndrome: correlation between vitreoretinal phenotypes and linkage to COL 2A1. Eye. (8), 609-14.

Snead MP, Yates JR. (1999). Clinical and Molecular genetics of Stickler syndrome. J Med Genet. (36), 353-9.

Sobetzko D, Eich G, Kalff-Suske M, Grzeschik KH, Superti-Furga A. (2000). Boy with syndactylies, macrocephaly, and severe skeletal dysplasia: not a new syndrome, but two dominant mutations (GLI3 E543X and COL2A1 G973R) in the same individual. Am J Med Genet. (90), 239-42.

Song F, Windsor LJ. (2005). Novel nonmatrix-metalloproteinase-mediated collagen degradation. Biochim Biophys Acta. (1721), 65-72.

Sood R, Porter AC, Ma K, Quilliam LA, Wek RC. (2000). Pancreatic eukaryotic initiation factor-2alpha kinase (PEK) homologues in humans, Drosophila melanogaster and Caenorhabditis elegans that mediate translational control in response to endoplasmic reticulum stress. Biochem J. (346), 281-93.

Spranger J, Winterpacht A, Zabel B. (1994). The type II collagenopathies: a spectrum of chondrodysplasias. Eur J Pediatr. (153), 56-65.

Stecher RM, Hersh AH, Hauser H. (1953). Heberden's nodes; the family history and radiographic appearance of a large family. Am J Hum Genet. (5), 46-60.

Stefansson SE, Jonsson H, Ingvarsson T, Manolescu I, Jonsson HH, Olafsdottir G, Palsdottir E, Stefansdottir G, Sveinbjornsdottir G, Frigge ML, Kong A, Gulcher JR, Stefansson K. (2003). Genomewide scan for hand osteoarthritis: a novel mutation in matrilin-3. Am J Hum Genet. (72), 1448-59.

Steplewski A, Brittingham R, Jimenez SA, Fertala A. (2005). Single amino acid substitutions in the C-terminus of collagen II alter its affinity for collagen IX. Biochem Biophys Res Commun. (335), 749-55.

Steplewski A, Ito H, Rucker E, Brittingham RJ, Alabyeva T, Gandhi M, Ko FK, Birk DE, Jimenez SA, Fertala A. (2004b). Position of single amino acid substitutions in the collagen triple helix determines their effect on structure of collagen fibrils. J Struct Biol. (148), 326-37.

Steplewski A, Majsterek I, McAdams E, Rucker E, Brittingham RJ, Ito H, Hirai K, Adachi E, Jimenez SA, Fertala A. (2004a). Thermostability gradient in the collagen triple helix reveals its multi-domain structure. J Mol Biol. (338), 989-98.

Stroh C, Schulze-Osthoff K. (1998). Death by a thousand cuts: an ever increasing list of caspase substrates. Cell Death Differ. (5), 997-1000.

Takahashi E, Hori T, O'Connell P, Leppert M, White R. (1990). R-banding and nonisotopic in situ hybridization: precise localization of the human type II collagen gene (COL2A1). Hum Genet. (86), 14-6.

Tasanen K, Eble JA, Aumailley M, Schumann H, Baetge J, Tu H, Bruckner P, Bruckner-Tuderman L. (2000). Collagen XVII is destabilized by a glycine substitution mutation in the cell adhesion domain Col15. J Biol Chem. (275), 3093-9.

Tchetina EV, Squires G, Poole AR. (2005). Increased type II collagen degradation and very early focal cartilage degeneration is associated with upregulation of chondrocyte differentiation related genes in early human articular cartilage lesions. J Rheumatol. (32), 876-86.

Temple IK. (1989). Stickler's syndrome. J Med Genet. (26), 119-26.

Tiller GE, Weis MA, Polumbo PA, Gruber HE, Rimoin DL, Cohn DH, Eyre DR. (1995). An RNA-splicing mutation (G+5IVS20) in the type II collagen gene (COL2A1) in a family with spondyloepiphyseal dysplasia congenita. Am J Hum Genet. (56), 388-95.

Vandenberg P, Khillan JS, Prockop DJ, Helminen H, Kontusaari S, Ala-Kokko L. (1991). Expression of a partially deleted gene of human type II procollagen (COL2A1) in transgenic mice produces a chondrodysplasia. Proc Natl Acad Sci USA. (88), 7640-4.

Vikkula M, Metsaranta M, Ala-Kokko L. (1994). Type II collagen mutations in rare and common cartilage diseases. Ann Med. (26), 107-14.

Vintiner GM, Temple IK, Middleton-Price HR, Baraitser M, Malcolm S. (1991). Genetic and clinical heterogeneity of Stickler syndrome. Am J Med Genet. (41), 44-8.

Vissing H, D'Alessio M, Lee B, Ramirez F, Godfrey M, Hollister DW. (1989). Glycine to serine substitution in the triple helical domain of pro-alpha 1 (II) collagen results in a lethal perinatal form of short-limbed dwarfism. J Biol Chem. (264), 18265-7.

Vogel BE, Doelz R, Kadler KE, Hojima Y, Engel J, Prockop DJ. (1988). A substitution of cysteine for glycine 748 of the alpha 1 chain produces a kink at this site in the procollagen I molecule and an altered N-proteinase cleavage site over 225 nm away. J Biol Chem. (263), 19249-55.

Vranka J, Mokashi A, Keene DR, Tufa S, Corson G, Sussman M, Horton WA, Maddox K, Sakai L, Bachinger HP. (2001). Selective intracellular retention of extracellular matrix proteins and chaperones associated with pseudoachondroplasia. Matrix Biol. (20), 439-50.

Wagner K, Poschl E, Turnay J, Baik J, Pihlajaniemi T, Frischholz S, von der Mark K. (2000). Coexpression of alpha and beta subunits of prolyl 4-hydroxylase stabilizes the triple helix of recombinant human type X collagen. Biochem J. (352), 907-11.

Wilkin DJ, Mortier GR, Johnson CL, Jones MC, de Paepe A, Shohat M, Wildin RS, Falk RE, Cohn DH. (1998). Correlation of linkage data with phenotype in eight families with Stickler syndrome. Am J Med Genet. (80), 121-7.

Williams CJ, Considine EL, Knowlton RG, Reginato A, Neumann G, Harrison D, Buxton P, Jimenez S, Prockop DJ. (1993). Spondyloepiphyseal dysplasia and precocious osteoarthritis in a family with an Arg75-->Cys mutation in the procollagen type II gene (COL2A1). Hum Genet. (92), 499-505.

Williams CJ, Ganguly A, Considine E, McCarron S, Prockop DJ, Walsh-Vockley C, Michels VV. (1996). A-2-->G transition at the 3' acceptor splice site of IVS17 characterizes the COL2A1 gene mutation in the original Stickler syndrome kindred. Am J Med Genet. (63), 461-7.

Wilson R, Freddi S, Chan D, Cheah KS, Bateman JF. (2005). Misfolding of collagen X chains harboring Schmid metaphyseal chondrodysplasia mutations results in aberrant disulfide bond formation, intracellular retention, and activation of the unfolded protein response. J Biol Chem. (280), 15544-52.

Wilson R, Lees JF, Bulleid NJ. (1998). Protein disulfide isomerase acts as a molecular chaperone during the assembly of procollagen. J Biol Chem. (273), 9637-43.

Wordsworth P, Ogilvie D, Priestley L, Smith R, Wynne-Davies R, Sykes B. (1988). Structural and segregation analysis of the type II collagen gene (COL2A1) in some heritable chondrodysplasias. J Med Genet. (25), 521-7.

Wu JJ, Eyre DR. (1995). Structural analysis of cross-linking domains in cartilage type XI collagen. Insights on polymeric assembly. J Biol Chem. (270), 18865-70.

Wu W, Billinghurst RC, Pidoux I, Antoniou J, Zukor D, Tanzer M, Poole AR. (2002). Sites of collagenase cleavage and denaturation of type II collagen in aging and osteoarthritic articular cartilage and their relationship to the distribution of matrix metalloproteinase 1 and matrix metalloproteinase 13. Arthritis Rheum. (46), 2087-94.

Wuttke M, Muller S, Nitsche DP, Paulsson M, Hanisch FG, Maurer P. (2001). Structural characterization of human recombinant and bone-derived bone sialoprotein. Functional implications for cell attachment and hydroxyapatite binding. J. Biol. Chem. (276), 36839-36848.

Yamada Y, Avvedimento VE, Mudryj M, Ohkubo H, Vogeli G, Irani M, Pastan I, de Crombrugghe B. (1980). The collagen gene: evidence for its evolutinary assembly by amplification of a DNA segment containing an exon of 54 bp. Cell. (22), 887-92.

Yang L, Carlson SG, McBurney D, Horton WE Jr. (2005). Multiple signals induce endoplasmic reticulum stress in both primary and immortalized chondrocytes resulting in loss of differentiation, impaired cell growth, and apoptosis. J Biol Chem. (280), 31156-65.

Yoshida H, Matsui T, Hosokawa N, Kaufman RJ, Nagata K, Mori K. (2003). A time-dependent phase shift in the mammalian unfolded protein response. Dev Cell. (4), 265-71.

Yoshida H, Matsui T, Yamamoto A, Okada T, and Mori K. (2001). XBP-1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. Cell (107), 881–891.

Zhang K, Kaufman RJ. (2004). Signaling the unfolded protein response from the endoplasmic reticulum. J Biol Chem. (279), 25935-8.

Zhuang J, Tromp G, Kuivaniemi H, Nakayasu K, Prockop DJ. (1993). Deletion of 19 base pairs in intron 13 of the gene for the pro alpha 2(I) chain of type-I procollagen (COL1A2) causes exon skipping in a proband with type-I osteogenesis imperfecta. Hum Genet. (91), 210-6.

# Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig und ohne unzulässige Hilfe angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit, die anderen Werken in Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass die Dissertation noch keiner anderen Fakultät oder Universität vorgelegt und noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Ablauf des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der geltenden Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Herrn Prof. Dr. Mats Paulsson betreut worden.

Köln, 28. September 2005

# Acknowledgements

I wish to express my deepest gratitude to my supervisor Professor Mats Paulsson for giving me the opportunity to prepare my thesis as a member of his research group. His knowledge, never-ending optimism and support have been an endless source of inspiration to me.

I owe my heart-felt gratitude to Dr. Frank Zaucke for his diligent guidance and methodological advice, which made this work possible. His instructive comments and meticulous editing of this thesis is greatly acknowledged. I will always remember him for helping me realize the importance of planning and keeping up with the schedules.

My sincere thanks are due to all the group leaders of Biochemie II for their invaluable suggestions, thorough discussions and constructive criticism throughout my work.

I also wish to express my special thanks to Dr. Matthias Morgelin and his staff, University of Lund, Sweden, for expert guidance in electron microscopy.

I am thankful to Prof. Maria Leptin and Prof. Siegfried Roth for providing the opportunity to perform my laboratory rotations and gain invaluable scientific experience as part of their group.

I would like to thank my group members Markus, Beecker, Christiane for providing an exciting atmosphere with relaxing scientific and non-scientific conversations, special thanks to Katrin for her valuable comments and suggestions in the preparation of the manuscript of this thesis.

I wish to extend my sincere thanks to all the staff and colleagues of Biochemie II. Especially, I wish to acknowledge Gerda for her help in smoothening out tedious administrative tasks and would like to mention my appreciation for Johanna (Hanni) for her technical advice.

I would like to thank Dr. Stefan Müller, Dr. Udo Roth and other members of the service facility of the Centre for Molecular Medicine, for their help in DNA sequencing and mass spectroscopy.

I am deeply grateful to all my friends outside the scientific world for giving me a chance to think about other issues than science.

My parents, my sister deserve my special thanks for their love, care, encouraging attitude and their unwavering support for all my endeavors.

No words are sufficient to express my love and gratitude for Suchitra (Guddi) for being my pillar of strength at all the times.

The spirit of Cologne, Carnival and Christmas parties will always be a part of my fond memories- *viva colonia.*

International graduate school in genetics and functional genomics, University of Cologne, is highly acknowledged for providing the opportunity and the financial support for this work.

# Lebenslauf

## Persönliche Daten

| | |
|---|---|
| Name | Chakkalakal Anandan Salin |
| Anschrift | Berrenrather Str, 133, 50937 Köln |
| Heimataddresse | 442, 15th-D cross, Mahalaxmipuram Bangalore, 560086, India |
| Geburtsdatum/-ort | 17.11.1974 / Trichur, India |
| Staatsangehörigkeit | Indisch |

## Schulbildung

| | |
|---|---|
| 1980-1987 | Deccan English School, Bangalore, India |
| 1987-1990 | Sheshadripuram High School, Bangalore, India |
| 1990-1992 | Sheshadripuram P U College, Bangalore, India |

## Studium (Universität)

| | |
|---|---|
| 1992-1996 | B. Pharm, Bangalore University, K L E College of Pharmacy, Bangalore, India |
| 1996-1998 | M. Tech (Biotechnology), Anna University Centre for Biotechnology, Chennai, India |

## Doktorarbeit

| | |
|---|---|
| Oktober 2001- Dezember 2005 | Promotion an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln im Fach Physiologische Chemie am Institut für Biochemie II der Medizinischen Fakultät bei Prof. Dr. Mats Paulsson |

Köln, 28. September 2005

# Appendix

## a. Collagen II cDNA sequence used in the present study

The codons which are marked in red represent arginines that have been exchanged to cysteines in the present study.

ATGATTCGCCTCGGGGCTCCCCAGTCGCTGGTGCTGCTGACGCTGCTCGTCGCCGCTGTCCTTC
GGTGTCAGGGCCAGGATGTCCGGCAACCAGGACCAAAGGGACAGAAAGGAGAACCTGGAGACAT
CAAGGATATTGTAGGACCCAAAGGACCTCCTGGGCCTCAGGGACCTGCAGGGGAACAAGGACCC
AGAGGGGATCGTGGTGACAAAGGTGAAAAAGGTGCCCCTGGACCTCGTGGCAGAGATGGAGAAC
CTGGGACCCCTGGAAATCCTGGCCCCCCTGGTCCTCCCGGCCCCCCTGGTCCCCCTGGTCTTGG
TGGAAACTTTGCTGCCCAGATGGCTGGAGGATTTGATGAAAAGGCTGGTGGCGCCCAGTTGGGA
GTAATGCAAGGACCAATGGGCCCCATGGGACCTCGAGGACCTCCAGGCCCTGCAGGTGCTCCTG
GGCCTCAAGGATTTCAAGGCAATCCTGGTGAACCTGGTGAACCTGGTGTCTCTGGTCCCATGGG
TCCCCGTGGTCCTCCTGGTCCCCCTGGAAAGCCTGGTGATGATGGTGAAGCTGGAAAACCTGGA
AAAGCTGGTGAAAGGGGTCCGCCTGGTCCTCAGGGTGCTCGTGGTTTCCCAGGAACCCCAGGCC
TTCCTGGTGTCAAAGGTCACAGAGGTTATCCAGGCCTGGACGGTGCTAAGGGAGAGGCGGGTGC
TCCTGGTGTGAAGGGTGAGAGTGGTTCCCCGGGTGAGAACGGATCTCCGGGCCCAATGGGTCCT
CGTGGCCTGCCTGGTGAAAGAGGACGGACTGGCCCTGCTGGCGCTGCGGGTGCCCGAGGCAACG
ATGGTCAGCCAGGCCCCGCAGGTCCTCCGGGTCCTGTCGGTCCTGCTGGTGGTCCTGGCTTCCC
TGGTGCTCCTGGAGCCAAGGGTGAAGCCGGCCCCACTGGTGCCCGTGGTCCTGAAGGTGCTCAA
GGTCCTCGCGGTGAACCTGGTACTCCTGGGTCCCCTGGGCCTGCTGGTGCCTCCGGTAACCCTG
GAACAGATGGAATTCCTGGAGCCAAAGGATCTGCTGGTGCTCCTGGCATTGCTGGTGCTCCTGG
CTTCCCTGGGCCACGGGGTCCTCCTGGCCCTCAAGGTGCAACTGGTCCTCTGGGCCCGAAAGGT
CAGACGGGTGAACCTGGTATTGCTGGCTTCAAAGGTGAACAAGGCCCCAAGGGAGAACCTGGCC
CTGCTGGCCCCCAGGGAGCCCCTGGACCCGCTGGTGAAGAAGGCAAGAGAGGTGCCCGTGGAGA
GCCTGGTGGCGTTGGGCCCATCGGTCCCCCTGGAGAAAGAGGTGCTCCCGGAAACCGCGGTTTC
CCAGGTCAAGATGGTCTGGCAGGTCCCAAGGGAGCCCCTGGAGAGCGAGGGCCCAGTGGTCTTG
CTGGCCCCAAGGGAGCCAACGGTGACCCTGGCCGTCCTGGAGAACCTGGCCTTCCTGGAGCCCG
GGGTCTCACTGGCCGCCCTGGTGATGCTGGTCCTCAAGGCAAAGTTGGCCCTTCTGGAGCCCCT
GGTGAAGATGGTCGTCCTGGACCTCCAGGTCCTCAGGGGGCTCGTGGGCAGCCTGGTGTCATGG
GTTTCCCTGGCCCCAAAGGTGCCAACGGTGAGCCTGGCAAAGCTGGTGAGAAGGGACTGCCTGG
TGCTCCTGGTCTGAGGGGTCTTCCTGGCAAAGATGGTGAGACAGGTGCTGCAGGACCCCCTGGC
CCTGCTGGACCTGCTGGTGAACGAGGCGAGCAGGGTGCTCCTGGGCCATCTGGGTTCCAGGGAC
TTCCTGGCCCTCCTGGTCCCCCAGGTGAAGGTGGAAAACCAGGTGACCAGGGTGTTCCCGGTGA
AGCTGGAGCCCCTGGCCTCGTGGGTCCCAGGGGTGAACGAGGTTTCCCAGGTGAACGTGGCTCT
CCCGGTGCCCAGGGCCTCCAGGGTCCCCGTGGCCTCCCCGGCACTCCTGGCACTGATGGTCCCA
AAGGTGCATCTGGCCCAGCAGGCCCCCCTGGCGCACAGGGCCCTCCAGGTCTTCAGGGAATGCC
TGGCGAGAGGGGAGCAGCTGGTATCGCTGGGCCCAAAGGCGACAGGGGTGACGTTGGTGAGAAA
GGCCCTGAGGGAGCCCCTGGAAAGGATGGTGGACGAGGCCTGACAGGTCCCATTGGCCCCCCTG
GCCCAGCTGGTGCTAACGGCGAGAAGGGAGAAGTTGGACCTCCTGGTCCTGCAGGAAGTGCTGG
TGCTCGTGGCGCTCCGGGTGAACGTGGAGAGACTGGCCCCCCCGGACCAGCGGGATTTGCTGGG

CCTCCTGGTGCTGATGGCCAGCCTGGGGCCAAGGGTGAGCAAGGAGAGGCCGGCCAGAAAGGCG
ATGCTGGTGCCCCTGGTCCTCAGGGCCCCTCTGGAGCACCTGGGCCTCAGGGTCCTACTGGAGT
GACTGGTCCTAAAGGAGCCCGAGGTGCCCAAGGCCCCCCGGGAGCCACTGGATTCCCTGGAGCT
GCTGGCCGCGTTGGACCCCCAGGCTCCAATGGCAACCCTGGACCCCCTGGTCCCCCTGGTCCTT
CTGGAAAAGATGGTCCCAAAGGTGCTCGAGGAGACAGCGGCCCCCCTGGCCGAGCTGGTGAACC
CGGCCTCCAAGGTCCTGCTGGACCCCCTGGCGAGAAGGGAGAGCCTGGAGATGACGGTCCCTCT
GGTGCCGAAGGTCCACCAGGTCCCCAGGGTCTGGCTGGTCAGAGAGGCATCGTCGGTCTGCCTG
GGCAACGTGGTGAGAGAGGATTCCCTGGCTTGCCTGGCCCATCGGGTGAGCCCGGCAAGCAGGG
TGCTCCTGGAGCATCTGGAGACAGAGGTCCTCCTGGCCCCGTGGGTCCTCCTGGCCTGACGGGT
CCTGCAGGTGAACCCGGACGAGAGGGAAGCCCCGGTGCTGATGGCCCCCCTGGCAGAGATGGCG
CTGCTGGAGTCAAGGGTGATCGTGGTGAGACTGGTGCTGTGGGAGCTCCTGGAGCCCCTGGGCC
CCCTGGCTCCCCTGGCCCCGCTGGTCCAACTGGCAAGCAAGGAGACAGAGGAGAAGCTGGTGCA
CAAGGCCCCATGGGACCCTCAGGACCAGCTGGAGCCCGGGGAATCCAGGGTCCTCAAGGCCCCA
GAGGTGACAAAGGAGAGGCTGGAGAGCCTGGCGAGAGAGGCCTGAAGGGACACCGTGGCTTCAC
TGGTCTGCAGGGTCTGCCCGGCCCTCCTGGTCCTTCTGGAGACCAAGGTGCTTCTGGTCCTGCT
GGTCCTTCTGGCCCTAGAGGTCCTCCTGGCCCCGTCGGTCCCTCTGGCAAAGATGGTGCTAATG
GAATCCCTGGCCCCATTGGGCCTCCTGGTCCCCGTGGACGATCAGGCGAAACCGGTCCTGCTGG
TCCTCCTGGAAATCCTGGGCCCCCTGGTCCTCCAGGTCCCCCTGGCCCTGGCATCGACATGTCC
GCCTTTGCTGGCTTAGGCCCGAGAGAGAAGGGCCCCGACCCCCTGCAGTACATGCGGGCCGACC
AGGCAGCCGGTGGCCTGAGACAGCATGACGCCGAGGTGGATGCCACACTCAAGTCCCTCAACAA
CCAGATTGAGAGCATCCGCAGCCCCGAGGGCTCCCGCAAGAACCCTGCTCGCACCTGCAGAGAC
CTGAAACTCTGCCACCCTGAGTGGAAGAGTGGAGACTACTGGATTGACCCCAACCAAGGCTGCA
CCTTGGACGCCATGAAGGTTTTCTGCAACATGGAGACTGGCGAGACTTGCGTCTACCCCAATCC
AGCAAACGTTCCCAAGAAGAACTGGTGGAGCAGCAAGAGCAAGGAGAAGAAACACATCTGGTTT
GGAGAAACCATCAATGGTGGCTTCCATTTCAGCTATGGAGATGACAATCTGGCTCCCAACACTG
CCAACGTCCAGATGACCTTCCTACGCCTGCTGTCCACGGAAGGCTCCCAGAACATCACCTACCA
CTGCAAGAACAGCATTGCCTATCTGGACGAAGCAGCTGGCAACCTCAAGAAGGCCCTGCTCATC
CAGGGCTCCAATGACGTGGAGATCCGGGCAGAGGGCAATAGCAGGTTCACGTACACTGCCCTGA
AGGATGGCTGCACGAAACATACCGGTAAGTGGGGCAAGACTGTTATCGAGTACCGGTCACAGAA
GACCTCACGCCTCCCCATCATTGACATTGCACCCATGGACATAGGAGGGCCCGAGCAGGAATTC
GGTGTGGACATAGGGCCGGTCTGCTTCTTGTAAAAACCTGAACCCAGAAACAACACAATCCGTT
GCAAACCCAAAGGACCCAAGTACTTTCCAATCTCAGTCACTCTAGGACTCTGCACTGAATGGCT
GACCTGACCTGATGTCCATTCATCCCACCCTCTCACAGTTCGGACTTTTCTCCCCTCTCTTTCT
AAGAGACCTGAACTGGGCAGACTGCAAAATAAAATCTCGGTGTTCTATTTATTTATTGTCTTCC
TGTAAGACCTTCGGGTCAAGGCAGAGGCAGGAAACTAACTGGTGTGAGTCAAATGCCCCCTGAG
TGACTGCCCCCAGCCCAGGCCAGAAGACCTCCCTTCAGGTGCCGGGCGCAGGAACTGTGTGTGT
CCTACACAATGGTGCTATTCTGTGTCAAACACCTCTGTATTTTTTAAA

## b. Collagen II protein sequence

The protein sequence of collagen II including, BM40 signal peptide, the hexahistidine motif (yellow) and the my epitope is shown. Arginines which are substituted to cysteines in the present study are marked in red. The triple helical domain is marked by the shaded area.

MRAWIFFLLCLAGRALAAPLVHHHHHHGPLVDVASNEQKLISEEDLASMTGGQQMGRDIEGRGL
AQDVRQPGPKGQKGEPGDIKDIVGPKGPPGPQGPAGEQGPRGDRGDKGEKGAPGPRGRDGEPGT
PGNPGPPGPPGPPGPPGLGGNFAAQMAGGFDEKAGGAQLGVMQGPMGPMGPRGPPGPAGAPGPQ
GFQGNPGEPGEPGVSGPMGPRGPPGPPGKPGDDGEAGKPGKAGERGPPGPQGARGFPGTPGLPG
VKGHRGYPGLDGAKGEAGAPGVKGESGSPGENGSPGPMGPRGLPGERGRTGPAGAAGARGNDGQ
PGPAGPPGPVGPAGGPGFPGAPGAKGEAGPTGARGPEGAQGPRGEPGTPGSPGPAGASGNPGTD
GIPGAKGSAGAPGIAGAPGFPGPRGPPGPQGATGPLGPKGQTGEPGIAGFKGEQGPKGEPGPAG
PQGAPGPAGEEGKRGARGEPGGVGPIGPPGERGAPGNRGFPGQDGLAGPKGAPGERGPSGLAGP
KGANGDPGRPGEPGLPGARGLTGRPGDAGPQGKVGPSGAPGEDGRPGPPGPQGARGQPGVMGFP
GPKGANGEPGKAGEKGLPGAPGLRGLPGKDGETGAAGPPGPAGPAGERGEQGAPGPSGFQGLPG
PPGPPGEGGKPGDQGVPGEAGAPGLVGPRGERGFPGERGSPGAQGLQGPRGLPGTPGTDGPKGA
SGPAGPPGAQGPPGLQGMPGERGAAGIAGPKGDRGDVGEKGPEGAPGKDGGRGLTGPIGPPGPA
GANGEKGEVGPPGPAGSAGARGAPGERGETGPPGPAGFAGPPGADGQPGAKGEQGEAGQKGDAG
APGPQGPSGAPGPQGPTGVTGPKGARGAQGPPGATGFPGAAGRVGPPGSNGNPGPPGPPGPSGK
DGPKGARGDSGPPGRAGEPGLQGPAGPPGEKGEPGDDGPSGAEGPPGPQGLAGQRGIVGLPGQR
GERGFPGLPGPSGEPGKQGAPGASGDRGPRGPPGPVGPPGLTGPAGEPGREGSPGADGPPGRDG
AAGVKGDRGETGAVGAPGAPGPPGSPGPAGPTGKQGDRGEAGAQGPMGPSGPAGARGIQGPQGP
RGDKGEAGEPGERGLKGHRGFTGLQGLPGPPGPSGDQGASGPAGPSGPRGPPGPVGPSGKDGAN
GIPGPIGPPGPRGRSGETGPAGPPGNPGPPGPPGPPGPGIDMSAFAGLGPREKGPDPLQYMRAD
QAAGGLRQHDAEVDATLKSLNNQIESIRSPEGSRKNPARTCRDLKLCHPEWKSGDYWIDPNQGC
TLDAMKVFCNMETGETCVYPNPANVPKKNWWSSKSKEKKHIWFGETINGGFHFSYGDDNLAPNT
ANVQMTFLRLLSTEGSQNITYHCKNSIAYLDEAAGNLKKALLIQGSNDVEIRAEGNSRFTYTAL
KDGCTKHTGKWGKTVIEYRSQKTSRLPIIDIAPMDIGGPEQEFGVDIGPVCFL