INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Novel Mutations of COL3A1 Resulting in Ehlers-Danlos Syndrome Type IV and Their Effect on the Folding of Type III Procollagen

by

Jayne A. Goldstein

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

1998

Approved by _____
(Chairperson of Supervisory Committee)

Program Authorized
to Offer Degree _____

Date _____

Doctoral Dissertation

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 1490 Eisenhower Place, P. O. Box 975, Ann Arbor, MI 48106, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature _____

Date _____

University of Washington

Abstract

Novel Mutations of COL3A1 Resulting in Ehlers-Danlos Syndrome Type IV and Their Effect on the Folding of Type III Procollagen

by Jayne A. Goldstein

Chairperson of the Supervisory Committee: Professor Peter H. Byers
Department of Pathology

Ehlers-Danlos syndrome type IV (EDS IV) is an inherited disease caused by mutation of COL3A1, the gene that encodes type III procollagen. Different mutations produce the disease: point mutations, exon skipping mutations, and genomic insertions and deletions. The effect of these different lesions on type III procollagen, the cells that produce it, and the ultrastructural aggregates of which it is a component are not known.

To better understand the consequences of COL3A1 mutation, protein secretion, protein folding, cell architecture, and dermal collagen fibril formation were examined following the characterization of novel EDS IV-causing mutations. I defined 15 mutations: skipping of exons 9, 14, 18, 22, 24, 34, and 38 and point mutations G373R, G769R, G829E, G904V, G1012R, G1015E, G1018V, G1021E all of which are in the triple helical portion of the protein. I determined the genomic origin of the exon 22

skip to be a 57 base pair deletion which removed 13 bases of intron 21 and 44 bases of exon 22. The remaining six exon skipping mutations were defined at the genomic level by another member of the lab.

The proteins from affected individuals displayed a gradient of overmodification; proteins with mutations closest to amino terminal end of the triple helix displayed the least amount with the degree of overmodification increasing as the mutation moved toward the carboxyl-terminal end of the protein concomitant with increased intracellular retention. The mutations nearest the carboxyl-terminal end of the triple helix, G1012R, G1015E, G1018V, G1018D, G1021E, G1021R, interfere with disulfide bond formation and perturb triple helix formation at 37° C suggesting that the domain for nucleating folding of the triple helix is within the last 20-25 amino acids of this region.

Electron microscopic examination of patient cells indicates that the ER is the site of retention of aberrant type III procollagen and is increasingly distended as mutation moves from N- to C- terminus. The fibrils from mutations in the first third of the triple helix display greater variability than normal and are also larger while those in the last third of the helix are much smaller than normal.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND BACKGROUND

INTRODUCTION

Ehlers-Danlos syndrome type IV (EDS IV) is a serious connective tissue disease caused by mutation of the gene encoding type III collagen. Disease-causing mutations in other fibrillar collagen genes have been extensively studied and the underlying genetics is understood. While some aspects of mutation-induced differences in collagen molecules, their processing and extracellular matrix contribution are known, it is not known how a particular mutation produces a given phenotype. Achieving an understanding of genotype/phenotype correlation requires further research into the intracellular and extracellular changes consequent to mutation. Disease-causing mutations have been characterized in few EDS IV patients, and little research has been devoted to studying the mutant protein. The work described in this dissertation is aimed at identifying mutations in EDS IV patients, and examining type III procollagen within patient cell strains to determine the effect of mutation on collagen folding and processing and whether mutant procollagen is retained within cells by any specific mechanism. Collagen fibril morphology is examined to investigate the effect of COL3A1 mutation on the extracellular matrix.

BACKGROUND

Ehlers-Danlos syndrome (EDS) is a heterogeneous group of connective tissue disorders typified by skin hyperextensibility, joint hypermobility and tissue fragility. First formally described by the physicians Edvard Ehlers and Henri-Alexandre Danlos near the beginning of this century, the syndrome is currently divided into nine distinct types; type IV is considered the most severe (Beighton, 1992).

Clinically, Ehlers-Danlos syndrome type IV (EDS IV) produces a wide range of consequences. Catastrophic rupture of large hollow organs: bowel, arteries or gravid uterus are the most dire results of the disease, often causing death in the third or fourth decade. Other manifestations of EDS IV are: thin skin, frequently with the underlying veins apparent; easy bruisibility; premature aging of the extremities; varicose veins; and hypermobility of small joints (Beighton, 1992). Inheritance of the disease is autosomal dominant.

The molecular defect responsible for EDS IV is mutation of COL3A1, the gene that encodes type III collagen (Pope et al., 1975; Byers et al., 1981). The disorder is not attributable to any single specific mutation. Instead EDS IV may result from point mutations, exon skipping mutations or genomic deletions, any of which can disrupt the folding of the collagen triple helix or assembly of collagen fibrils.

In fibrillar collagens the triple helix is composed of about 330 repeating triplets of the amino acids glycine-X-Y, where proline is often in the X position and hydroxyproline is frequently occupies the Y slot (Linsenmayer, 1991). Folding of the triple helix requires three collagen precursor chains, known as proα chains, to associate by hydrophobic and electrostatic interactions at carboxyl-terminal propeptides. These globular propeptides ensure chain alignment so that the hydrogen bonds and water bridges that stabilize the triple helix can form. Proper alignment or registration of the last few triplets in the helix domain is essential as they serve as the nucleus for folding of this region (Engel and Prockop, 1990; Lees and Bulleid, 1994; Bulleid et al., 1996).

Several elements are necessary for correct folding of collagen. Structurally, every third amino acid in the triple helical domain must be a glycine as it is the only residue that fits in the center of the helix (Piez, 1984; Byers et al., 1991; Kuivaniemi et al., 1991). Stable helix formation requires hydroxylation of some prolyl moities in the X position by prolyl 3-hydroxylase while almost all of those in the Y position are modified by prolyl 4-hydroxylase, the β subunit of protein disulphide isomerase. A third enzyme, lysyl hydroxylase, transforms several lysyl residues in the Y position into hydroxylysine. These are further modified by addition of galactose or glucosylgalactose. Only chains not yet in a triple helical conformation are susceptible to the action of these modifying enzymes (Olsen, 1991). A final requirement for proper folding is the isomerization of *cis* peptide bonds to their *trans* conformation (Engel and Prockop, 1991).

The effect of mutation on collagen folding has been extensively studied in individuals with osteogenesis imperfecta (OI), a disease caused by mutation of the genes encoding type I collagen. Generally, the chains of abnormal procollagens display increased hydroxylation of lsyl moieties resulting in superfluous glycosylation of these residues amino-terminal to the site of the mutation. Although folding of molecules containing abnormal chains is slowed, it appears that overmodification is a consequence of conformational alteration of the area just amino-terminal to the mutation site which leaves the chains of the triple helix accessible to collagen modifying enzymes (Chessler and Byers, 1992; Raghunath, et al., 1994). The aberrant procollagen molecule has a lower melting temperature (Williams and Prockop, 1983; Bateman et al., 1984; Bonadio and Byers, 1985; Bonadio et al., 1985) and therefore may be unstable at physiologic temperatures, and has a longer residence time within the cell as it is not secreted as efficiently (Barsh and Byers, 1981; Bonadio et al., 1985). The endoplasmic reticulum was suspected as the site of retention due to dilatation of the organelle obvious in electron micrographs (Willing et al., 1988). Chessler and Byers (1992, 1993) have found that some abnormal proα1(I) chains are retained within the ER by proteins residing within that organelle.

Monitoring the fitness of procollagen for secretion is the responsibility of the quality control system of the endoplasmic reticulum. The system is comprised of ER-resident proteins that oversee the folding, oligomerization, and modification of proteins destined for secretion. Proteins that are percieved as normal by the quality

control machinery of the ER are ultimately secreted unless they possess a signal directing them to a particular organelle such as the C-terminal KDEL peptide that determines residence within the ER or the signal patch that impels proteins to lysosomes (Munro and Pelham, 1987, Alberts, 1993). A protein that misfolds as a result of errant biosynthesis or is improperly or only partially assembled is rapidly degraded within the ER. Proteins that misfold as a consequence of mutation often aggregate and form complexes in association with molecular chaperones. Chaperone proteins aid other proteins in folding without becoming a part of the final protein product (Laskey et al., 1978). Structural motifs such as exposed hydrophobic regions, exposed free sulfhydryl groups, partially glucose-trimmed oligosaccharides, and the formation of large aggregates attract ER-resident chaperones.

The ER-resident chaperone BiP has been implicated in the retention of abnormal type I procollagen in osteogenesis imperfecta while calnexin associates with abnormal proteins in cystic fibrosis, and hereditary juvenile emphysema. BiP, the immunoglobulin binding protein is a 78kd member of the hsp70 family of stress proteins that associates transiently with many exocytic proteins via a short hydrophic sequence (Blond-Elguindi, 1993). A longer-lived association is observed between BiP and abnormal proteins in both model systems and in patients with osteogenesis imperfecta. BiP binds type I procollagens that have impaired chain association as a result of mutations affecting their C-propeptides but does not bind molecules with mutations in the triple helical domain (Chessler and Byers, 1993). Increased synthesis

of another stress protein, GRP 94, which is coordinately regulated with BiP and has been detected complexed with it and unassembled immunoglobulin subunits (Melnick et al., 1992) was also observed in the OI patient cell strains with C-propeptide mutations (Chessler and Byers, 1993).

Calnexin is a recently discovered chaperone that is a type I integral membrane protein found in association with three other type I integral membrane proteins: signal sequence receptor (SSR) α, SSRβ, and gp25L. Like BiP, calnexin transiently associates with normal proteins while they fold and find their proper subunits. Calnexin, too, is more permanently associated with misfolded proteins and those unable to form oligomers. It appears to specialize in proteins exhibiting N-linked glycosylation (Bergeron, 1994) making it an attractive candidate for binding aberrant proα chains as there is a asparagine-linked glycosylation site in the propeptide of fibrillar collagens (Lamande and Bateman, 1995). Additionally, calnexin has been observed to bind proteins that have been partially folded as a consequence of their interaction with BiP, indicating that chaperones may form a kind of folding production line.

Another candidate for binding abnormal procollagens from EDS IV individuals is the molecular chaperone hsp47 or colligin, found in endoplasmic reticulum of collagen producing cells. Hsp47 was isolated by its binding to gelatin, denatured types I and III collagen (Nagata et al., 1986). *In vitro* binding examined by surface plasmon resonance indicates hsp47 binds to native types I, II, III, IV and V

with approximately the same avidity (Natsume et al., 1994). Transient binding has been observed in chicken and mouse 3T6 cells when radiolabeled cell extracts are immunoprecipitated with either hsp47 or type I procollagen antibodies. Other ER residents proteins are associated with the hsp47-procollagen complex including BiP, GRP94, protein disulfide isomerase, and cyclophilin B (in 3T6 cells) (Nakai et al., 1992; Smith et al., 1995). A longer lived association occurs when triple helix formation is abolished due to treatment with α'α' dipyridyl, analogous to that between BiP or calnexin and abnormal proteins (Nakai et al., 1992). Whether hsp47 mediates intracellular retention of procollagens that are abnormal as a consequence of mutation has yet to be examined.

Ensuring proper protein folding and assembly is also the responsibility of ER-resident proteins that perform specific enzymatic reactions. The β subunit of the collagen modifiying enzyme prolyl-4-hydroxylase is protein disulfide isomerase (PDI) which catalyzes the formation and isomerization of disulfide bonds in the propeptides of procollagen and other proteins that fold in the endoplasmic reticulum. PDI has been implicated in the retention of abnormal type I procollagen in the cells of an osteogenesis imperfecta patient (Chessler and Byers, 1992). The disease-causing mutation was a 4.5 kilobase deletion in COL1A2 affecting the triple helix of type I procollagen. Molecules incorporating the shortened proα2(I) chain were stable and

retained within the cell. The retained molecules associated with PDI but not with BiP indicating that more than one type of ER-resident protein is capable of interacting with abnormal type I procollagen (Chessler and Byers, 1992).

Collagen fibril formation is also affected by mutation. Normally, procollagen molecules traverse the secretory pathway, and are finally deposited outside the cell where the C- and N-terminal propeptides are clipped off by specific proteinases, leaving behind the telopeptides on each end of the triple helical domain. These collagen molecules, 300 nm long and 1.5 nm in diameter, self-assemble, lining up laterally and longitudinally to form collagen fibrils of 10 to 300 nm in diameter (Olsen, 1991, Kuhn, 1987). In one OI patient where cysteine substituted glycine at position 748 the procollagen secreted by his fibroblasts containing two mutant chains were disulfide bonded. This structural change resulted in a kink at the site of substitution and a phase-shift between the three chains that perturbed the N-proteinase site over 700 amino acids (more than 225 nm) away. Fibrils formed *in vitro* from the kinked molecules appeared branched rather than straight rods of normal fibrils (Vogel et al., 1988).

Mutations that interfere with the structural integrity of type III collagen, a large, structural protein can be catastrophic for the individuals who harbor them. How does mutation translate into disease? This broad question can be reduced to smaller, more discrete queries: How does mutation affect folding of type III procollagen? Is secretion of type III procollagen affected by mutation? Are there accessory proteins

responsible for the retention of aberrant type III procollagen molecules? Is cellular architecture changed as a result of the production of an abnormal type III procollagen? Is fibrillogenesis altered in cells of EDS IV individuals? These are the questions I have asked to better understand the mechanisms underlying Ehlers-Danlos syndrome type IV and the chapters that follow are organized within the framework of these questions.

INTRODUCTION TO CHAPTERS II - V

This dissertation describes new mutations in COL3A1 that result in Ehlers-Danlos syndrome type IV and examines the effect of mutation at the molecular, cellular and ultrastructural levels. Chapter II details the characterization of eight mutations that result in the substitution of glycine by another amino acid with a larger side chain, six exon skipping mutations at the cDNA level, and a 57 base pair genomic deletion. The effect of the glycine substitutions on type III procollagen folding is examined in Chapter III; mutations within the last 20 amino acids of the triple helix interfere with the formation of the disulfide bonds at the end of the region as well as disrupting the sequence necessary for folding of the triple helix. Chapter IV describes investigations into whether any ER-resident proteins such as molecular chaperones or collagen modifying enzymes are involved in the intracellular retention of aberrant type III procollagen. The effect of COL3A1 mutations on cellular architecture, as well as collagen fibril diameter and morphology is presented in Chapter V.

CHAPTER II

MUTATION ANALYSIS

EXPERIMENTAL QUESTION

What is the underlying molecular basis for Ehlers-Danlos syndrome type IV in a given patient? Is there any correlation between type (single-base substitutions vs. exon skipping) or location of a mutation and the clinical features or complications of the affected individuals?

INTRODUCTION

The individuals studied were selected because they either exhibited symptoms characteristic of EDS IV or had a family history of the disease. The first step in confirming a suspected diagnosis of EDS IV is to obtain a dermal punch biopsy from the patient, which is explanted and from which fibroblasts are cultured. Freshly plated patient cells are then labeled with ^{3}H-proline, the proteins harvested and analyzed by gel electrophoresis and autofluorography. Type III procollagens and collagens analyzed this way may display structural abnormalities, altered mobility, dimunition in the amount of these proteins or all of these characteristics.

When a diagnosis of EDS IV is confirmed biochemically the COL3A1 gene is screened to find the disease-causing mutation. The first step in screening is the isolation of RNA from patient fibroblasts. The RNA is then reverse-transcribed into

cDNA, fragments of which are amplified by the polymerase chain reaction using a set of overlapping primers that covers the entire COL3A1 cDNA. Exon skipping mutations and genomic insertions or deletions may be obvious at this stage of the detection process. If no such abnormality is observed the PCR fragments, sized 400 - 500 base pairs, are applied to a polyacrylamide gel specifically designed to separate denatured DNA fragments that differ by as little as one base; a technique known as single-stranded conformational polymorphism (SSCP) analysis (Orita et al., 1989). When a difference in the pattern of fragment migration is seen, the fragment is sequenced using direct double-stranded sequencing. Once a mutation is suspected and the appropriate fragment sequenced and found to contain a nucleotide substitution, the mutation is confirmed by an endonuclease restriction digest or allele-specific oligonucleotide hybridization.

RESULTS

Altered mobility and inefficient secretion of type III procollagen

Figure 2.1 illustrates the biochemical findings indicative of EDS IV in eight patients with single amino acid substitutions. The procollagens exhibit the biochemical diversity that is a hallmark of this disease. Type III procollagen from patients G1021R, G1018V, G1015E, and G1012R display delayed electrophoretic mobility and intracellular retention resulting from excess hydroxylation and glycosylation that typify collagens with mutations near the carboxyl terminus. The type III procollagens isolated

from the remaining patient fibroblasts have less obvious abnormalities but are generally secreted less efficiently than type III procollagen from control cells. The type III procollagen from patients with exon-skipping mutations exhibited the same pattern as the glycine substitutions: dimished secretion and an increasing delay in electrophoretic mobility as the exon deletion approached the carboxyl-terminus. In addition, the separation between type III procollagen and proα1 chain of type I procollagen blurred in most cell strains due to the faster migration of molecules composed of three shortened proα(III) chains (data not shown).

Mutations span the triple helical domain

I have identified eight point mutations, six exon skipping mutations at the cDNA level, and one exon skipping mutation at the genomic level. The cDNA-defined exon skipping mutations have been characterized at the genomic level by Ulrike Schwarze, a senior fellow in the lab. The mutations all affect exons that encode portions of the triple helical domain of type III procollagen.

Point mutations

84-152. Sequencing of dermal fibroblast cDNA revealed a mutation in exon 24, a transition at nucleotide 1741 where guanine is replaced by adenine. The codon, GGA, that encodes glycine at residue 373 of the triple helix (the triple helix begins at amino acid 209) becomes AGA which encodes arginine. The mutation eliminates an *Nci* I restriction digestion site, CCG(C)GG, at positions 1738 - 1742. The confirmation is

observed at the cDNA level in a fragment created with primers 22S-28A (data not shown) but is clearest in the genomic fragment generated by primers 23S-25A (figure 2.2).

82-095. A G→C transversion at nucleotide 2929, in exon 40, is responsible for the arginine for glycine substitution observed in this cell strain. The glycine encoded by GGT at position 769 becomes CGT. An *Mnl* I site, CCTCN(6), is created by the mutation at base 2929; the normal sequence GCCTGGT is not cut while the mutant CCTCGT is cut. Figure 2.3 shows the confirming restriction digest on the genomic fragment created by primers 40S-41A.

89-176. Glutamic acid replaces glycine at amino acid 829 as a result of a guanine to adenine transition at nucleotide 3110 in exon 42. *Drd* I was used to the confirm this mutation; the site is eliminated in the mutated allele but present in the normal allele as shown by the cDNA digest in figure 2.4. Digestion of the PCR product 42S-47S by *Drd* I produces fragments of 409 and 48 base pairs in normal alleles; the mutated allele remains uncut yielding a fragment 457 bases in length.

93-057. A G→T transversion at base 3335, in exon 47, changes the glycine codon GGC to GTC resulting in a valine at amino acid 904. Although it appeared as if the transversion occurred at the second site, it was not unequivocal. A first site mutation would produce a cysteine substitution rather than valine. No restriction enzyme was available to differentiate between these two possibilities so allele-specific oligonucleotide hybridization was performed. Oligonucleotides incorporating either no

change or the first or second position thymine were generated. These were then hybridized to patient and normal PCR products defined by primers 42S-47A. As expected, the normal oligonucleotide hybridized to both the patient sample which possesses one normal allele and to the normal control. The oligo with the first position transversion (cysteine oligo) hybridized to neither sample and the probe specific for valine, the second position substitution, hybridized to the patient sample but not the control, confirmation that the mutation results in a valine substitution (figure 2.5).

94-196. A first position G→C transversion in the codon for glycine 1012 in exon 49 results in the substitution of arginine for glycine. The 446 base pair fragment of COL3A1 generated by primers 49J and 50A was used for digestion with *Hae* III. The GGCC cut site of this enzyme is eliminated in the patient's mutant allele resulting in a 62 base pair fragment rather than two smaller pieces of 44 and 18 bases (figure 2.6).

82-115. Glycine codon GGG is changed to GAG by a G→A transition at nucleotide 3670 in exon 49 resulting in the substitution of glutamic acid at residue 1015. No informative restriction digestion site encompassed this mutation so the sequence was modified to generate one. A mismatch primer, 5'AGAGGATCTGAGGGCTCCCCAGGCCACCAA 3', was used to create the CCT(A)T(A)GG site of restriction endonuclease *Sty* I which is eliminated in the mutant allele of the patient. The patient digest exhibits the fragments representing the normal alllele of 282 and 28 bases and an uncut fragment of 310 (figure 2.7).

77-008. Another mutation in exon 49 results in a G→T transversion at base 3677 changing amino acid 1018 from glycine to valine. The mutation eliminated a *Hae* III site in the patient's abnormal allele combining two fragments of 62 and 279 bases into a 341 base segment. The PCR product created by primers 46S-49A was used for this restriction digest (figure 2.8).

91-579. A G→A transition at base 3685 in exon 49 changes the glycine codon GGA to a GAA codon for glutamic acid at amino acid 1021. The mutation creates a new site for the restriction endonuclease *Ava* II resulting in 93 base pair fragment that is not present in the normal allele (figure 2.9).

Exon-skipping mutations

A summary of exon-skipping mutations is presented in table 2.1.

86-052. The PCR fragment defined by primers 6S-13A showed a deletion of approximately 54 base pairs when analyzed on a polyacrylamide gel. Direct double stranded sequencing of this cDNA fragment indicated that exon 9 was skipped. The genomic origin of the abnormality proved to be a substitution of guanine for thymine in position +2 of the intron 9 5' donor consensus splice sequence.

84-174. A G→T transversion at the +3 position of the 5' donor splice site in intron14 causes the skipping of exon 14.

93-200. The mutation resulting in the skipping of exon 18 in this patient is a consequence of a G→C transversion at the -1 position of the 3' splice site acceptor consensus sequence in intron 17.

91-336. A 57 base pair genomic deletion causes exon 22 to be skipped; 13 nucleotides are deleted from the 3' end of intron 21 and 44 from exon 22.

88-039. Exon 24 is skipped due to the G→A transition in the +1 position of intron 24.

82-026. A T→C transition in intron 34, at the +2 position of the 5' splice donor sequence, causes the skipping of exon 34, which was identified in a fragment generated by primers 32S and 38A.

93-344. Sequencing of PCR amplified fragment 37S-41A revealed that exon 38 is skipped in this patient. Genomic sequencing proved that exon 38 was skipped as a result of substitution by adenine for the guanine at position +5 of the 5' donor splice site of intron 38.

DISCUSSION

I have characterized the disease-causing mutation in 15 individuals with Ehlers-Danlos syndrome type IV. These mutations span virtually the entire triple helical portion of type III procollagen and emphasize the genetic heterogeneity associated with the disease. Phenotypic heterogeneity is also a hallmark of EDS type IV and is observed in the people whose mutations have been characterized. There is no phenotype-genotype correlation in the mutations I have characterized or other mutations identified in the Byers lab or published in the literature. In fact, point mutation G034R (Goldstein et al., 1994) identified by Ulrike Schwarze highlights the different possible outcomes of the same mutation in type III collagen. The male

proband died at age 34 of a ventricular rupture, his sister died at age 28 of an aortic aneurysm, yet their approximately 60-year old affected mother is alive having only minor symptoms of disease.

Biochemical heterogeneity is also evident in the type III collagens and procollagens from the 15 individuals whose mutations I identified. The position of the mutation not its nature, substitution of obligate glycine, exon-skipping etc., determines how the collagen is handled by collagen modifying enzymes. Those persons with mutations close to the carboxyl-terminal end of the triple helix have proteins that are posttranslationally overmodified to the greatest extent with concomitant intracellular retention. The degree of overmodification and intracellular retention lessen as the mutation nears the amino-terminal end of the helix. Cell strains harboring mutations closest to the carboxyl-terminal end of the helix displayed an usual pattern of migration on SDS-PAGE gels. Under non-reducing conditions, the type III collagen from individuals with mutations G1021E, G1018V, G1015E, and G1012R migrated in the position of a type III procollagen monomer instead of the expected position of a trimer. The mutations in this subset of patients interfere with the formation of the normal trimer and are examined at length in chapter IV.

The disease-causing mutations I characterized represent the two most common types of mutation seen in collagen genes: substitution of obligate glycines within the triple helical domain and exon skipping due to a lesion within a splice site. Both are due to single base substitutions which is the most common genomic alteration,

accounting for over 40 human genetic diseases (Cooper et al., 1993). Several of the mutations I have identified are reported in unrelated individuals in the literature or identified by others in the Byers lab. The arginine for glycine substitution that occurs at residue 373 has been identified by Ulrike Schwarze of the Byers lab (Goldstein et al., 1994) in a second EDS type IV individual. The substitution occurs at a mutational hotspot, a CpG dinucleotide.

It is estimated that 35% of human disease-causing point mutations originate at CpG sites (Cooper and Youssoufian, 1988). Mutations arise when cytosine is methylated and then deaminated to form thymine. Collagen genes with their CG rich coding regions and high glycine content, which is encoded by the sequence GGN, might be expected to have many CpG sites. However, as is seen in the coding regions of many vertebrate genomes, there is a paucity of CpG sites (Cooper and Youssoufian, 1988). The dearth of this sequence in the genome is thought to result from the tendency of 5-methylcytosine to undergo deamination, forming thymidine. This hypothesis is bolstered by the observation that 90% of the disease-causing point mutations at CpG sites are G→A or C→T transitions (Cooper and Youssoufian, 1988). Deamination of 5-methylcytosine may produce uracil with an equal frequency but consequent disease is probably avoided by the removal of the anomolous base by uracil DNA glycosylase whereas thymidine is recognized as legitimate and is left in place by DNA repair machinery (Cooper et al., 1993). Of the 341 glycine codons in type III procollagen only 16 within the triple helical region are preceded by a C. The

deamination mechanism predicts that the majority of substitutions at these sites would result in either serine or arginine. CpG sites are implicated in a number of previously reported mutations of COL3A1 including G619R associated with familial aneurysm (Kontusaari et al., 1990), G415S (Anderson et al., 1997) which has also been characterized in a second unrelated individual (Schwarze and Byers, unpublished), and the glycine at position 16 which has 7 mutations associated with it, six are serine substitutions and one is a replacement with cysteine (Schwarze and Byers, unpublished).

Six of the exon skipping mutations I identified at the cDNA level were further characterized by Ulrike Schwarze and found to result from single base substitutions within the splicing sites. Lesions of this type are the source of approximately 15% of the total point mutations causing human genetic disease (Cooper et al., 1993). Mechanistically, accurate removal of introns from mRNA requires recognition of 5' donor and 3' consensus splice site sequences by small nuclear ribonucleoproteins and other splicing factors (Valcarcel et al., 1996; Singh et al., 1995; Wang et al., 1996; MacMillan et al., 1997). Splicing accuracy is further influenced by RNA secondary structure (Estes et al., 1992) and intron and exon length (Sirand-Pugnet et al., 1995; Cogan et al., 1996; Sterner and Berget, 1993; Peterson et al., 1994; Van Oers et al., 1994; Del Gatto et al., 1996).

Analysis of splice site mutations in human genetic disease indicates a non-random distribution of single-base substitutions with the majority of mutations affecting the 5' donor site (Cooper et al., 1993). This pattern is true of the mutations I have identified as six individuals have mutations in the 5' site while only one has an error in the 3' sequence. When considering EDS type IV splice site lesions overall, of the 43 separate mutations identified only 3 are in the 3' acceptor site (Pope et al., 1996, Schwarze et al., 1997). Investigators theorize that the concentration of 5' splice site sequence mutations is probably due to an ascertainment bias whereby phenotypic severity of a given disease brings patients to a physician's notice and there is, in fact, probably no difference in the frequency of mutation at 3' and 5' splice sites (Cooper et al., 1993). This may well be the case for EDS IV patients as the possible outcomes for the two splice site changes differ. Both 5' and 3' splice site mutations can result in exon-skipping. Because the exons encoding the triple helix all begin with glycine and are in multiples of nine, exon skipping can produce in-frame mRNA that encodes a shortened but functional protein. Alteration of the 5' splice site also leads to read through whereby the intronic sequence becomes part of the mature RNA, and use of a cryptic donor site from the preceding exon or upstream intron. Read through resulting in the inclusion of an entire intron in mature COL3A1 mRNA has been observed with an intron 20 $G^{-1}\rightarrow A$ transition (Kuivaniemi et al., 1990) and although it is a possible outcome of the disruption of the 5' splice site of intron 34 it does not appear to be used as an alternative (Schwarze et al., 1997). Of the other 5' mutations identified,

inclusion of an entire intron would produce an exon larger than the 300 nucleotide limit required for proper splicing, so skipping would be the likely result (Schwarze et al., 1997). An alternative with 3' splice-site mutations, is the use of an alternative acceptor site within the intron or the succeeding exon. Use of an alternative acceptor site could have quite a different outcome. Schwarze et al. (1997) examined exons 6-49 (those that encode the triple helix) for the c/tag splice-acceptor sequence and found that of the 84 potential sites, 76 would result in out-of-frame mRNA sequences that would be expected to be retained within the nucleus or be unstable (Maquat, 1995, 1996; Stover et al., 1993). Thus, it is possible that 3' splice-site mutations of COL3A1 most often result in "null" alleles which gives rise to a milder phenotype than EDS IV and, therefore the majority of individuals harboring such a mutation would not be included in a study such as this. For those individuals who are diagnosed with EDS IV, phenotypic diversity, if present at all, is subtle regardless of which splice site is involved.

MATERIALS AND METHODS

Clinical Summary - The clinical features of each individual are provided in Table 2.2.

Selection of patient samples - Cells selected for mutation analysis all had evidence of synthesis or secretion of abnormal type III procollagen and were derived from individuals with clinical signs of EDS type IV or with a family history of the disease (Table 2.2)

Procollagen and collagen analysis - Patient and control dermal fibroblast cultures were established with appropriate consent and maintained as previously described (Bonadio et al., 1985). When added, ascorbate was used at a final concentration of 50 μg/ml.

Procollagens and collagens were labeled with [^{3}H]proline (Amersham Corp.) and separated by SDS-PAGE as previously described (Bonadio et al., 1985) except that proteins were precipitated with 70% ethanol (v/v final concentration) for one hour on ice after addition of unlabeled collagen (Sigma Chemical Co.) rather than dialyzed into ammonium bicarbonate.

RNA isolation, cDNA transcription, and amplification - Total cytoplasmic RNA was isolated from cultured fibroblasts by the guanidinium chloride method (Chomczynski and Sacchi, 1987). The total RNA was reverse transcribed to cDNA using random hexamers (Feinberg and Vogelstein, 1984) which was then PCR amplified (Saiki et al., 1988) using overlapping primer pairs that covered the entire COL3A1 gene. (Table 2.3)

Single-Strand Conformation Polymorphism (SSCP) Analysis - Amplified fragments ranging in length from 300 - 500 bp were applied to a MDE™ gel (FMC Bio Products) following the methods of Orita et al. (1989). Electrophoresis was performed overnight at constant power, 8 W, at room temperature. The fragments were visualized with silver stain (Promega Corp.).

DNA sequence determination - PCR-amplified fragments were gel purified in NuSieve® agarose and sequenced by the dideoxy-chain termination method (Sanger et al., 1977) using T7 polymerase (Sequenase®, U. S. Biochemicals).

Mutation confirmation - Mutations were confirmed by restriction digestion of an existing site changed by mutation, by introduction of a sequence change in one amplification primer that created an informative site when the mutant allele was amplified, or by allele specific oligonucleotide (ASO) hybridization to the normal and mutant alleles.

Table 2.1 **EXON-SKIPPING MUTATIONS**

Patient ID number	Aberrant PCR fragment	Exon skipped in cDNA	Genomic basis for skipped exon
86-052	6S-13A	9	Intron 9 $T^{-2}\rightarrow G$
93-272	12S-19A	14	Intron 14 $G^{+3}\rightarrow T$
93-200	12S-19A	18	Intron 17 $G^{-1}\rightarrow C$
91-336	22S-28A	22	13bp intron 21, 44bp exon 22
88-039	22S-28A	24	Intron 24 $G^{+1}\rightarrow A$
82-026	32S-38A	34	Intron 34 $T^{+2}\rightarrow C$
93-344	37S-41A	38	Intron 38 $G^{+5}\rightarrow A$

Table 2.1 **Ehlers-Danlos Syndrome Type IV: Clinical Manifestations**

Age = age when EDS IV diagnosis was made on clinical grounds

FH = family history; + is one or more affected family members; - is neither parent affected

+, -, blank = + is presence and - absence of trait, a blank indicates a lack of information

ID	Mutation	Sex	Age	FH	Face	Joints	Skin Manifestations			Blood Vessels		Complications			Death	
							thin trans-lucent	cigarette paper scars	bruising bleeding	venous aricosities	neurysm	arterial rupture	bowel rupture	spleen rupture	age	cause
84-152	G373R	M	31	+								+				
82-095	G769R	F	25	+			+									
89-176	G829E	M	14 mo	-	+	+	-	-	+	-	-	-	-			
93-057	G904V	M	30	+								+	-			
94-196	G1012R	F	4	-	-	-	+	-	+	-	-	-	-			
82-115	G1015E	M	10	-	+	+	+	-	+	-	-	-	-			
77-008	G1018V	F	26	-	+	-	+	+	+	-	+	-	+		48	
91-575	G1021E	M	2	-	+	-	+	-	+	-	-	-	-			
94-341	exon 9	F	45	+		+	+		+			+				
93-272	exon 14	M	34	-		+	+	+	+	+	+	+				
93-200	exon 18	F	13	-		+		+	+							
91-336	exon 22	M	17	-	+	+	+	+	+	+	+	+				
88-039	exon 24	F	24	-	+	+	+		+		+					
93-177	exon 34	F	19													
93-344	exon 38	F	13	+		+	+		+				+			

Table 2.3. **COL3A1 PRIMERS**

Exon	Sense Sequence
1	CTGAAGGGCAGGGAACAACTTG
2	GTCAGTCCTATGCGGATAGAGA
6	AACTATTCTCCCCAGTATGATTCATATGA
12	CGAAATGGAGAAAAGGGTGAAACA
18	GGCCCTCCTGGGATTAATGGTA
21	TAAAGGCGAAGATGGCAAGGATG
23	GCTGCTGGAGAACCTGGCAG
26	TCCTGGTAAGAATGGAGAACGA
32	AGGAATGCCTGGAGAAAGGAG
37	GCTCCTGGACAGAATGGTGA
40	CAGCGGTTCTCCAGGCAAGG
42	ACCAGGAGCTACGGTCTCAG
46	TCCTGGTCCTCAAGGCCCACGTGGTGACAA
48	GGACCTGTTGGACCCAGTGGA
49	CTCATTAGTCCTGATGGTTCTC
49J	CAGGGCCTCGAGGTAACAGAGGTGAA
50	AGAAACACGTTTGGTTTGGAGAG

Exon	Antisense Sequence
3	CAAGGCCCCAAGGGAGATCCA
7	ATCTGGTCATCCTGGTTCCCC
13	TGAAAATGGTCTTCCAGGCGAAAA
19	GGAACCAGCCGGTGCTAATGGT
23	GATGGCGTCCCTGGAGGTCCA
25	TCGACCACTTTCTCCTTGACTT
28	ACAAAGGAGACACAGGACCC
33	TGGTGTCCCAGGGAAAGATGGC
38	TCCCAAGGTGTCAAAGGTGAA
41	GCCCTGGCCCTCAGGGTGT
44	GATCGTGGTGAAAAATGGCTCT
47	GGTGCAATCGGCAGTCCAGGA
49	GCCATCCTGAACTCAAGACT
50	AAGTGCCAATCCTTTGAATGTTCCACGGAA
51	ATAGCATTGCATACATGGATCA
52	TGTTCCTCTTGTTCTAATCTTGTCAACC

Figure 2.1 (a) **Procollagens from EDS IV patient and control cells .** Cells were incubated with [^{3}H]proline for 16 hours and medium and cell layer proteins were harvested separately. Proteins were separated by SDS-PAGE under reducing conditions. Lanes 1 and 10 are control cells. Lanes 2-9 are point mutations: G373R, G769R, G829E, G904V, G1012R, G1015E, G1018V, and G1021R respectively. The proα1(III) chains in lanes 3-9 are poorly secreted and retained with the cells. These proteins are also overmodified as the wider, diffuse band indicates.

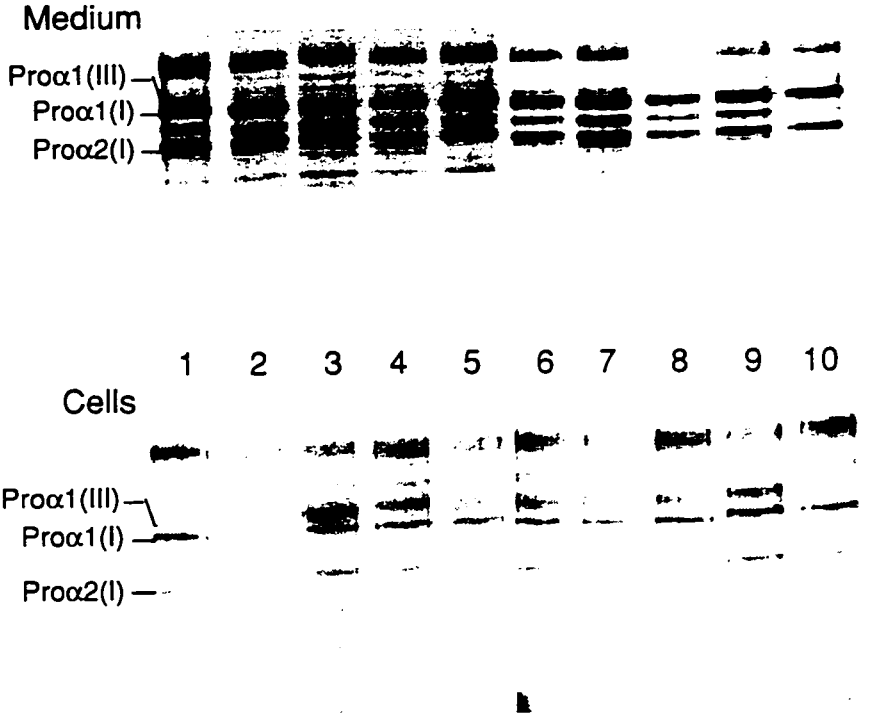

Medium
Proα1(III)
Proα1(I)
Proα2(I)
1 2 3 4 5 6 7 8 9 10
Cells
Proα1(III)
Proα1(I)
Proα2(I)

Figure 2.1 (b). **Collagenous proteins from EDS IV patient and control cells.** Cells were incubated with [^{3}H]proline for 16 hours and medium and cell layer proteins were harvested separately. Proteins were separated by SDS-PAGE under non-reducing conditions after digestion with pepsin. Lanes 1 and 10 are control cells. Lanes 2-9 are point mutations: G373R, G769R, G829E, G904V, G1012R, G1015E, G1018V, and G1021R respectively. Note the wide, diffuse band migrating in the position of a monomer of type III collagen in the cell layer proteins of cell strains with mutations: G1012R, G1015E, G1018V, and G1021R.

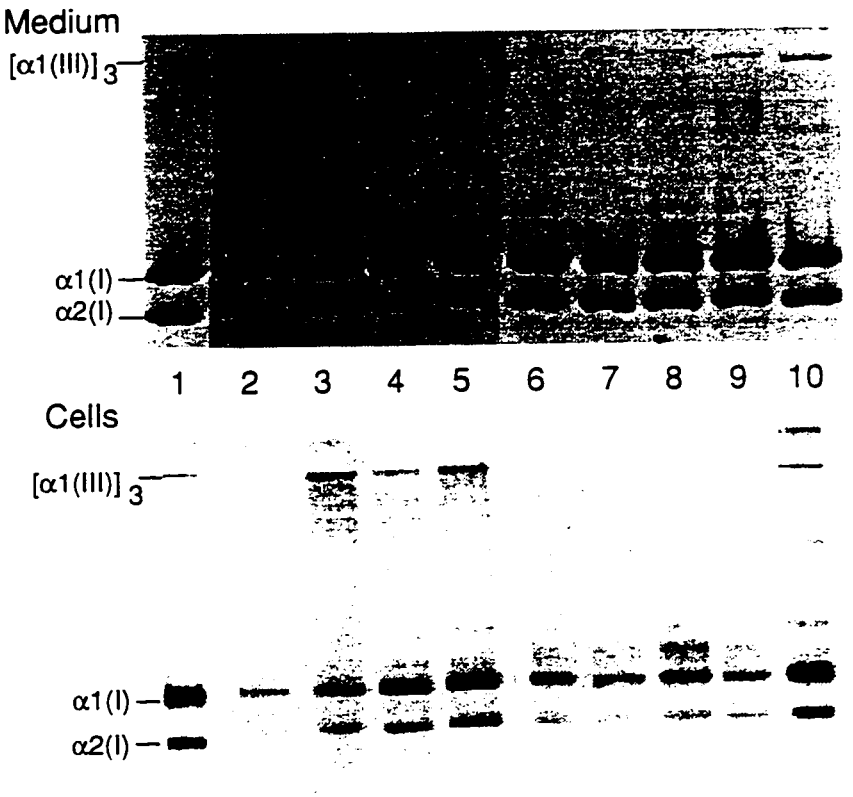

Medium
[α1(III)]3
α1(I)
α2(I)
1 2 3 4 5 6 7 8 9 10
Cells
[α1(III)]3
α1(I)
α2(I)

Figure 2.2. **G373R confirming restriction digest.** Genomic DNA amplified by primers 23S and 25A was digested with Nci I to confirm the G→A transition at nucleotide 1741. The *Nci* I site, CCG(C)GG is eliminated by the mutation in patient 84-152, lane 1, while the enzyme is active in the controls, lanes 2 and 3. Lane 4 shows uncut patient DNA. Lanes 5 and 6 are uncut control DNA. Lane 7 is λ DNA digested with *Pst* I.

1 2 3 4 5 6 7

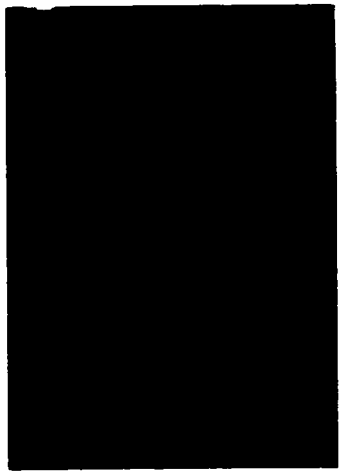


Figure 2.3. **G769R confirming restriction digest.** The G→C transversion at nucleotide 2929 of patient 82-095 was confirmed using genomic DNA amplified with primers 40S and 41A. Lane 1 contains uncut control DNA, lane 2 uncut patient DNA. The DNA in lanes 4 and 5 was digested with *Mnl* I. The control DNA remains uncut while the patient DNA exhibits fragments of approximately 130 and 210-220 base pairs resulting from the creation of the enzyme's CCTCN(6) site in addition to an uncut band of approximately 339 nucleotides representing the mutant and normal alleles, respectively. Lane 1 contains λ DNA digested with *Pst* I.

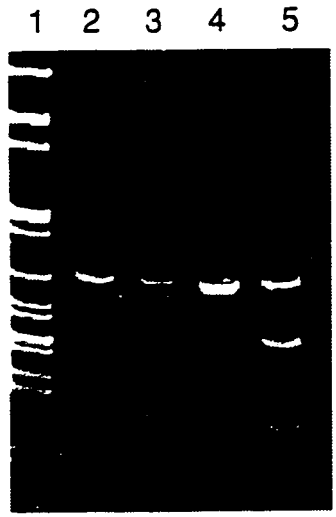

1
2
3
4
5

Figure 2.4. **G829E confirming restriction digest.** Drd I was used to confirm the G→A transition at nucleotide 3110 in patient 89-176. cDNA was amplified with primers 42S and 47A producing a 457 base pair fragment. Uncut control DNA is shown in lanes 2 and 3, patient uncut DNA is in lane 4. After digestion with Drd I, fragments of 490 and 48 base pairs are seen in the control lanes (5 and 6) as well as lane 7 where it represents the normal allele of the patient. The fragment marked with the * is the uncut fragment generated by the mutated allele. Lane 1 contains λ DNA digested with *Pst* I.

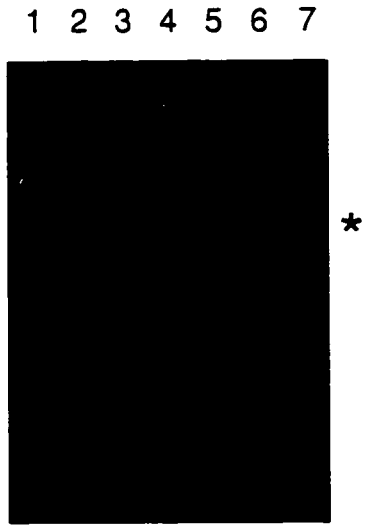

1 2 3 4 5 6 7
*

Figure 2.5. **G904V confirmation by allele specific oligonucleotide hybridization.** Allele specific oligonucleotide hybridization was used to differentiate between a first and second site G→T transversion in codon 904 of patient 93-057 as no restriction endonuclease was available. A cDNA fragment encompassing codon 904 was generated by PCR amplification using primers 42S and 47A. Oligonucleotides incorporating either no change, 5'GCCCTGCTGGCCCTGCT3', a first position change resulting in a valine substitution, 5'GCCCTGCTTGCCCTGCTG3'or a cysteine created by a second position mutation, 5'GCCCTGCTGTCCCTGCT3' were hybridized to a normal control and the patient cDNA. The normal probe hybridized to the control (lane 1) and the patient (lane 2) cDNA. The cysteine probe hybridized to neither sample (lanes 3 and 4) and the valine probe hybridized only to the patient cDNA in lane 6 and not the normal control in lane 5.

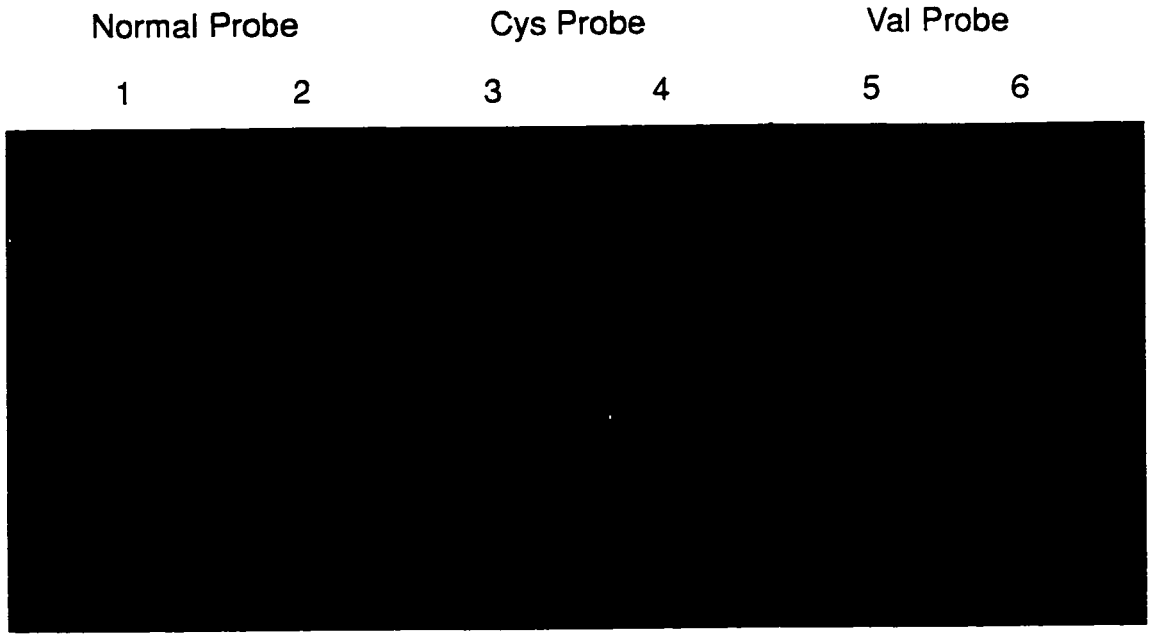

Normal Probe
Cys Probe
Val Probe
1
2
3
4
5
6

Figure 2.6. **G1012R confirming restriction digest.** A first position G→C transversion resulting in the substitution of arginine for glycine in exon 49 was confirmed with *Hae* III. A 446 base pair fragment of COL3A1 generated by primers 49J and 50A was used for digestion. The *Hae* III GGCC site is eliminated creating a 62 base pair fragment in the mutant allele (* in lane 5) in addition to two smaller pieces of 44 and 18 bases representing the normal allele and present in the control (lane 4). Lane 2 is uncut control DNA, lane 3 uncut DNA from patient 94-196. Lane 1 is λ DNA digested with *Pst* I.

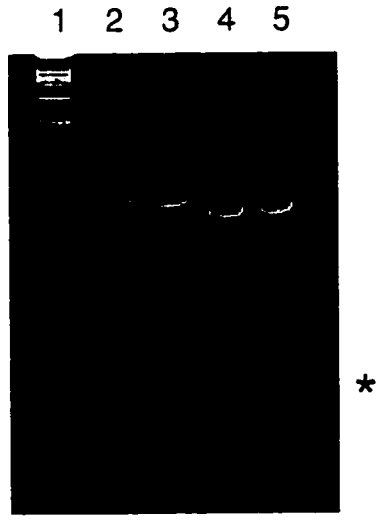
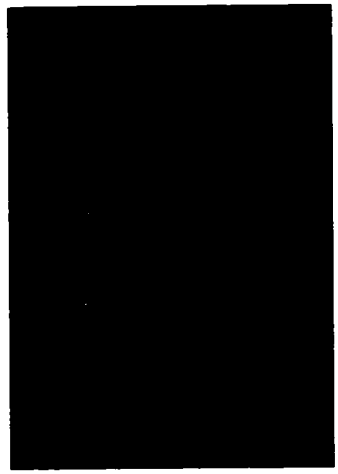

1 2 3 4 5
*

Figure 2.7. **G1015 confirming restriction digest.** Because no informative restriction digestion site existed confirming the change from GGG to GAG generated by the G→A transition at nucleotide 3670 in exon 49 one was created with a mismatch primer. Primer 5'AGAGGATCTGAGGGCTCCCCAGGCCACCAA3' was used to create the *Sty* I site of CCT(A)T(A)GG. The site in eliminated in the mutant allele of patient 82-115 generating a 310 base pair fragment (* in lane 5) while the normal allele and the control DNA (lane 4) is cut creating two fragments of 282 and 28 bases. Undigested control DNA and patient DNA are in lane 2 and 3 respectively. Lane 1 contains λ DNA digested with *Pst* I.

1 2 3 4 5



*

Figure 2.8. **G1018V confirming restriction digest.** A G→T transversion at nucleotide 3677 in exon 49 of patient 77-008 results in a valine for glycine substitution at residue 1018. A *Hae* III restriction site is eliminated in the patient's abnormal allele combining two fragments of 62 and 279 bases into a 341 base pair fragment in the COL3A1 segment generated by amplification with primers 46S and 49A (* in lane 5). Control DNA digested with *Hae* III is in lane 4; undigested control DNA in lane 2 and undigested patient DNA in lane 3. Lane 1 contains λ DNA digested with *Pst* I.

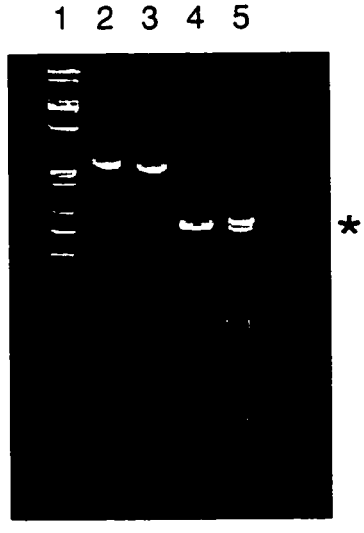
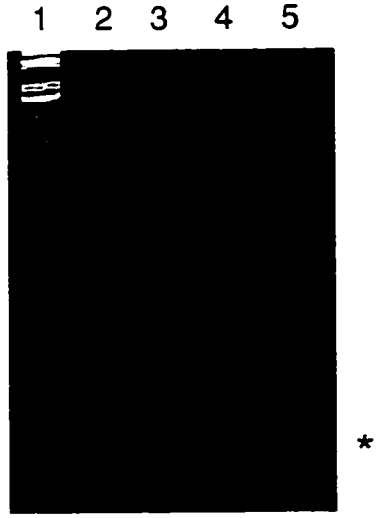

1 2 3 4 5
*

Figure 2.9. **G1021E confirming restriction digest.** A G→A transition at base 3685 changes the glycine codon GGA to a GAA codon for glutamic acid at amino acid 1021 in exon 49 of patient 91-579. The mutation creates a new site for the restriction endonuclease *Ava* II resulting in 93 base pair fragment that is not present in the normal allele (* in lane 5). Lane 4 contains digested control DNA. Lanes 2 and 3 contain undigested DNA from a control and the patient, respectively. Lane 1 contains λ DNA digested with *Pst* I.


1 2 3 4 5
*

CHAPTER III

DEFECTIVE HELIX INITIATION OF TYPE III PROCOLLAGEN WITH MUTATIONS IN THE FAR END OF THE CARBOXYL TERMINUS OF THE TRIPLE HELIX

EXPERIMENTAL QUESTION

In this chapter I explore the effect of COL3A1 mutation at the molecular level by asking the question: Does mutation interfere with procollagen folding in a position-specific manner?

INTRODUCTION

Collagen folding is unusual. The complex multidomain protein requires three precursor chains, α chains, to unite to fold into its final conformation. Chain association is mediated by globular propeptide domains at the carboxyl ends of fibrillar colllagen. The three chains then wind around each other forming a right handed super helix that is propagated from the C-terminus to the N-terminus in a zipper-like fashion. Kinetically, folding occcurs in two phases: an initial fast phase involving the first 25 or so tripeptides and a slow phase, the rate of which is determined by conversion of *cis*-peptide bonds to their required *trans* conformation (Engel and Prockop, 1991).

Individuals with EDS IV exhibit a variety of mutations that alter the primary structure of type III procollagen and interfere with its proper folding. Genomic insertions or deletions and exon-skipping mutations disrupt the registration of the three proα chains while point mutations that replace obligate glycines cause local perturbations and interfere with folding as their bulky side chains do not fit in the center of the helix. The consequence of any of these mutations is to expose proα chains to the action of collagen modifying enzymes for a prolonged time resulting in excess hydroxylation and glycosylation of the molecule. In the cell, abnormal chains associate with both abnormal and normal chains, amplifying the effect of the mutation. Because type III procollagen is a homotrimer the ratio of abnormal:normal molecules is 7:1. The malfolded procollagens are usually retained within the cell rather than being secreted.

Not all mutations affect collagen in the same way. In two individuals with perinatal lethal osteogenesis imperfecta (OI II) replacement of obligate glycines with serines had very different effects on thermal stability even though the residues were only 35 amino acids apart. The effect on stability was assayed by judging the resistance to proteases of type I collagen. The collagens are heated and subjected to proteases that are active only on monomeric chains not the triple helix. The resulting fragments are electrophoretically analysed to determine at what temperature the helix destabilizes. The individual with serine α1-631 had a thermal unfolding temperature of 41°C, comparable to control cells, while the cells from the individual with serine α1-598

produced collagens that unwound at 20°C (Westerhausen et al, 1990). This drastic difference in melting temperatures corroborates experimental results indicating that collagen is not a stabile rod but is instead composed of microenvironments that "breathe" or unfold at physiologic temperatures (Privalov, 1982).

I identified a group of EDS IV patients with mutations very near the carboxyl-terminal end of the triple helix, G1012R, G1015E, G1018V, and G1021R whose dermal fibroblasts produce procollagens that are poorly secreted and are instead retained within the cells. After partial proteolysis with pepsin a novel protein band that migrates in the position of α1(III) monomers appeared. Although it was not disulfide-bonded into either a dimeric or trimeric molecule it contains the cysteine residues at the end of the triple helix. Thus, these mutations perturb the triple helix carboxyl-terminal to the site of mutation sufficiently to interfere with the formation of the nearby interchain disulfide bonds. In fact, it appears that this set of mutations interrupts the sequence required for initiation of folding, or nucleation, of the triple helix.

RESULTS

Mutations near the carboxyl-terminal end of the triple helix alter folding

The mutations closest to the end of the triple helix, G1012R, G1015E, G1018V, G1021R, (and G1018D, G1021E characterized in the Byers lab) had an unexpected effect on proα chain assembly and interchain disulfide bond formation. When the intracellular proteins were digested with pepsin at 15 °C and the protease resistant helical proteins were separated under non-reducing conditions, there was a band that migrated just above the α1(I) chains of type I collagen (figure 3.1; see figure 2.1 for procollagens and secreted collagens). Peptide mapping showed that this band was monomeric α1(III) (data not shown). The α1(III) chains in a type III collagen molecule normally are in a disulfide-bonded trimer linked through two cysteine residues that form the last residue of the triple helix and the first non-helical residue. These findings suggested that either the cysteine-containing region was removed by proteolysis or that the mutations did not allow disulfide bonds to form.

To determine if a new pepsin site was introduced by these mutations resulting in the removal of the cysteines at the end of the triple helix, cells with mutations G1012R, G1018D and G1021E were labeled in parallel with [^{35}S]cysteine and [^{3}H]proline. Type III procollagen labeled only with [^{35}S]cysteine was immunoprecipitated with antibody directed against the N-terminal propeptide, and the precipitated protein was digested with pepsin and examined by SDS-PAGE (figure

3.2). The monomeric α1(III) chain contained cysteine which indicated that the pepsin cleavage site was external to the carboxyl-terminal end of the triple helix. Thus, although the helix could form and produce a pepsin resistant molecule, the cysteine residues were too far apart to allow disulfide bonds to form. When the cells with the G1021E mutation were labeled with [^{3}H]proline, the proteins harvested by ethanol precipitation, and the procollagens analyzed by SDS-PAGE under nonreducing conditions, the proα!(III) chains associated into trimeric molecules which indicated that the mutations in the triple helix did not alter inter-chain disulfide bond formation in the carboxyl-terminal propetide (figure 3.3).

To determine if the mutations in the carboxyl-terminal end of the triple helix altered the structure of the end of the helix, cells with glutamic acid for glycine substitutions at residues 61, 385, and 1021 were incubated with [^{3}H]proline for 16 hours at 37°C, the medium changed, and the cells shifted to 4°C, 20°C, or 37°C for 40 minutes to allow the triple helix to propagate. The intracellular proteins were harvested by cell lysis and treated with trypsin/chymotrypsin at 20°C for 2 minutes to remove the terminal propeptides (figure 3.4). In the cells with the G061E mutation and those with the G385E substitution, a disulfide-bonded trimeric protein with a mobility faster than the normal trimer was seen at all temperatures. Triple helix was nucleated and propagated to the region of the mutation (figure 3.5). The cells harboring the 1021E mutation displayed a small amount of trimeric type III procollagen but no helical molecules migrating as monomers were evident at any temperature (figure 3.4). These

experiments suggest that helix does not nucleate at the temperatures and times used in these studies with mutations near the carboxyl-terminus (figure 3.5); re-nucleation apparently is a slow process that can occur at lower temperatures.

DISCUSSION

Mutations that create substitutions for glycine close to the C-terminal end of the triple helix, produce procollagens that are retained within the cell rather than secreted. After partial proteolysis with pepsin and SDS-PAGE under non-reducing conditions a novel protein band that migrates in the position of the α1(III) monomer appears but contains the cysteine residues at the end of the triple helix. These mutations perturb the triple helix carboxyl-terminal to the site of mutation sufficiently to interfere with the formation of the nearby interchain disulfide bonds. One explanation for this observation is that folding of the triple helix is initiated, proceeds past the mutation and then unwinds so that disulfide bonds cannot stablize the carboxyl-terminal end of the triple helix. Alternatively, triple helix nucleation does not occur at the labeling temperatures and the helix, which is resistant to pepsin, forms only when the molecules are cooled for a long period. My studies suggest that helix nucleation does not occur at the usual temperature at which proteins are labeled and there is no re-nucleation amino-terminal to a disruption in the triple helix at 37°C.

The effects of mutations on the folding of the triple helix of collagens have, in general, been inferred from alterations in thermal stability (Bateman et al., 1992; Deak et al., 1991; Pack et al., 1989; Westerhausen et al., 1990). In these studies, proteins are cooled following biosynthesis, partially purified with protease digestion at low temperatures under acid conditions, and then their susceptibility to proteolysis at physiologic pH determined. The extent to which such studies provide information about the folding of molecules that contain chains with sequence alterations is, however, not entirely clear, especially given the observations that the rate of folding *in vitro* is far slower than in the cell and displays other non-uniform properties (Davis and Bachinger, 1993). There are few studies in which the effects of mutations on molecular assembly and triple helix propagation have been investigated in detail.

Amino acid alterations or truncation in the carboxyl-terminal propeptide of the proα1(I) chain of type I procollagen have been shown to delay the incorporation of the chains into trimeric molecules, allowing overmodification within the triple helical domain (Chessler and Byers, 1993; Lamande et al., 1995). In a study of mutations within the triple helical domain, Raghunath and colleagues (1994) demonstrated that some substitutions of cysteine for glycine in the proα1(I) chain altered the rate at which triple helix propagated to the full extent of the molecule. They found that cysteine for glycine substitutions at positions 988, 691, and 223 markedly delayed triple helix formation at 37°C while those at 526 and 94 did not. Although delayed, helix formation did proceed to the full length of the triple helix under the conditions they

used. Their findings contrast with mine in which propagation of triple helix beyond the site of the substitution of glycine by glutamic acid in type III procollagen appears to be very inefficient in most molecules that contain one or more of the abnormal chains (see figure 3.5 for a model of folding). I am uncertain if this effect reflects the different amino acid substitution or the molecule in which the mutation occurs. Substitutions of cysteine for glycine within the triple helical domain of the proα1(I) chain result in a kinked molecule, apparently due to conformational distortion induced by the interchain disulfide bond between two abnormal α1(I) chains (Vogel et al., 1988); substitutions by other residues do not appear to have the same effect (Lightfoot et al., 1992). It is possible that the cysteine residues permit re-nucleation to occur amino-terminal to the disulfide bond.

Studies using a minigene version of type III procollagen (Lees and Bulleid, 1994; Bulleid et al., 1996) have recently shown that triple helix formation is dependent both on complete hydroxylation and on the presence of at least two of the six triplets that contain hydroxyproline in the very carboxyl-terminal end of the triple helix (equivalent to those positions 1012, 1015, 1018, 1021, 1024 and 1027 in the full-length molecules); the mutations I and others identified at positions 1009, 1012, 1015, 1018, and 1021 all interfere with nucleation. Thus, not only is the hydroxylation of the residues important, but the conformation of the helical domain appears to be essential.

Neither the formation of interchain disulfide bonds at the carboxyl-terminal end of the triple helix, nor in the propeptide region of the chain (the usual domain for interchain disulfide bonds) is necessary for triple helix nucleation or propagation (Bulleid et al., 1996). In the molecules I studied with substitutions for glycine within the last 20 residues of the triple helix, interchain disulfide bonds formed in the carboxyl-terminal propeptide but the telopeptide and triple helical cysteines did not form interchain links. My data, combined with that from the COL3A1 minigene experiments (Lees and Bulleid, 1994; Bulleid et al., 1996) suggest that the order of events during assembly is probably propeptide disulfide bond formation, then helix nucleation and propagation, with telopeptide disulfide bond formation a late event.

How far away can a mutation occur and perturb the disulfide bond formation at amino acids 1029 and 1030 in type III collagen? Nuytinck, et al. (1994) noticed a monomer in cells from a patient with a glycine to valine substitution at amino acid 1009 but were uncertain of the mechanism by which it had been produced. Reports of other mutations in this region, G1000V (Richards et al., 1993), G1003D (Richards et al., 1993), G1006E (Richards et al., 1993; Mackay et al., 1996), G1006E, (Richards et al., 1992; Johnson, et al., 1992), G1018D (Kontusaari et al., 1992), and G1021E (Narcisi et. al., 1993) do not describe an abnormal band after proteolysis and in the reports of the G1018D (Kontusaari et al., 1992) and G1021E (Narcisi et al., 1993) mutations the monomer is not apparent. Although mutations in the span from 1009 to 1021 are sufficient to alter helix formation and produce such an effect, conditions of labeling

may vary enough that the effect is not always seen. From my studies, a mutation 125 residues away is beyond the realm of the effect as substitution G904V did not affect disulfide bond formation.

Bateman and colleagues (1987) demonstrated that substitution of arginine for glycine at amino acid 391 in the α1(I) chain of type I collagen resulted in hydroxylation of a normally non-modified lysyl residue 17 amino acids carboxyl-terminal from the site of mutation. Although it is tempting to interpret such findings as consistent with a retrograde "unwinding" effect, they are probably better explained by the concept of cooperative folding blocks within the triple helical domain (Privalov, 1982). Westerhausen and colleagues (1990) demonstrated the significance of the microenvironment in which a mutation resides in two individuals with perinatal lethal osteogenesis imperfecta (OI II) where replacement of obligate glycines with serines had very different effects on thermal stability even though the residues were only 35 amino acids apart. The effect on stability was assayed by judging the resistance to proteases of type I collagen. The collagens are heated and subjected to proteases that are active only on monomeric chains not the triple helix. The resulting fragments are electrophoretically analysed to determine at what temperature the helix destabilizes. The individual with serine α1-631 had a thermal unfolding temperature of 41°C, comparable to control cells, while the cells from the individual with serine α1-598 produced collagens that unwound at 20°C (Westerhausen et al., 1990). This drastic

difference in melting temperatures corroborates experimental results indicating that collagen is not a stabile rod but is instead composed of microenvironments that "breathe" or unfold at physiologic temperatures (Privalov, 1982).

Additionally, thermodynamic considerations (Johnson et al., 1992) and refolding experiments with type III collagen molecules anchored by the carboxyl-terminal disulfide bonds (Davis and Bachinger, 1993) suggest that the folding could be "punctuated" in the sense that there are regions of high helix stability that alternate with those of relatively low stability. Helix propagates through the highly stable regions quickly and then pauses as the next region of high stability associates, allowing the intervening regions to form helix as well. The carboxyl-terminal region of the triple helix appears to be one of high stability in which the concentrations of proline and hydroxyproline are higher than in other domains. Substitutions for glycine residues within this domain create regions that must have very low stability and cannot act as nucleation sites. Nucleation brings chains into proximity, initiates helix formation and allows disulfide bonds to form. Mutations in the last 25 residues of the triple helix, which are the first to form triple helix, appear to be able to abrogate nucleation at physiologic temperatures so that disulfide bond formation does not occur. Nucleation would have to occur in a more amino-terminal domain, but happens only at lower temperature with alcohol precipitation and with prolonged incubation used for pepsin digestion. The absence of data on folding of chains that carry mutations between 1009

and 904 makes it difficult to assign the dimensions of the first block of cooperative folding. An exon 47-skipping mutation that removes the sequences from amino acid 955-972 does not appear to interfere with the initial phase of nucleation in that no monomeric chains are released following pepsin digestion (Schwarze et al., 1997). It is possible that this mutation lies outside of the nucleation sequence but also possible is that the sequence can be deleted but not interrupted.

If helix propagation is disrupted during molecular assembly, can it be re-initiated at the usual cellular temperatures? My data suggest that while folding can be propagated at low temperatures with long incubations, short incubations at 4°C, 20°C, or 37°C are insufficient to allow helix re-initiation. Such failure to re-initiate may explain the phenotypic effects of mutations near the amino-terminal end of the triple helix in that the molecules are not secreted as well as might be expected and, no doubt, do not aggregate well into fibrillar structures which require a very regular contour of the molecule (see chapter V for an examination of dermal fibrils).

MATERIALS AND METHODS

Cell strains, tissue culture, and antibodies - Patient and control dermal fibroblast cultures were established with appropriate consent and maintained as previously described (Bonadio et al., 1985). When added, ascorbate was used at a final

concentration of 50 µg/ml. A polyclonal antibody directed against the amino terminal propeptide of human type III procollagen was used for immunoprecipitations (Niemela et al., 1985).

Procollagen and collagen analysis - Procollagen and collagen labeling with [^{3}H]proline (Amersham Corp.) and separation by SDS-PAGE was performed as previously described (Bonadio et al., 1985) except that proteins were precipitated with 70% ethanol (v/v final concentration) for one hour on ice after addition of unlabeled collagen (Sigma Chemical Co.) rather than dialyzed into ammonium bicarbonate. To label with [^{35}S]cysteine (ICN Pharmaceuticals, Inc.), cells were pre-pulsed for 2-4 hours in DMEM (GIBCO BRL) without serum or cysteine but with ascorbate. Cells were labeled with 30µCi [^{35}S]cysteine (>1000 Ci/mmol) overnight in 0.7 ml serum-free, cysteine-free DMEM with ascorbate and 10% complete DMEM (Irvine Scientific). Immunoprecipitations were performed at 4° C. Antigen-antibody complexes were precipitated with protein-G sepharose (Pharmacia LKB Biotechnology, Inc.) following the protocol of Maniatis (Sambrook et al., 1989).

Folding of triple helix - The extent to which a triple helix was able to form was measured by its resistance to the proteases trypsin (Worthington Biochemical Corp.) and chymotrypsin (Boehringer, Mannheim, Germany). Cell strains were labeled with 100 µCi [^{3}H]proline for 16 hours. They were then incubated for 40' at 4°, 20° or 37° C. Following incubation, cells were washed with PBS and lysed in 1% NP-40 (Sigma Chemical Co.), 2mM EGTA buffer and then scraped into eppendorf tubes to which

80μg/ml trypsin, 200 μg/ml chymotrypsin were added and incubated for 2' at 20° C. The reaction was stopped by the addition of 500 μg/ml soybean trypsin inhibitor (Sigma Chemical Co.) and 100 μg/ml PMSF (Sigma Chemical Co.). The samples were separated by SDS-PAGE following the addition of 2X Laemmli sample buffer. To compare the collagenous proteins of the folding experiments with those harvested under typical conditions, cell strains were labeled in parallel and harvested as described above under ***procollagen and collagen analysis.***

Figure 3.1. **Proα1(III) chains from EDS IV individuals with mutations near the carboxyl-terminal end of the triple helix migrate as monomers.** Dermal fibroblasts were labeled with [^{3}H]proline for 16 hours, the proteins harvested and subjected to partial proteolysis with pepsin and then separated without reduction by SDS-PAGE. The association of proα1(III) chains into disulfide bonded trimers is seen in the control cells of lane 1 and 5. Interchain disulfide bonds do not form in cells with mutations G1012R, G1018D, and G1021E (lanes 2-4, respectively), instead the proα1(III) molecules migrate as monomers.

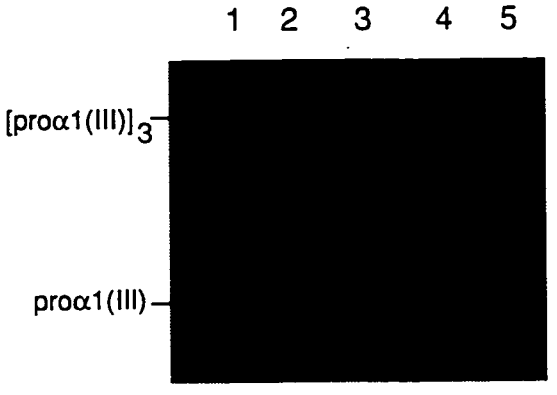

1 2 3 4 5
[proα1(III)]3
proα1(III)

Figure 3.2. **Mutations near the C-terminal end of the triple helix interfere with disulfide bond formation.** Fibroblasts with mutations G1012R, G1018D and G1021E were labeled in parallel with [^{3}H]proline or [^{35}S]cysteine. Type III procollagen was immunoprecipitated from the [^{35}S]cysteine-labeled cell layer with an antibody directed against the amino-terminal propeptide of type III procollagen. The immunoprecipitated protein was cleaved with pepsin to yield collagen. [^{3}H]proline-labeled collagens were generated by partial proteolysis of ethanol-precipitated cell layer proteins harvested from each cell line. The collagens were separated without reduction by SDS-PAGE. The normal migration of type III collagen as a trimer is evident in the [^{3}H]proline-labeled control but little or no trimer is seen in the EDS IV cells. Patient cells exhibit a protein that migrates in the position of a type III monomer (top arrow). The delay in the mobility of the [^{35}S]cysteine-labeled sample appears to be an artifact of immunoprecipitation as it was also seen in a [^{3}H]proline-labeled sample that was immunoprecipitated under the same conditions (data not shown).

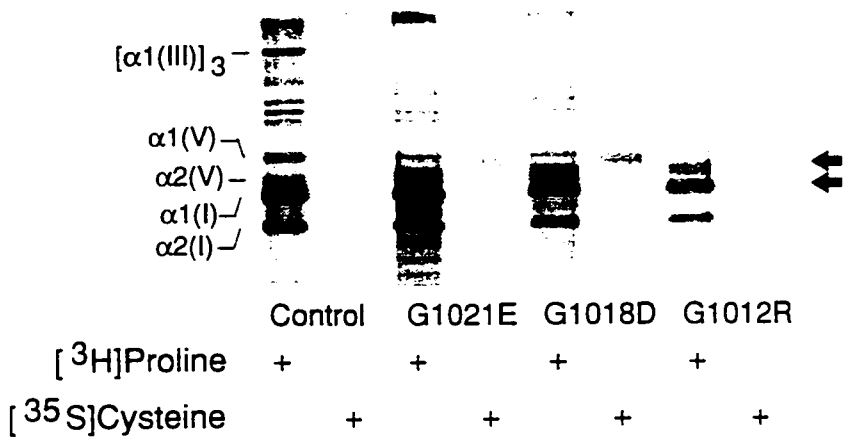

[α1(III)]3
α1(V)
α2(V)
α1(I)
α2(I)
Control
G1021E
G1018D
G1012R
[3H]Proline
[35S]Cysteine
+
+
+
+
+
+
+
+

Figure 3.3. **Type III procollagen chain assembly is not impaired.** Cells harboring the G1021E mutation were labeled for 16 hours with [^{3}H]proline. The medium and cell layer proteins were harvested and precipitated with ethanol and then separated after reduction of inter-and intrachain disulfide bonds (+DTT) or without reduction (-DTT). Under reducing conditions there is an abundant band in the cells (arrow) that migrates above the proα1(I) band which is the over-modified proα1(III) chains. This band is not present in the absence of reduction (-DTT) which indicates that the interchain disulfide bonds in the carboxyl-terminal propeptide form normally.

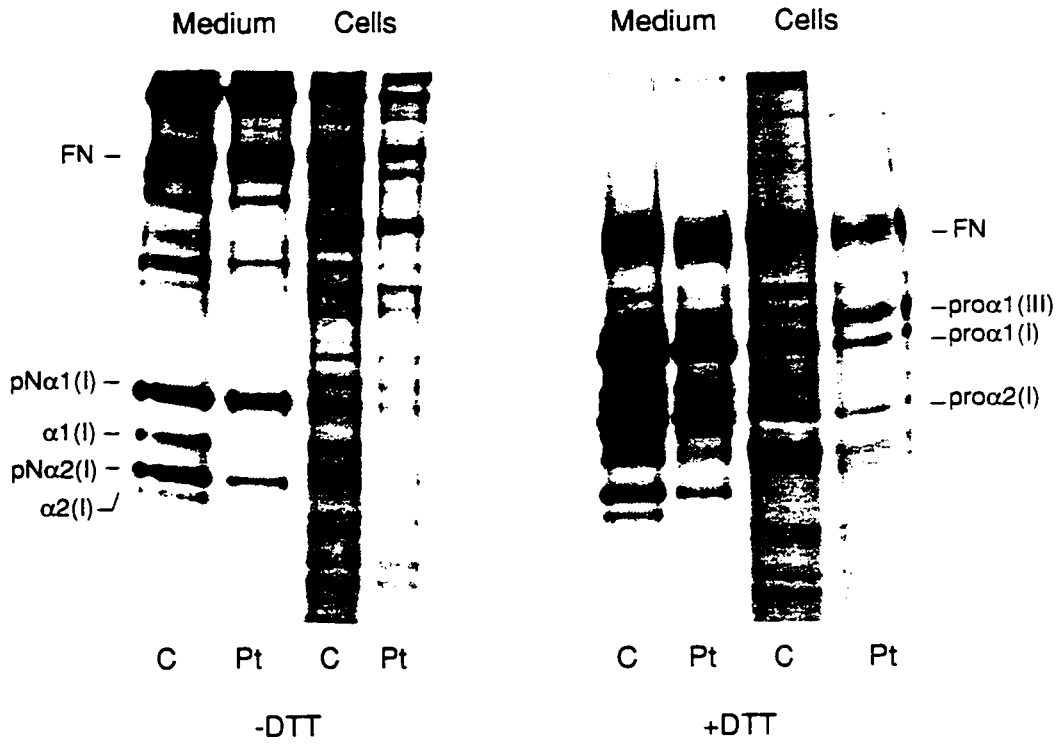

Medium
Cells
FN –
pNα1(I) –
α1(I) –
pNα2(I) –
α2(I) –
C
Pt
C
Pt
-DTT
Medium
Cells
– FN
– proα1(III)
– proα1(I)
– proα2(I)
C
Pt
C
Pt
+DTT

Figure 3.4. **Analysis of triple helix propagation.** Dermal fibroblasts with mutations G061E, G385E, and G1021E were labeled with [^{3}H]proline at 37°C for 16 hours, the medium changed and the cells shifted to 4°C, 20°C, or 37°C for 40 minutes to allow the triple helix to propagate. The intracellular proteins were harvested by cell lysis and treated with trypsin/chymotrypsin at 20°C for 2 minutes and the proteins separated without reduction by SDS-PAGE. At all temperatures the G061E and G385E molecules could be assembled and triple helix propagated to the region of the mutation but no further (arrows). The molecules with the G1021E mutation did not form a stable triple helix at any temperature except for a very small amount representing the normal chains. For comparison, cell layer proteins were harvested from each cell strain in parallel by ethanol-precipitation and collagens were generated by partial proteolysis with pepsin (left lane in each panel).

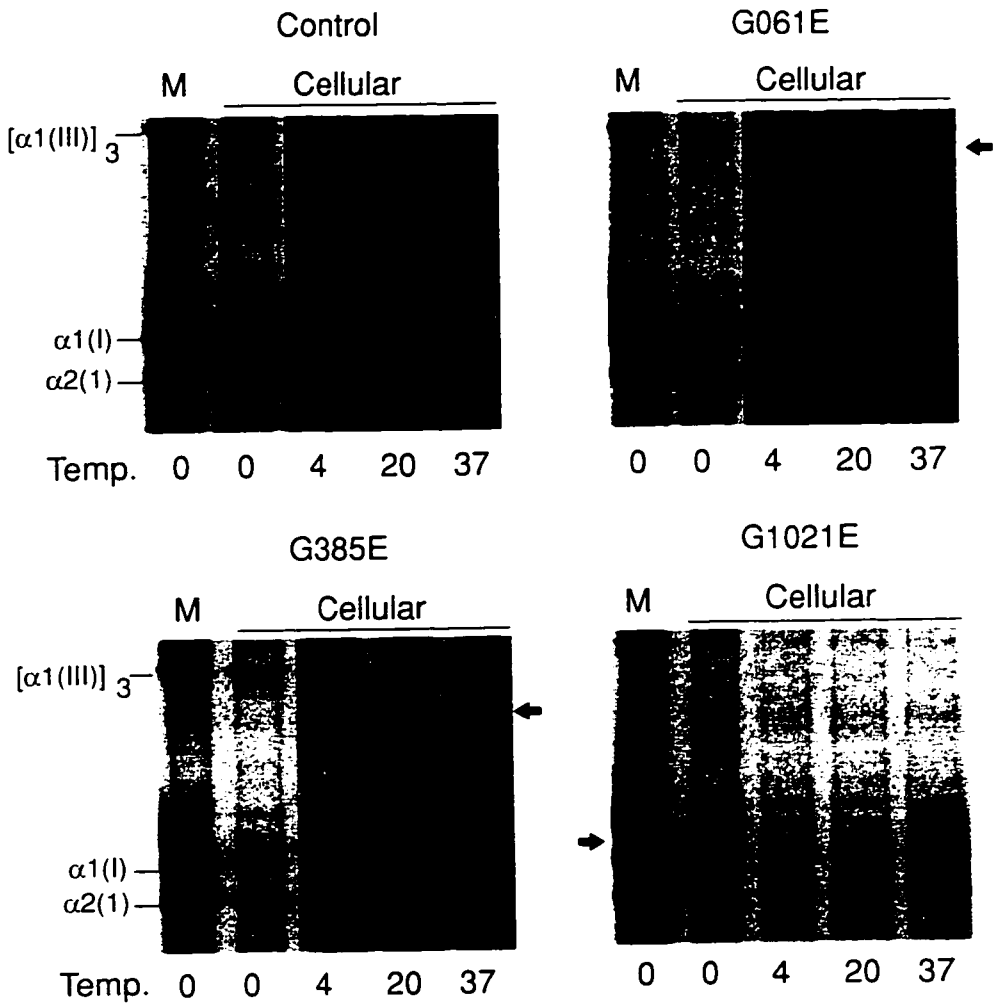
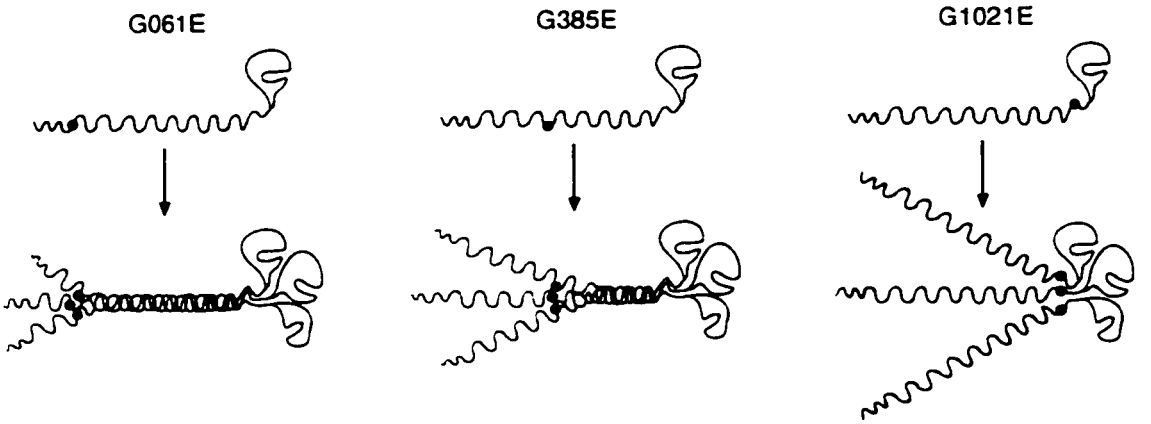

Control
M
Cellular
[α1(III)] 3
α1(I)
α2(1)
Temp. 0 0 4 20 37
G061E
M
Cellular
0 0 4 20 37
G385E
M
Cellular
[α1(III)] 3
α1(I)
α2(1)
Temp. 0 0 4 20 37
G1021E
M
Cellular
0 0 4 20 37

Figure 3.5. **Model of the effect of mutation on the nucleation and propagation of the triple helix in type III procollagen molecules.** The large dots represent the positions of the mutations in cells that harbor the G061E, G385E, and G1021E mutations. For the G061E and G385E mutations chain assembly and triple helix nucleation occur normally but propagation of helix does not extend beyond the mutant residue and at 37°C re-nucleation does not occur. In cells that harbor the G1021E mutation, chain assembly ocurs but triple helix nucleation does not and there are no additional sites of nucleation in the more amino-terminal regions of the chains. This model is for molecules that contain 3 abnormal chains; the effect on molecules with 1 or 2 abnormal chains is uncertain.


G061E
G385E
G1021E

CHAPTER IV

INTERACTION OF TYPE III PROCOLLAGEN WITH RESIDENT PROTEINS OF THE ENDOPLASMIC RETICULUM

EXPERIMENTAL QUESTION

Abnormal type III procollagen is often overmodified and retained within the cell as a result of mutation, especially when the mutation occurs in the region of COL3A1 that corresponds to the carboxyl-terminal end of the molecule. This chapter addresses the question: is there an ER-resident protein(s) involved in the retention of malfolded type III procollagen in the cells of EDS IV individuals?

INTRODUCTION

The endoplasmic reticulum is the first stop for proteins traversing the secretory pathway. It is here that nascent proteins fold, assemble, and are initially modified. When one of these processes goes awry, the misfolded, incompletely assembled or improperly modified protein is retained within the ER rather than exported, a phenomenon termed "quality control" (Hurtley and Helenius, 1989). Components of the quality control machinery are folding enzymes such as protein disulfide isomerase, peptidylprolyl isomerase, modifying enzymes including lysyl hydroxylase, glucosidases I and II, and molecular chaperones such as BiP, GRP94, and calnexin.

Type III procollagen is a trimeric secretory protein that is retained within cells of individuals with Ehlers-Danlos syndrome type IV (EDS IV) as a result of improper folding or assembly due to mutation. Chessler and Byers (1992, 1993) studied abnormal type I procollagens (the prototype of fibrillar collagens) from osteogenesis imperfecta (OI) patients and found that they associate with the ER-resident proteins, BiP (GRP 78), and protein disulfide isomerase (PDI), the β subunit of the collagen modifying enzyme prolyl hydroxylase. Their work indicates that the choice of binding partner depends on the type of procollagen mutation. The stress response protein, GRP 78/BiP, binds to procollagens that have impaired chain association as a result of mutations affecting their C-propeptides. GRP78/BiP does not bind to procollagens with mutations affecting the triple helical domain; instead procollagens with this type of mutation have been seen associated with prolyl hydroxylase (Chessler and Byers, 1992).

To better understand the molecular consequences of EDS IV, I examined the role of members of the ER quality control system in retaining abnormal type III procollagen. I concentrated on PDI because all of the type III collagen mutations discovered thus far affect the triple helix, calnexin because it mediates the retention of malfolded proteins in other human genetic diseases (Bergeron et al., 1994), and hsp47 because it is a collagen-specific chaperone thought to mediate the prolonged retention of unstable type I procollagen molecules (Nakai et al., 1992). I chose to study cell

strains representing the same amino acid substitution in different regions of the triple helix, exon skipping mutations affecting either end of the helix and point mutations resulting in highly overmodified proteins exhibiting intracellular retention.

RESULTS

Radiolabeled EDS IV cell strain lysates and immunoprecipitations with anti-type III procollagen antibodies reveal no known associating proteins

To identify proteins involved in the retention of abnormal procollagen, I first examined EDS IV cell lines for the induction of proteins. Patient and control cells were metabolically labeled with Tran^{35}S-Label™ (composed of [^{35}S] methionine and [^{35}S] cysteine), lysed, and the labeled proteins separated by SDS-PAGE. EDS IV cells chosen for study include those with mutation G1021R which has unique folding properties (see Chapter III), cell line G904V which produces a highly overmodified protein that is retained within the cell, a cysteine substitution at residue 385 that results in two species of retained procollagen and, finally, two multiexon skipping mutations, one at each end of the triple helix (8-11 and 34-48). Intracellular and secreted proteins were electrophoresed on mini-gels and compared. No differences were apparent between control cells and the cells derived from EDS IV patients (figure 4.1). Intracellular proteins from another set of EDS IV patient cells including point mutations G1021R, G1021E, G769R, G757C, G373R, and G061E were electrophoresed on a larger size gel and compared with control proteins. In the cell

line with the G061E mutation a protein migrating at approximately 65 - 66kD stands out (lane 7, figure 4.2). This band is present in the control cells and other patient samples but its intensity relative to other bands appears quite different in the G061E cell strain. Protein disulfide isomerase has a molecular weight of 65kD and is a candidate for mediating intracellular retention of aberrant procollagens. However, anti-PDI antibodies did not co-precipitate proα1(III) in this cell strain (see below).

To establish whether a specific interaction occurs between mutant type III procollagen retained in EDS IV cell strains and any ER-resident proteins, antibodies directed against type III procollagen were used for immunoprecipitations and the antigen-antibody complexes analyzed by SDS-PAGE. Figure 4.3 shows the results of immunoprecipitations with antibody PIIINP, directed against the pN III peptide of bovine procollagen that cross-reacts with human. The SDS-PAGE analysis depicted in figure 4.3(a) shows EDS IV cell strains containing an arginine for glycine substitution at amino acids 34, 373, 769, and 1021 along the type III procollagen triple helix. The patient proteins precipitated are the same as those in the control cells with the exception of a protein of approximately 110-120kD in the cell strain with mutation G769R. There is also a very faint band migrating at this position in the cell strain harboring mutation G373R. Figure 4.3(b) features two patient cell lines with a protein migrating at this spot, G904V and G1021E. The other patient cell lines, G685C and the two strains with exon-skipping mutations show no evidence of a band in the 110-120kD range, nor do the control cells. Immunoprecipitations with another anti-type III

procollagen antibody, SIF 98, also brought down a protein of 110-120kD (figure 4.4). Using this antibody, the 110-120kD protein was observed in the control cells and all of the patient cell strains. Judging the abundance of the protein and how it differed between the various cell lines is difficult because SIF 98 precipitates many proteins in addition to type III procollagen creating a high background signal.

To investigate whether the 110-120kD protein is an ER-resident and might, therefore, be implicated in retaining abnormal type III procollagen in EDS IV patients, Western blots were performed. Permanent residence within the ER requires a COOH-terminal retention signal, usually composed of the amino acids lysine, aspartic acid, glutamic acid, leucine (KDEL), or association with a protein that possesses this signal (Munro and Pelham, 1987). A monoclonal antibody specific for the KDEL amino acid sequence was used to probe lysates of patient cell strains with arginine for glycine substitutions along the triple helix, G034R, G769R, G1021R. The antibody recognizes several proteins possessing the KDEL signal migrating at approximately 55, 78, 87, 94, and 120 kD (figure 4.5). The bands migrating at 55, and 87 kD are uniform while those migrating at 94, and 120kD vary in intensity. The 120kD protein is the only one that differs significantly between control and patient cells, it appears to be relatively uninduced in control cells and much more prevalent in the cells derived from EDS IV patients.

Mutant type III procollagen does not associate with PDI

Previously, Chessler and Byers (1992) demonstrated that malfolded type I procollagen associates with protein disulfide isomerase (PDI), the β subunit of the collagen modifying enzyme prolyl-4-hydroxylase, in the cells of OI patients. To determine whether a similar association occurs between abnormal type III procollagen and PDI, immunoprecipitations were performed using anti-PDI antibodies. Control and patient cells were incubated with TranS^{35}Label™, lysed, and the radiolabeled proteins immunoprecipitated, separated by SDS-PAGE and examined by autoradiography. PDI is easily identified as the protein migrating at 65kD but the protein that migrates in the position of type III procollagen in lanes 4 and 5 is very faint (figure 4.1). Two proteins migrate around the position of fibronectin (FN). One is probably fibronectin and the other may be PDI complexed with other proteins. Another high molecular weight complex is visible near the very top of the gel and it, too, may be PDI complexed with other proteins or it could be a large aggregrate that is not reduced and does not easily enter the gel. The same high molecular weight bands are visible in lane 5 containing proteins from the EDS IV individual harboring mutation G1021E, although no band appears in the position of type III procollagen (figure 4.6). Proteins from individuals with mutations G034R, G373R, G769R, G1021R, G061E, G385E, G1021E, and a deletion of exons 34-48 were immunoprecipitated and examined with similar results (data not shown).

Calnexin does not associate with type III procollagen

To determine whether abnormal type III procollagen associates with the integral membrane protein chaperone calnexin, lysates from radiolabeled patient and control cells were immunoprecipitated with antibodies against calnexin. The cell strains chosen include glutamic acid for glycine substitutions in different regions of the triple helix (G385E, G1021E), a highly overmodified and retained type III procollagen (G 904V). Figure 4.2 shows an autoradiograph of the SDS-PAGE analysis of these proteins labeled with TranS^{35}Label™. A dense band migrating at approximately 88kD represents calnexin. A band comigrating with that of type III procollagen appears in the control cells but not in any of the EDS IV-derived patient cells, however, the background labeling is extensive in this experiment making interpretation of the results difficult. To clarify the results of these immunoprecipitations, two EDS IV cell lines were labeled with [^{3}H]proline and the immunoprecipitations repeated. Labeling with proline should eliminate the high background as it is a rare amino acid in proteins other than collagens. To ensure that the immunoprecipitation procedure was successful, a TranS^{35}Label™ labeled lysate from the control cell line was included. As before, a band migrating in the position of type III procollagen was seen in the [^{35}S]-labeled control cell line. The [^{3}H]proline lysates revealed some background labeling in the area where the proα1(III) and proα1(I) migrate but no distinct band is apparent. It appears from these experiments that calnexin and type III procollagen, either in its usual conformation or in a misfolded form, do not associate (figure 4.3) .

Hsp47

To investigate the role of the collagen-specific chaperone, hsp47, in the retention of type III procollagen in the cells of EDS IV individuals immunoprecipitations were performed on patient and control lysates. Following overnight labeling with TranS^{35}Label™, lysates were made, immunoprecipitated with antibody against hsp47 and the proteins separated by SDS-PAGE. The autoradiograph shown in figure 4.9 indicates that in the control fibroblasts, hsp47 coprecipitates a protein that migrates as a wide band in the positions of types I and III procollagen (lane 3). The patient cell lines display proteins that migrate in the region of procollagen as does the irrelevant control. Type III procollagen from the patient with mutation G904V is a broad band migrating well above the control protein in collagen screening gels (see figure 2.1). A band reminiscent of this patient's type III procollagen is seen in lane 7, co-incident with the 205 kD marker, but the Tsp irrelevant antibody also immunoprecipitates a protein indicated by the faint band in this area as does the A8496 control and the patient with exons 34-48 deleted.

DISCUSSION

The results presented above are ambiguous. The only conclusion I can draw confidently is that calnexin does not mediate retention of aberrant type III procollagen in EDS IV cell lines. Although a coprecipitating protein migrates in the place of type

III procollagen in cells labeled with TranS^{35}Label™, it is not apparent in [^{3}H]proline-labeled samples. I also performed immunoprecipitations on [^{3}H]proline-labeled cell lines G769R and G757C. In this experiment, a very faint band migrating in the position of calnexin appeared in the EDS IV and control cell lanes but there was no evidence in any lane of type III procollagen even though the procollagen is much richer in proline than calnexin (data not shown).

The data relevant to the other ER-residents that were candidates for retaining malfolded type III procollagen is equivocal. A protein of the approximate molecular weight of PDI appeared to be induced in cell line harboring the G061R mutation but I was unable to confirm this with anti-PDI antibodies. The high molecular weight complexes seen in patient cells upon immunoprecipitation with anti-PDI antibodies may contain PDI but it is unclear from these experiments. Two separate problems contribute to the uncertainty of the experiments and can possibly be corrected. First, the PDI antibody is different than that used by Chessler and Byers when they successfully demonstrated an association between PDI and type I procollagen from OI patients (Chessler and Byers, 1993). When I do parallel immunoprecipitations with an antibody obtained from StressGen and compare the co-precipitating proteins to those brought down by the antibody that was a gift of Dr. Richard Berg used in the OI experiments, fewer bands are apparent with the former and the procollagen brought down is vanishingly small (data not shown). It is possible that less stringent conditions would improve the association between the PDI immunoprecipitated with the

commercial antibody and other cellular proteins. Another possible solution is to obtain antibodies from Dr. Berg once again. A second problem is that of background labeling. A shorter exposure to radioactivity may eliminate the labeling of proteins with a long half-life resulting in less overall background. The disadvantage to a shorter labeling period is that type III procollagen may still not be obvious because it is less abundant than the type I visualized by this technique.

The results of the immunoprecipitations with Hsp47 are also difficult to interpret due to high background labeling. Again, changing the labeling protocol may make it possible to unequivocally conclude that the bands apparent in cell strains G940V and G385E are type III procollagen. Another approach is to perform Western blots on cell lysates from these patients. Generally, radiolabeling is much more sensitive as well as easier to quantitate but blotted proteins visualized by chemiluminescence can also be quantified. Western blotting could also be used to determine the amount of Hsp47 present when lysates are immunoprecipitated with anti-type III procollagen antibodies. This too has a drawback; cross-reactivity is sometimes seen between the immunoglobulin chains present in the immunoprecipitating protein and that detected by the probe antibody.

The 120kd protein co-precipitated with type III procollagen by the PIIINP anti-type III procollagen antibody is a promising candidate for an interacting protein. It appears in five patient cell strains in two separate sets of immunoprecipitations. Its molelcular weight parallels that of a ER-resident protein and it is more abundant in

EDS IV-derived cells than the control cell strain. Neither the literature nor the supplier of the antibody have any suggestions as to this protein's identity so the next logical step is to identify it. The unknown protein can be partially purified with affinity chromatography (using the G904V mutant protein for binding) followed by electrophoresis and transfer to a polyvinylidene difluoride membrane for partial sequence determination. Then a human fibroblast libray could be probed to identify the encoding gene.

Several alternative methods could be used to investigate the interaction between ER proteins and type III procollagen. First, cross-linking the proteins within the ER would capture a fleeting interaction or one disrupted by the conditions of immunoprecipitation. The membrane-permeant thiol-cleavable cross-linker dithiobis(succinimidyl propionate) has been successfully used for *in situ* crosslinking of PDI to immunoglobulins and calnexin to Class I histocompatibility molecules (Roth and Pierce, 1987; Degen and Williams, 1991). However, background labeling is also a drawback to crosslinking as, again, a long labeling period is required to adequately label type III procollagen and crosslinkers often link many proteins not just those of interest, conceivably producing even more background.

Another option is the creation of a protein expression system. The human kidney tumor cell line HT1080 has been used for the expression of type II procollagen, another of the fibrillar collagens (Fertala et al., 1994, Ganguly et al., 1994). The system has advantages over using primary cell lines, any mutation can be introduced so

that a systematic examination of the effect of different mutations is easily accomodated; because HT1080 cells are immortal they do not senesce like primary cells and they grow rapidly. Overexpression, or at least, higher expression of type III procollagen might allow it and proteins with which it interacts to be distinguished from background labeling. It is possible, however, that the procollagen associates with a very small component of the total population of a given ER-resident protein so increasing the level of type III procollagen will not improve its perceptibility. It is also possible that overexpression of procollagen will swamp any protein with which it associates.

Alternatively, protein interactions can be assessed by the genetically-based yeast two hybrid system. Each hybrid is composed of a protein of interest, a reporter gene, and either a transcriptional activation or DNA binding domain. When the two hybrids are expressed in the same cell and the two proteins associate, the activation domain is brought close to the binding domain resulting in the creation of a functional activator that is measured by a reporter gene (Fields and Song, 1989). The system has been used successfully in yeast and other mammalian cells for a wide variety of proteins including extracellular matrix proteins. As the genes for the ER-resident proteins BiP, calnexin, hsp47 and collagen modifying enzymes prolyl and lysyl hydroxylase have been isolated, the two hybrid system could be used to investigate the association of these proteins with mutant type III procollagen.

MATERIALS AND METHODS

Cell strains and antibodies

Patient and control dermal fibroblasts were established with appropriate consent and maintained as previously described (Bonadio et al., 1985). The control cell strain, A8496, was derived from an infant who died of hyaline membrane disease in the perinatal period. When added, ascorbate was used at a final concentration of 50μg/ml.

To immunoprecipitate type III procollagen two antibodies were used. A rabbit polyclonal antibody directed against the human amino-terminal propeptide known as PIIINP was a gift of Dr. Leila Risteli (University of Oulu, Finland). SIF 98, guinea pig antisera against pepsin-treated bovine type III collagen that cross reacts with human was a gift of Dr. Betty Nusgens (University of Liege, Belgium).

A mouse monoclonal antibody was used for the detection of PDI on Western blots and for immunoprecipitation (Affinity Bioreagents, Inc).

A rabbit polyclonal antibody directed against the carboxyl terminus of canine calnexin and cross reacting with human was used for immunoprecipitations and Western blotting (StressGen Biotechnologies Corp.).

Monoclonal antibody SPA-470 was used for immunoprecipitation of Hsp47 (StressGen Biotechnologies Corp.).

Metabolic Labeling

Newly confluent 35 mm dishes of cells were pre-incubated in 1ml of Dulbecco-Vogt Eagle's medium containing 50μg/ml ascorbate for several hours. The medium

was then replaced with fresh serum-free medium containing the appropriate label and incubated overnight. To label collagenous proteins 750 µl of medium containing 100 µCi of [2,3,4,5-^{3}H]proline (101 Ci/mmol; Amersham Corp.) was added. To label other proteins 700 µl of methionine and cysteine deficient medium containing 100 µCi of Tran^{35}S-Label™ (ICN Pharmaceuticals), 10% complete media, and ascorbate was added.

Immunoprecipitations

Immunoprecipitations were performed at 4° C according to the protocol of Maniatis (Sambrook et al., 1989). Cells were lysed in 1% NP-40 single detergent buffer containing 10 mM phenylmethylsulfonyl fluoride. Antigen-antibody complexes were precipitated with protein-G sepharose (Pharmacia LKB Biotechnology, Inc.). Bovine serum albumin was used instead of gelatin in NET-gel buffer.

Proteins were separated by SDS-PAGE in the presence of 2M urea and detected either by autoradiography or autoradiofluorography with EN^{3}HANCE (DuPont-New England Nuclear Research Products).

Western Blotting

Proteins were transferred from either 5% or 7.5% SDS-PAGE gels to 0.2 µ nitrocellulose (Schleicher and Schuell, Inc.) according to the protocol of Maniatis (Sambrook et al., 1989). Blots were blocked with 5% (w/v) nonfat milk, probed with

primary antibody followed by alkaline phosphatase-conjugated secondary antibody. Protein were detected with chemiluminescence (Tropix, Inc.) according the manufacturer's protocol.

Figure 4.1. **There is no increase in the synthesis of known proteins in EDS IV cells.** Dermal fibroblasts from patients and a control were labeled with TranS^{35}Label™ and incubated overnight in the presence of ascorbate and 10% complete DMEM. Samples were analyzed by SDS-PAGE under reducing conditions. Odd numbered lanes contain cell layer proteins; even numbered lanes are secreted proteins. Lanes 1 and 2 are control cells, 3 and 4 are G1021E, 5 and 6 are from a patient with uncharacterized mutation, 7 and 8 are G904V, 9 and 10 are G685C, 11and 12 are skipped exons 34-48, and 13 and 14 are skipped exons 8-11. The positions of molecular mass markers (kD) are shown on the right.

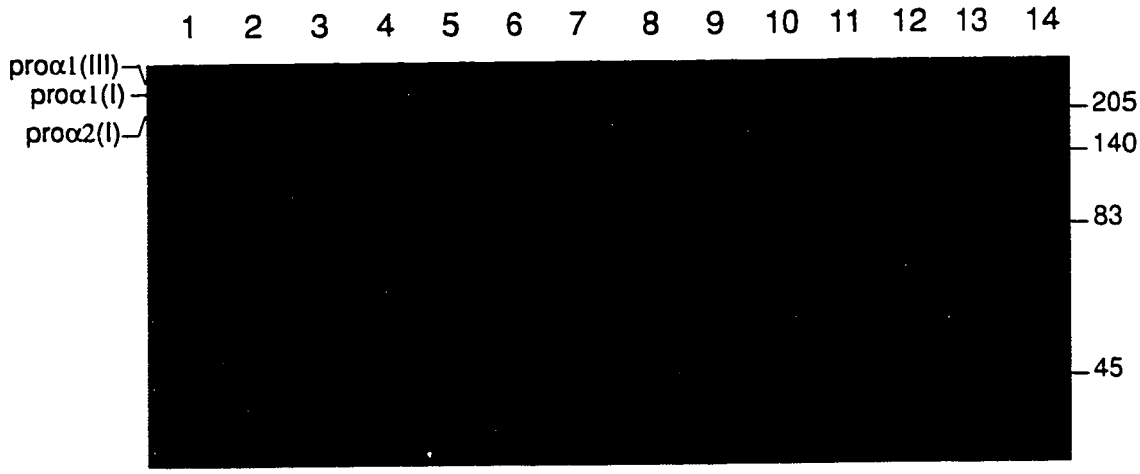

1 2 3 4 5 6 7 8 9 10 11 12 13 14
proα1(III)
proα1(I)
proα2(I)
205
140
83
45

Figure 4.2 **Increased synthesis of a 66kD protein in a cell strain with a G061E mutation.** Dermal fibroblasts from patients and a control were labeled with TranS^{35}Label™ overnight in the presence of ascorbate and 10% complete DMEM. Cell layer proteins were separated by SDS-PAGE under reducing conditions. Lane 1 contains a control cell line, lanes 2-7 contain cell strains harboring mutations G1021R, G1021E, G769R, G757C, G373R, G061E, respectively. The positions of molecular mass markers (kD) are shown on the right. Compared to the control the EDS IV-derived cells do not display increased synthesis of any protein with the exception of G061E (lane 7). This cell strain displays an increased amount of a 65-66kD (arrow) protein relative to its other proteins; this pattern of proteins is different than the control or any of the other EDS IV cell lines.

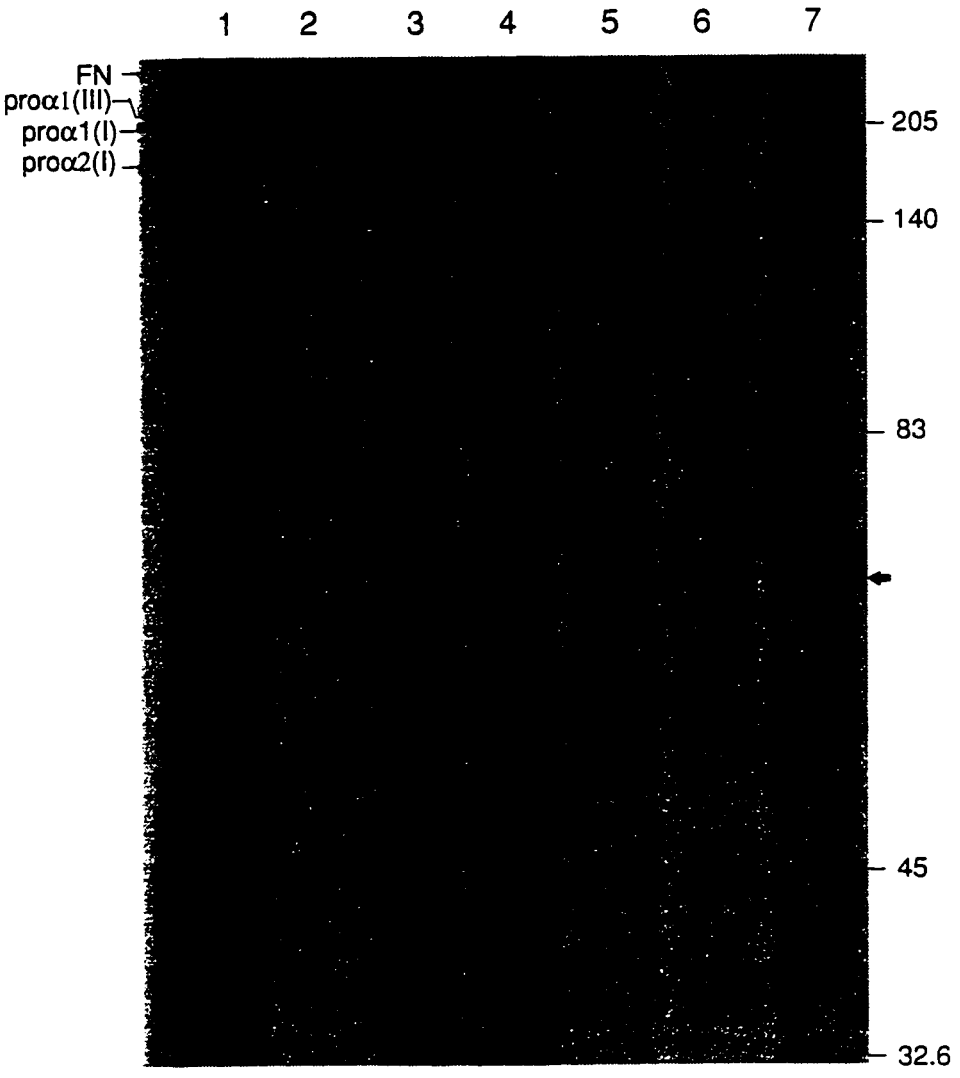

1
2
3
4
5
6
7
FN
proα1(III)
proα1(I)
proα2(I)
205
140
83
45
32.6

Figure 4.3. **SDS-PAGE analysis of immunoprecipitations with anti-type III procollagen antibody PIIINP.** Cells from EDS IV patients were labeled with TranS^{35}Label™ overnight in the presence of ascorbate and 10% complete DMEM. Cells were lysed in 1% NP-40 single detergent buffer, lysates immunoprecipitated with anti-type III procollagen antibody PIIINP and the precipitated proteins analyzed by SDS-PAGE under reducing conditions. a) [^{3} H]proline labeled media was harvested from the A8496 control cell line and immunoprecipitated with PIIINP (lane 3) and an anti-type I collagen antibody (LF9) (lane 1) to easily visualize migration of procollagens; a control lysate immunoprecipitated with LF9 is also included (lane 2). Controls include immunoprecipitations of A8496 lysates with PIIINP (lane 4), with an irrelevant antibody (against thrombospondin) (lane 9) and a lysate incubated with only protein G-sepharose beads used as a matrix to precipitate antigen-antibody complexes (lane 10). Proteins from patients with mutations G034R, G373R, G769R, G1021R are in lanes 5-8, respectively. The only notable difference in precipitated proteins is seen in lanes 6 and 7 where a band with an approximate molecular weight of 120 kD is seen. The positions of molecular mass markers (kD) are shown on the right.

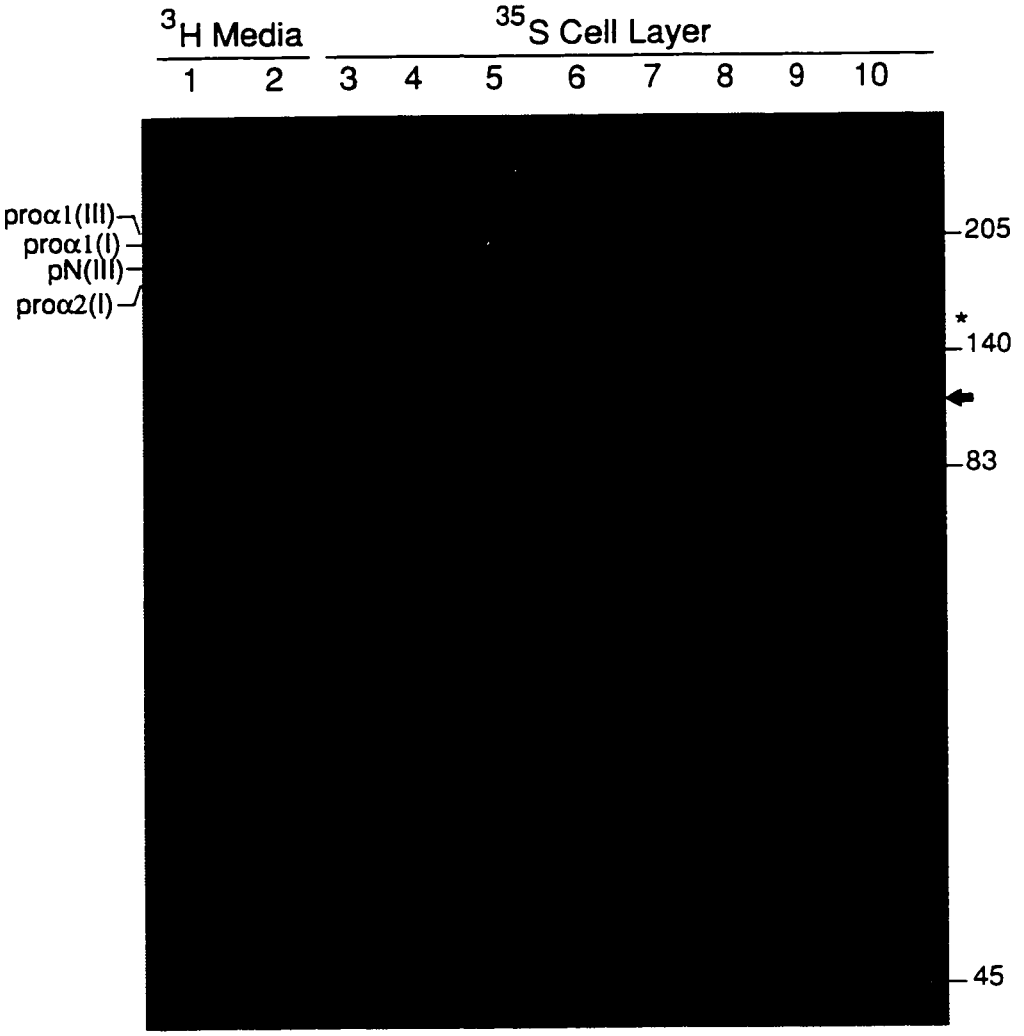

^{3}H Media
^{35}S Cell Layer
1
2
3
4
5
6
7
8
9
10
proα1(III)
proα1(I)
pN(III)
proα2(I)
205
*
140
83
45

Figure 4.3. b) A 120kD protein (←) is also seen in cells with mutations G904V (lane 7) and G1021E (lane 8). Lane 3 contains the control cell line. Lanes 4 and 5 contain cells with exon skipping mutations that delete exon 8-11 and 34-48, respectively. Lane 6 contains mutation G685C. [^{3}H]proline labeled media harvested from A8496 control cells and immunoprecipitated with PIIINP (lane 2) or LF9 (lane 1) shows the placement of procollagens. The irrelevant antibody and protein G-sepharose bead controls are shown in lanes 9 and 10, respectively. The positions of molecular mass markers (kD) are shown on the right.

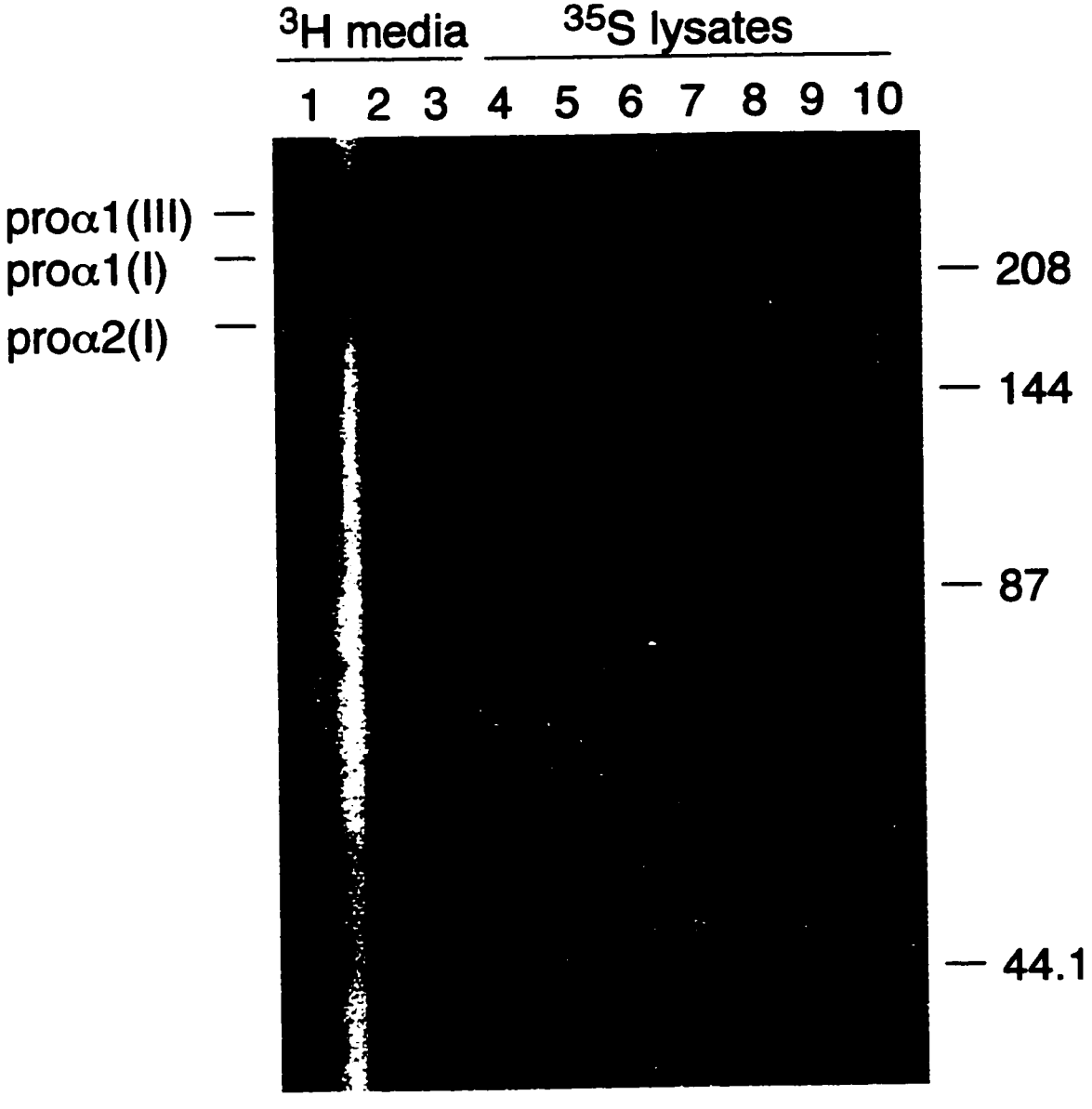

^{3}H media
^{35}S lysates
1 2 3 4 5 6 7 8 9 10
proα1(III)
proα1(I)
proα2(I)
208
144
87
44.1

Figure 4.4. **SDS-PAGE analysis of immunoprecipitations with anti-type III procollagen antibody SIF98.** Cells from EDS IV patients were labeled with TranS^{35}Label™ overnight in the presence of ascorbate and 10% complete DMEM. Cells were lysed in 1% NP-40 single detergent buffer, lystates immunoprecipitated with anti-type III procollagen antibody SIF98 and the precipitated proteins analyzed by SDS-PAGE under reducing conditions. A protein of approximately 120 kD seen in immunoprecipitations with another anti-type III procollagen antibody is apparent in all patient lysates as well as the control cells (←). This protein is not seen in control cells immunoprecipitated with anti-thrompospondin (lane 7), protein G-sepharose alone (lane 8) or in [^{3} H]proline labeled media immunoprecipitated with the anti-type III procollagen antibody SIF98 (lane 2) or anti-type I procollagen antibody LF9 (lane 1). The positions of molecular mass markers (kD) are shown on the right.

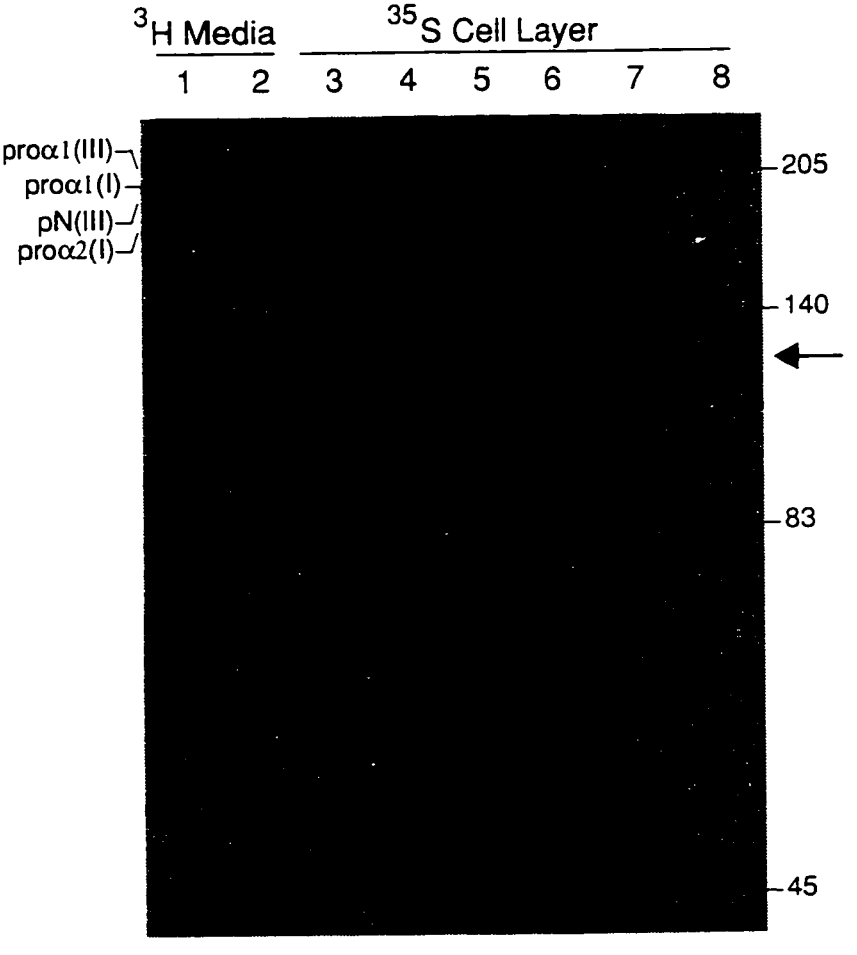

^{3}H Media
^{35}S Cell Layer
1
2
3
4
5
6
7
8
proα1(III)
proα1(I)
pN(III)
proα2(I)
205
140
83
45

Figure 4.5. **Western blot probed with anti-KDEL antibody.** Cell lysates were separated by SDS-PAGE, transferred to 0.2μ nitrocellulose, probed with anti-KDEL primary antibody followed by an alkaline-phosphatase conjugated secondary antibody and detected by chemiluminescence. EDS IV patients with mutations G034R, G769R, G1021R (lanes 2-4, respectively) exhibit increased synthesis of an 120kD protein compared to A8496 control cells (lane 1). Molecular weight markers are shown on the left.

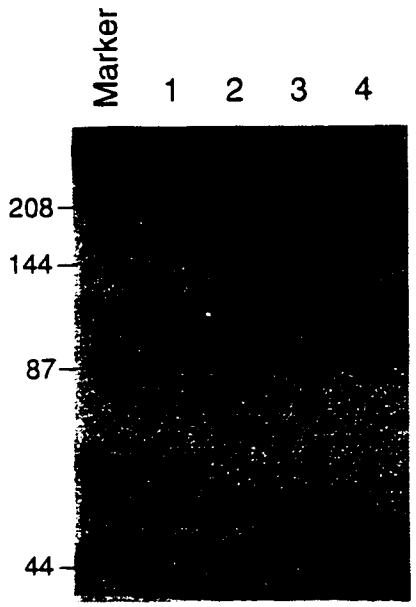

Marker
1
2
3
4
208
144
87
44

Figure 4.6. **SDS-PAGE analysis of immunoprecipitations with anti-PDI antibodies.** Dermal fibroblasts from a control and EDS IV patients with mutations G904V (lane 4) and G1021E (lane 5) were labeled overnight with [^{35}S]methionine and cysteine in the presence of ascorbate. They were harvested by cell lysis and the resulting samples immunoprecipitated with anti-PDI antibody and analyzed by SDS-PAGE. Lane 3 contains control fibroblasts, lane 6 an irrelevant antibody, lane 7 a lysate incubated with protein G- sepharose beads only. Lanes 1 and 2 contain [^{3}H]proline labeled media that was immunoprecipitated with antibodies against anti-type I procollagen, and anti-type III procollagen, respectively. In lane 4 (mutation G904V) there is a band migrating slightly above the control type III procollagen (lane 2) which is probably overmodified type III procollagen. There are also higher molecular weight bands that may be PDI complexed with other proteins. A high molecular weight protein is also seen in lane 5 (mutation) G1021E. The positions of molecular mass markers (kD) are shown on the right.

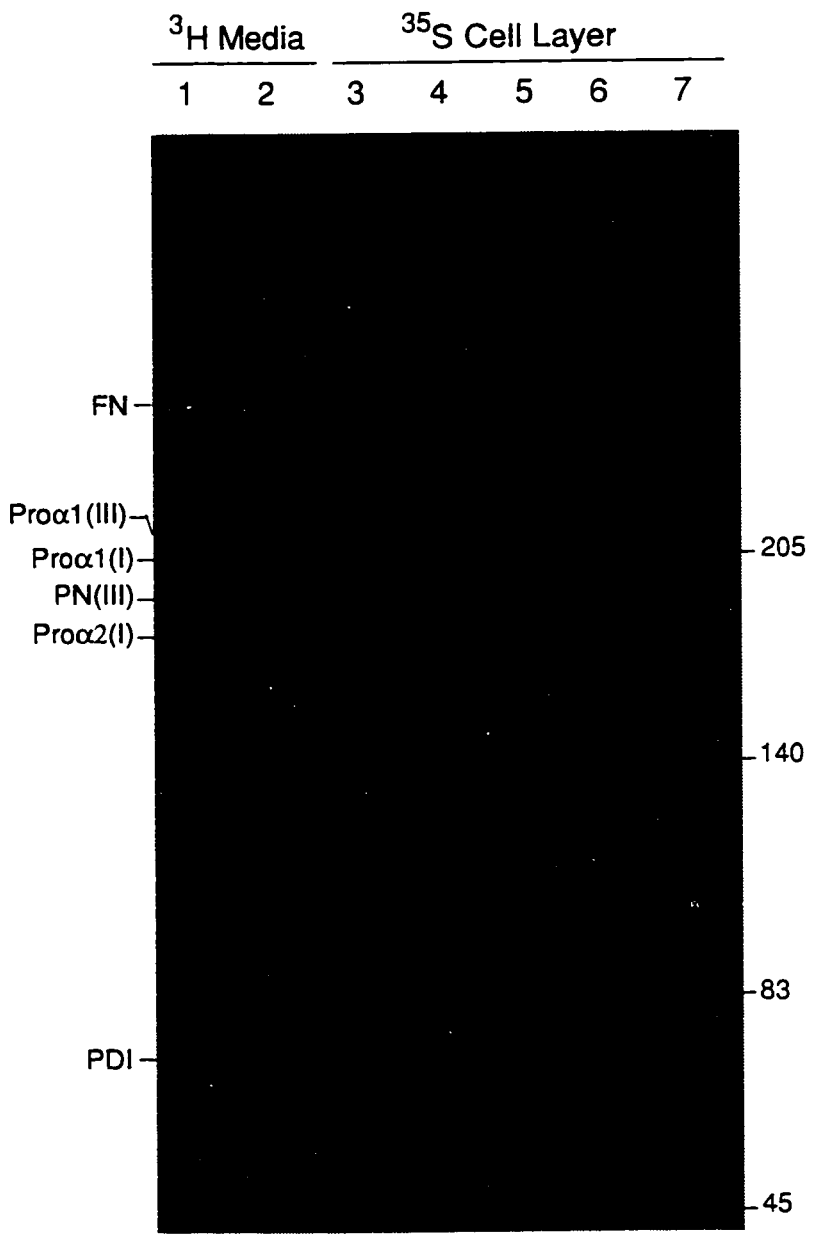

^{3}H Media
^{35}S Cell Layer
1
2
3
4
5
6
7
FN
Proα1(III)
Proα1(I)
PN(III)
Proα2(I)
PDI
205
140
83
45

Figure 4.7. **SDS-PAGE analysis of anti-calnexin immunoprecipitations.** Control and patient dermal fibroblasts were labeled with TranS^{35}Label™ overnight in the presence of ascorbate and 10% complete DMEM. Lysates were made, immunoprecipitations performed with anti-calnexin antibody, and proteins separated by SDS-PAGE. Migration of type I and type III procollagen is shown in lanes 1 and 2, respectively, by control cell medium that was immunoprecipitated with anti-type I or anti-type III antibodies. An irrelevant antibody control is shown in lane 8 and control lysate incubated only with protein G-sepharose beads in lane 9. Calnexin is seen in both the control and patient cells G385E, G904V, G1021E (lanes 4-6, respectively). However, co-precipitation of type III procollagen is seen only in the A8496 control cells. The positions of molecular mass markers (kD) are shown on the right.

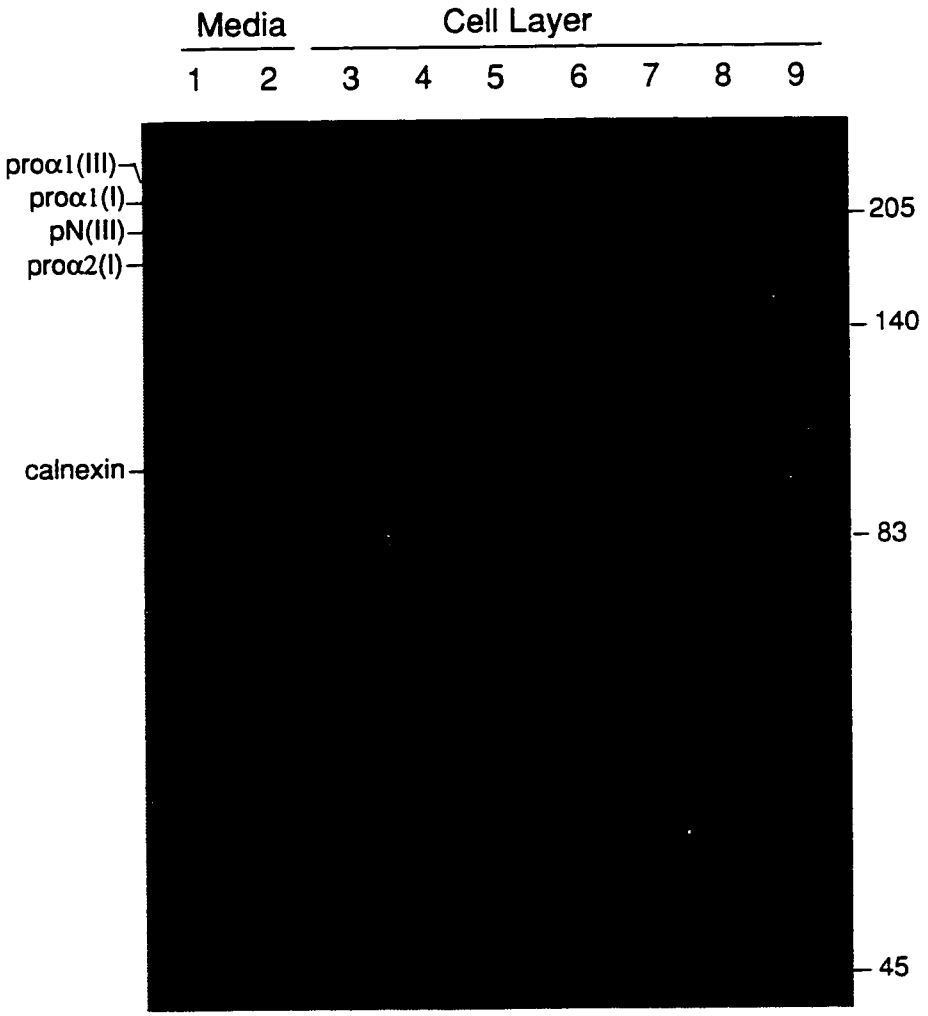

Media
Cell Layer
1
2
3
4
5
6
7
8
9
proα1(III)
proα1(I)
pN(III)
proα2(I)
calnexin
205
140
83
45

Figure 4.8. **[^{3}H]proline labeled samples show no evidence of an association between calnexin and type III procollagen.** Patient and control cells were labeled with [^{3}H]proline or TranS^{35}Label™ (with 10% complete DMEM) overnight in the presence of ascorbate. Cells were lysed with NP-40 and the resulting lysates immunoprecipitated with anti-calnexin antibodies. The [^{35}S]-labeled control in lane 3 clearly shows the successful precipitation of calnexin as well as a protein migrating in the position of type III procollagen. However, this protein is not seen in the [^{3}H]proline-labeled control (lane 4) or in either the cell line with mutation G757C (lane 5) or an as yet unidentified mutation from an EDS IV patient with a highly overmodified type III procollagen (lane 6). Migration of procollagen was corroborated by comparision with [^{3}H]proline-labeled media that was immunoprecipitated with anti-type I or type III procollagen antibodies (lanes 1 and 2, respectively). Lane 7 contains control lysate incubated with protein G-sepharose beads alone. The positions of molecular mass markers (kD) are shown on the right.

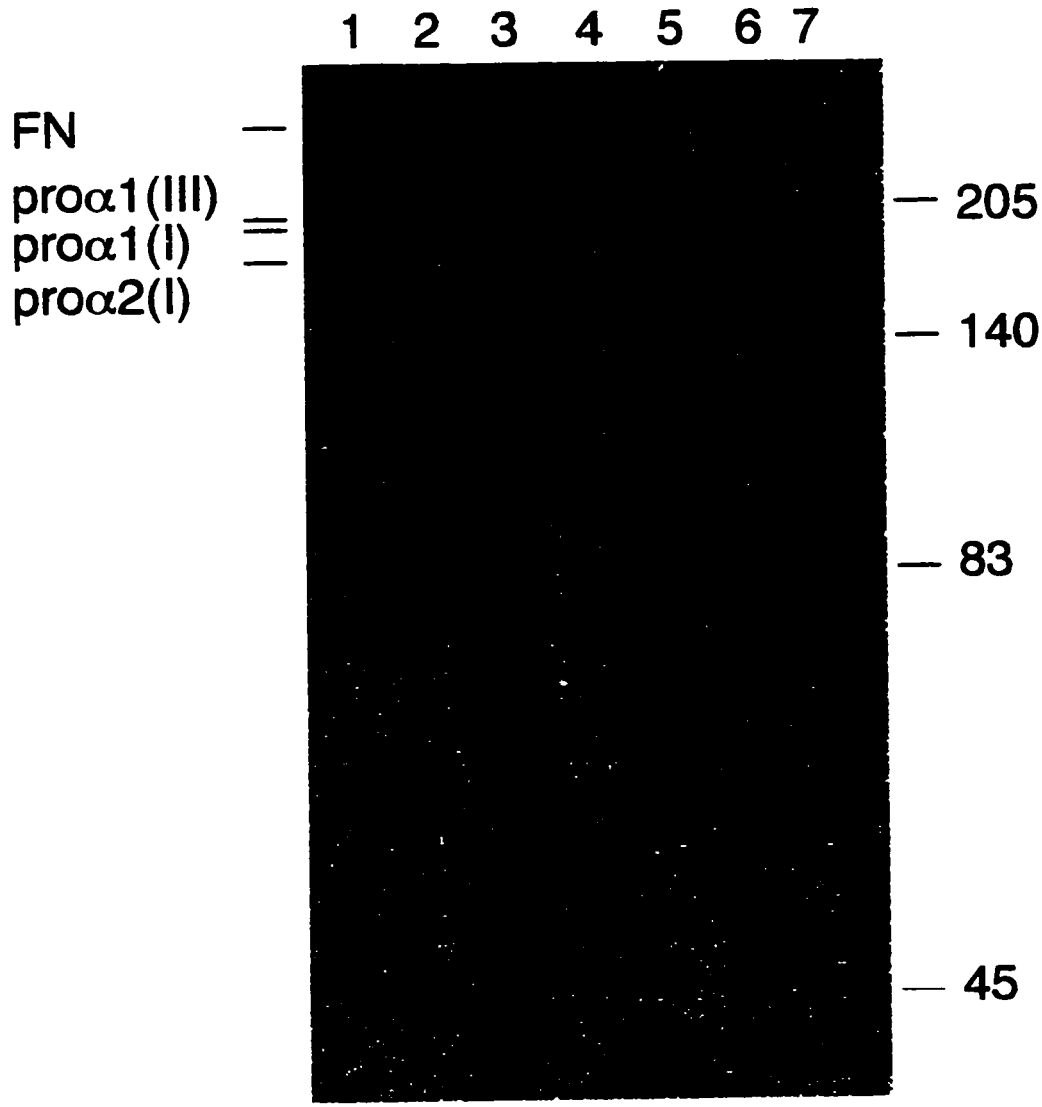
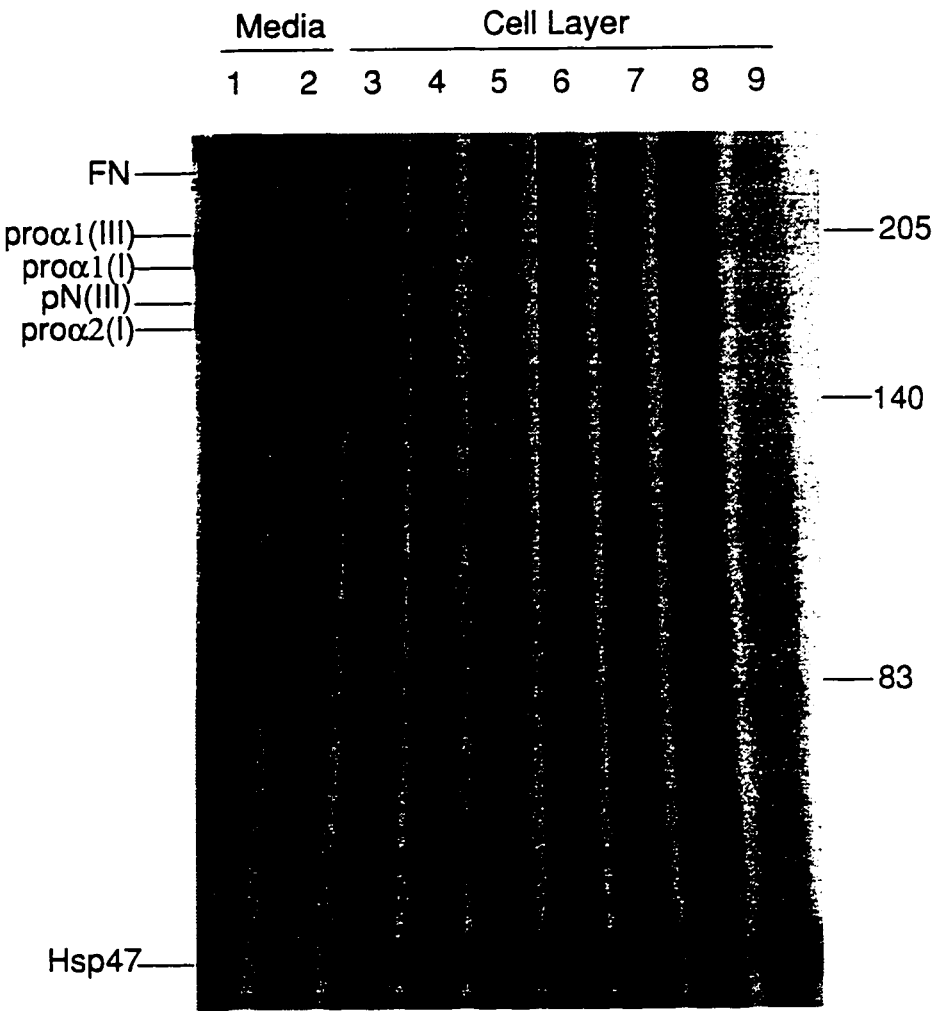

1 2 3 4 5 6 7
FN
proα1(III)
proα1(I)
proα2(I)
205
140
83
45

Figure 4.9. **SDS-PAGE analysis of anti-hsp47 immunoprecipitations.** Dermal fibroblasts from patients and a control were labeled with TranS^{35}Label™ and 10% complete DMEM and ascorbate overnight. The cells were lysed, the resulting sample immunoprecipitated with antibody against hsp47 and the proteins analyzed by SDS-PAGE. Lane 3 contains a control cell line, lanes 4-7 contain mutations skipped exons 34-48, G385E, G1012E, G904V, respectively. Lane 8 contains an irrelevant antibody control, lane 9 control lysate incubated with protein G-sepharose beads alone.

[^{3} H]proline-labeled control medium was immunoprecipitated with anti-type I or anti-type III procollagen antibodies and appear in lanes 1 and 2, respectively. A diffuse band in seen in the approximate position of type III procollagen in the control cells. This band is also present in the cell line with mutation G385E and in the irrelevant antibody control. There are several other proteins migrating in the region but none are specific to immunoprecipitation with hsp47 antibodies. The positions of molecular mass markers (kD) are shown on the right.


Media
Cell Layer
1 2 3 4 5 6 7 8 9
FN
proα1(III)
proα1(I)
pN(III)
proα2(I)
Hsp47
205
140
83

CHAPTER V

DERMAL COLLAGEN FIBRILS OF EDS IV PATIENTS ARE ALTERED DUE TO COL3A1 MUTATIONS

EXPERIMENTAL QUESTION

The clinical manifestations of Ehlers-Danlos syndrome type IV are heterogeneous including a variety of skin manifestations. Because type III collagen associates with type I collagen in dermal fibrils, this chapter focuses on collagen fibrillogenesis, asking the question: Is fibrillogenesis altered in cells of EDS IV individuals? Does position of mutation influence fibril characteristics?

INTRODUCTION

Ehlers-Danlos syndrome type IV results from mutation of COL3A1, the gene that encodes type III procollagen. To date, all mutations characterized affect the triple helical domain of this fibrillar collagen, altering its structural integrity. Mutations include those that affect splice junctions resulting in exon-skipping or other splicing anomolies, point mutations that substitute obligate glycines with residues possessing larger side chains, and genomic insertions and deletions. Derangement of type III collagen results in a range of clinical manifestations including thin skin with the

underlying veins often visible, skin that bruises and scars easily, rupture of large hollow organs such as bowel, arteries, and pregnant uterus, acrogeria, and a gaunt appearing face with large eyes.

Tissues and cells from EDS IV-affected individuals have been examined to understand how perturbations of type III collagen bring about the clinical manifestations observed in these patients. Examination of collagen fibrils and cell structures from patients vary. Dilated endoplasmic reticulum containing abnormal type III procollagen and small fibrils has been reported in some EDS IV individuals (Laurent and Agache, 1974; Holbrook and Byers, 1981; Byers et al., 1979) while other studies report patients with small (Hausser and Anton-Lamprecht, 1994) or varied diameter fibrils (Byers et al., 1979) associated with normal cellular morphology. To determine whether the observed differences result from the type or location of COL3A1 mutation, nine patient skin biopsies and nine cultured fibroblast EDS IV cell lines were examined. The biopsies showed a pattern of small fibril size associated with mutations in the latter third of the triple helix, while more amino terminal mutations were identified with fibrils of varying diameters. Exon-skipping and point mutations displayed the same pattern of fibril diameter. Enlargement of the RER was most often seen in the carboxyl-terminal mutations. Cultured fibroblasts from EDS IV patients were examined only for cell morphology; no clear cut patterns were seen with these samples.

RESULTS

Fibroblast morphology is changed in EDS IV

The fibroblasts of EDS IV patient biopsies had endoplasmic reticulum that was dilated to various degrees (figure 5.1). Those with mutation closest to the carboxyl-terminal end of the triple helix had the greatest dilatation with the distention lessening moving toward the amino terminus. Both skipped exons and replacement of glycines produced similar effects on ER architecture as indicated in figure 5.1 which includes biopsies from patients whose mutations were characterized by other members of the Byers lab. This pattern corresponds with that seen when collagenous proteins are examined by SDS-PAGE; patients with mutations that alter sequences near the carboxyl-terminal end of the molecule have procollagens that are the most overmodified and retained within the cell to the greatest degree (see chapter II).

Other cells examined did not display the varied ER dilatation of fibroblasts. Mast cells, macrophages, endothelial cells and smooth muscle cells all had endoplasmic reticulum comparable to that of control cells. Many patient biopsies displayed a larger number of lamellae around dermal blood vessels, especially post-capillary venules than was seen in control cells.

The effect of COL3A1 mutation on cellular architecture was also examined using cultured fibroblasts. Because fibroblasts were available from all patients, these cells provided a means to systematically assess the effect of mutation (exon skipping, point mutation) or substituting residue along the length of the triple helix. Fibroblasts

from patients with skipped exons 6, 14, 18, and 38 were compared. The cell lines exhibited small variations in ER distension but in all cases it was not greatly different from that seen in control cells (data not shown). I also examined a glutamic acid for glycine substitution at positions 061, 385, and 1021 (patient 94-202). The cell line harboring the G385E mutation showed moderate ER dilatation but neither the more amino-terminal G061E nor the more carboxyl-terminal G1021E substitutions produced a phenotype different from control cells (data not shown). Three cell lines with mutations that interfered with nucleation of the triple helix were also examined. In cell strains with mutations G1021R and G1021E (a different patient than the abovementioned individual with a glutamic acid for glycine substition -patient 91-579), marked distension of the ER was observed while this organelle resembled that of control cells in cells with mutation G1018V (data not shown).

Collagen fibril diameter reflects the position of the mutation in the triple helix

Collagen fibrils in the reticular dermis of EDS IV individuals were more heterogeneous in diameter, ranging from 67 ± 9 to 121 ± 18 than fibrils in controls, which ranged from 82 ± 8 to 103 ± 8nm (table 5.1). Mutations that affected the last third of the triple helix produced fibrils that were either in the small end of the normal range or smaller than control fibrils whether the mutation was a substitution of an obligate glycine or an exon-skipping mutation (figure 5.2). Mutations affecting the rest of the helix were associated with fibrils varying widely in diameter. Mutation resulting in the skipping of exons 14, 18, and 24 produced fibrils that were larger than

control fibrils with a wide standard deviation. The shape of fibrils manifest in cross section deviated from the normal rounded shape and there was a subtle divergence of shape in EDS IV collagen fibrils from normal. Fibrils were not examined from the cells in culture as these cells do not assemble them in abundance and cutting fibrils cleanly in culture dishes is difficult, limiting meaningful assessment.

Dermal thickness and collagen fiber density is decreased in EDS IV

Dermal thickness in EDS IV skin ranged from 0.66 to 1.54 mm compared to the thicker skin of controls which had a mean of 1.25 ± 0.25 mm. In addition, the number of elastic fibers present relative to collagen fibers was greater in EDS IV skin than in controls. The EDS IV collagen fibers were less organized and finer than controls as shown in figure 5.3.

DISCUSSION

EDS IV is a result of different mutations in COL3A1 and is biochemically, and clinically heterogeneous. The structural studies presented here suggest that type and location of the disease-causing mutation in COL3A1 influences cellular morphology as well as the profile of the collagen fiber and fibril. Mutations near the carboxyl-terminal end of the triple helix, either single base changes that substitute obligate glycines with amino acids possessing larger side chains or changes that result in a skipped exon, produce fibrils that are much smaller than normal. Cellularly, these mutations

correspond with abnormal retention of type III procollagen and ER distension. Mutations closer to the amino-terminal end of the triple helix correlate with cells displaying less evidence of abnormal protein retention and fibrils that vary widely in size.

Collagen fibrillogenesis and fibril assembly are complex processes governed by the collagen types involved, procollagen processing, and interactions with cellular and matrix proteins (Birk and Trelstad, 1986). Fibrils generated in cell-free systems vary in diameter according to the type of collagen employed; homotypic type I fibrils being generally thicker than type III (Birk and Trelstad, 1986). When type III collagen was added to type I collagen *in vitro* the diameter of the assembled fibril decreased with increasing concentrations of type III collagen (Lapiere, 1975). The same effect has been observed using procollagens in a fibril-forming system that included C- and N-procollagen proteinases; fibril diameter decreased from an average of 180 nm for those composed solely of type I to 80 nm when the molar ratio of type III procollagen to type I procollagen was 2:1 (Romanic et al., 1991). Another fibril-forming collagen, type V, has been shown to decrease fibril diameter when combined with type I collagen in a cell-free system. Again, heterotypic fibril diameter decreased with increasing amounts of type V collagen while fibrils assembled solely from type I collagen were the largest (Birk, et al., 1990).

Dermal collagen fibril composition and diameter has been examined *in vivo* by immunoelectron microscopy. Antibodies against the amino propeptide of type III procollagen [pNα(III)] and those directed against the amino propeptide of type I procollagen [pNα(I)] label the 20-30 nm fibrils of fetal skin. Fibrils of 25-45 nm in adult skin react with the [pNα(I)] antibodies but larger fibrils of 90-100 nm do not. This suggests that the amino propeptide of type I procollagen must be removed for the formation of a larger fibril which is consistent with the thin ribbon-like fibrils observed in dermatosparaxis (Smith et al., 1992) and failure to remove the N-propeptide *in vitro* (Holmes et al., 1993). These larger fibrils do react with the anti-type III procollagen and collagen antibodies. Exclusive surface labeling of fibrils in skin, amnion, aorta, Achilles tendon, and cornea by anti-type III collagen has been seen with antibodies directed against an epitope within the triple helix even when 8M urea was used to disrupt the fibril. Antibodies against type I collagen revealed that it was present throughout the fibril (Holmes et al., 1993). The labeling proportions matched those predicted biochemically with type I collagen representing 85-90% and type III collagen 10-15% of the fibril (Fleischmajer et al., 1990). The observed immunolabeling suggests that type III procollagen is added to the outside of a fibril.

It is clear that there is wide variability in the diameter of collagen fibrils in the EDS IV individuals presented in this study. If the role of pN-III collagen is to participate in fibrillogenesis by adding girth to the surface of mature fibrils, smaller fibrils might be the expected outcome when little to no type III procollagen is secreted

as is the case with the EDS IV individuals with mutations closest to the carboxyl-terminus. The variability seen in fibril diameter from individuals with mutations more amino terminal is most pronounced in exon-skipping mutations. These mutations produce 0-3 abnormal chains. Those trimers with one or two abnormal chains are generally retained while those composed solely of normal or shortened chains are secreted. It may be that the shortened molecules interact differently with type I collagen or any of the other components involved in fibrillogenesis.

MATERIALS AND METHODS

Tissue Samples

Skin biopsies of the upper inner arm were obtained with appropriate consent from individuals with EDS IV and from unaffected control subjects.

Cell strains and tissue culture

Patient and control dermal fibroblast cultures were established with appropriate consent and maintained as previously described (Bonadio et al., 1985). For electron microscopic examination, cells were plated at a density of 250,000 in 35mm dishes and incubated at 37° C in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum and 50 μg/ml ascorbate. The medium was replaced daily with fresh, serum and ascorbate-containing DMEM for 10 days at which time the cells had formed a thick monolayer.

Light and electron microscopy of skin biopsies

Skin biopsies were either fixed at the time of biopsy or frozen in Dulbecco's modified Eagle's medium with 7.5% dimethyl sulfoxide (DMSO)and stored for from one to several years in liquied nitrogen prior to thawing and fixing. The biopsies were fixed in 2.5% glutaraldehyde and 2% paraformaldehyde in sodium cacodylate buffer, post-fixed in osmium tetraoxid, en bloc stained in saturated uranyl acetate, dehydrated in graded ethanol and propylene oxide, processed into Epon, sectioned, sequentially stained with phosphotungstic acid, uranyl acetate and lead citrate, and viewed using a Philips 420 STEM electron microscope at 60KV (Smith et al., 1992). A carbon grating replica with 2,160 line/mm (Ernest F. Fullam, Inc.) was used to calibrate the magnification of the electron micrographs. For light microscopy, 1μm thick sections were stained with Richardson's blue and observed using a Zeiss photomicroscope. For scanning electron microscopy fixed specimens were critical point dried, mounted using carbon paste, coated with gold-palladium and viewed using a JEOL JM6300F field emission scanning electron microscope.

Samples that had been frozen and stored in liquid nitrogen prior to fixation were compared to samples from the same individual that were fixed at the time of biopsy to control for artifacts due to freezing and thawing. For two individuals with EDS IV there was no significant difference in fibril diameter when frozen and rapidly processed samples were compared. Qualitative evaluation by light microscopy and

scanning electron microscopy also did not suggest fiber disruption related to freezing and thawing. The handling of some tissue specimens prior to fixation and the fixative varied because biopsies were received from physicians at a number of institutions.

Skin thickness was measured at standard magnifications using a reticule and under light microscopy using the lowest level of the stratum corneum and the sub-dermal fat layer to define the margins.

Electron microscpy of fibroblast cultures

Duplicate dishes of cultured fibroblasts were washed several times in phosphate-buffered saline and fixed *in situ* in half-strength Karnovsky's fixative for a minimum of 24 hours at 4° - 8° C. They were washed in 0.1M sodium cacodylate buffer and post fixed for one hour at room temperature in 1% aqueous osmium tetroxide. Cultures were then washed with water and stained en bloc with 0.1% uranyl acetate. The dishes were dehydrated with a series of graded ethanol from 35-100% and embedded into epoxy resin using ethanol: epoxy resin ratios of 3:1 for 8 hours, 2:1 for 12 hours, 1:1 for 8 hours, 100% resin for 12 hours and then hardening at 65°C for 12-24 hours. The embedded cells were then removed from the dishes, cut and mounted for sectioning. En face and transverse sections were collected onto copper grids and stained sequentially with 1% phosphotungstic acid, saturated uranyl acetate, and lead citrate. The sections were examined and photographed using a Philips 420 STEM operated at 60kV.

Image Analysis

Electron micrographs of cross sections of collagen fibrils in the reticular dermis were taken at 30,000X magnification. Micrographs were digitized using an Epson 1200 scanner and Adobe Photoshop with Scantastics software. The digitized images were imported into NIH Image version 1.52, modified by threshholding, binarized, and the particles were analyzed to measure fibril area. Between 400-900 diameters were measured for each specimen. These data were exported to Excel and the fibril diameters, calculated as $2\sqrt{(area/\pi)}$, mean fibril diameter and standard deviation (SD) were calculated and graphs were plotted. The standard error of the mean (SEM) was <1% indicating that the sample size was adequate and that the error in estimating the mean diameter was less than the error due to measuring each fibril diameter.

Table 5.1 **Skin Thickness, Fibril Diameter and RER Dilatation in EDS IV**

Patient	Age (years)	Sex	COL3A1 Mutation	Skin (mm)[a]	Fibril Diameter (mm)	n[b]	RER[c]
84-153	31	M	G373R	1.08 ± 0.08	120 ± 12[d]	654	1
82-095	24	F	G769R	0.88 ± 0.08	84 ± 11	622	2
83-100	4	F	G1012R	0.84 ± 0.06	78 ± 13	868	3
78-008	26	F	G1018V	0.76 ± 0.25	71 ± 10[e]	409	3
"	30	F	G1018V	0.66 ± 0.15	67 ± 09[d]	695	3
84-174	34	M	exon 14	0.84 ± 0.10	114 ± 14[f]	669	1
93-200	18	F	exon 18	0.84 ± 0.06	121 ± 18[d]	543	2
"	18	F	exon 18	0.87 ± 0.05	119 ± 16[d]	540	2
87-253	2	M	exon 24	0.76 ± 0.11	99 ± 17	753	1
78-042	24	M	exon 34	0.92 ± 0.25	90 ± 20	540	2

[a]Skin thickness (mean ± SD) was calculated from measurements of five sections. At the 95% confidence level, none of the EDS IV values is significantly different from the mean control value of 1.25 ± 0.25 mm.
[b]Number of fibril diameters measured.
[c]RER dilation in fibroblasts: 1, normal; 2, slightly increased; 3, extreme dilation.
[d]Significantly different from controls; $p < 0.001$.
[e]Significantly different from controls; $p < 0.01$
[f]Significantly different from controls; $p < 0.0027$

Figure 5.1. **Dermal fibroblasts from EDS IV patients have dilated RER.** Electron micrographs of dermal fibroblasts from skin biopsies show the typical size of the endoplasmic reticulum in control cells (a) (→). EDS IV cells with mutation G034R have little dilation (b) while moderately enlarged ER is seen with an exon 18 skipping mutation (c) (→), and great dilation with point mutation G1018V (d) (→), skipped exon 45 (e) (→), and multi-exon deletion 34-38 (f) (→).

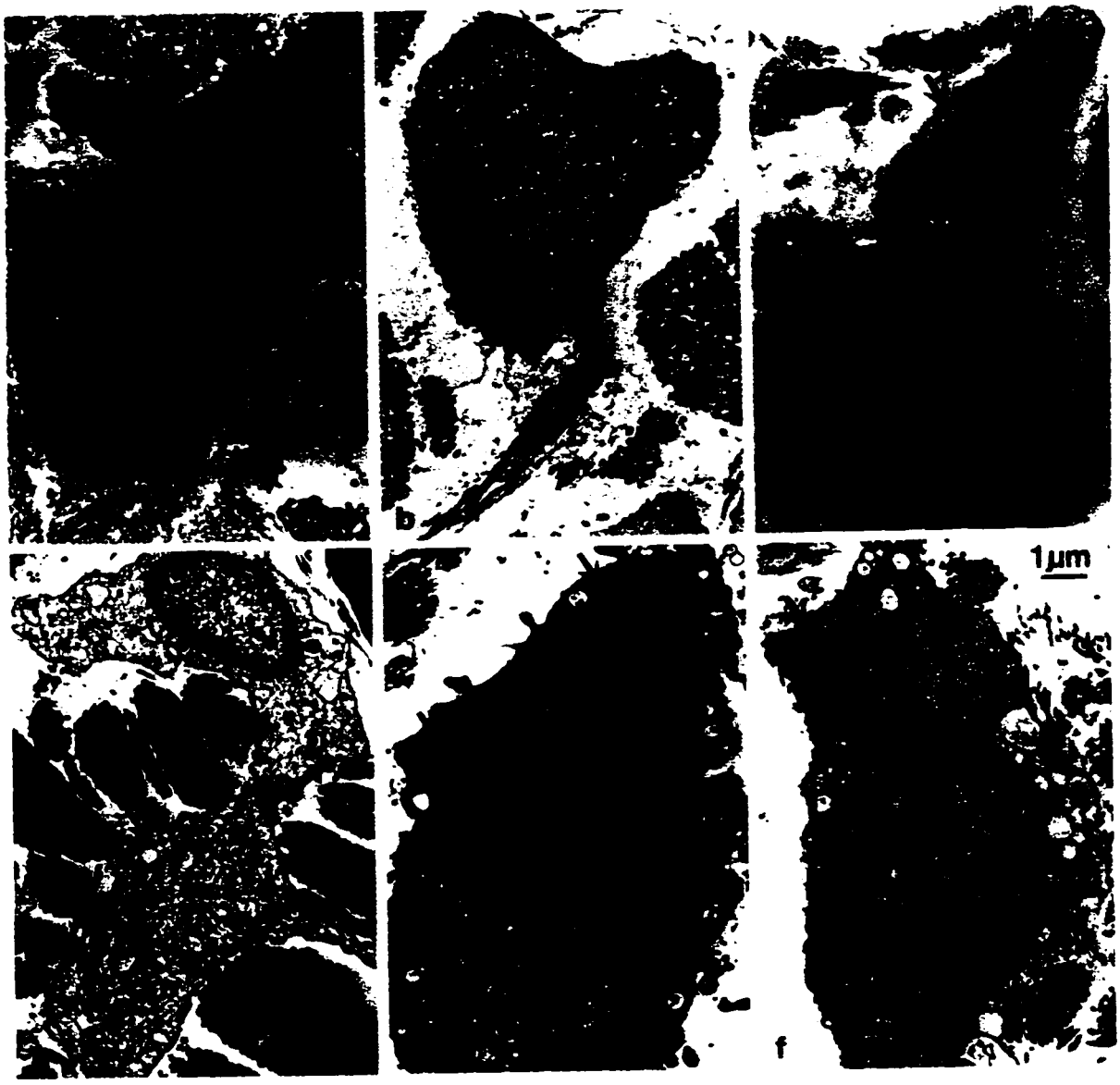

b
1μm
f

Figure 5.2. **Collagen fibrils from EDS IV vary in size and packing.** Transmission electron micrographs of collagen fibrils show that control (a) fibrils are regular in size, shape and packing while EDS IV patients with mutations near the amino-terminal end of the triple helix whether a point mutation, G034R (b) or skipped exon, 7, (d) have fibrils of variable sizes that are larger than controls. Skipping of exon 27 and deletion of exons 34-48 also result in fibrils of larger and more varied size (e, f). Point mutation G1018V near the carboxyl-terminal end of the triple helix produces fibrils of smaller than control size (c).

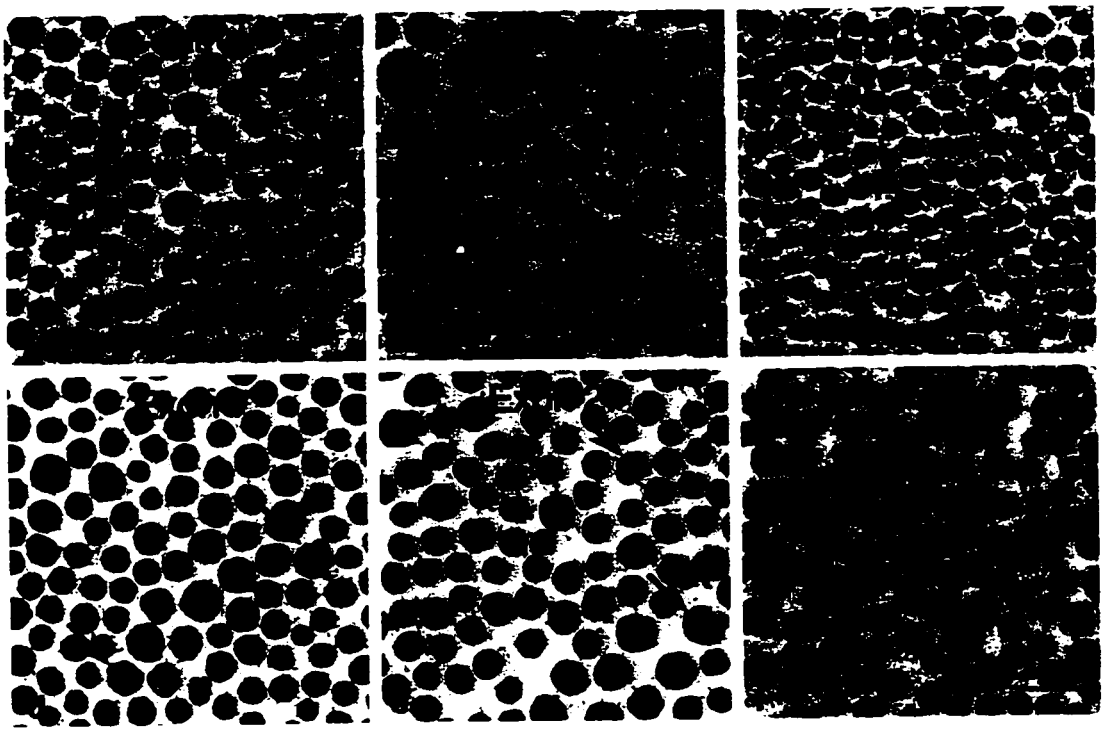
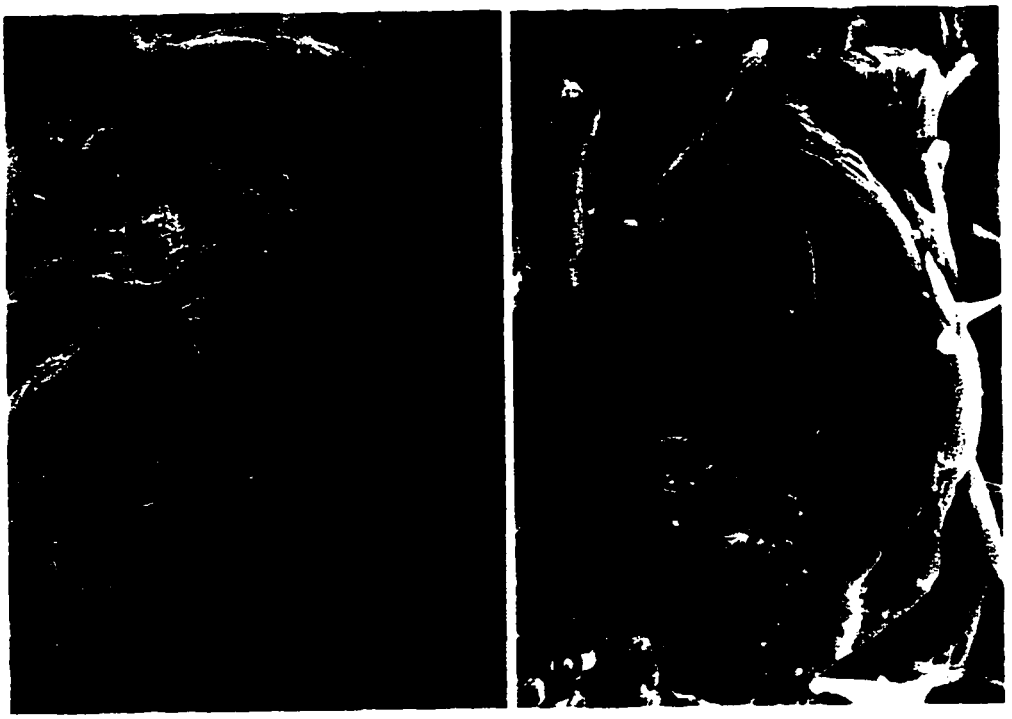


Figure 5.3. **Collagen fibers are more loosely organized from EDS IV patients.** Scanning electron microscopy shows that collagen fibers in the reticular dermis are more loosely organized than control (a) in an EDS IV patient with point mutation G373R (b).



CHAPTER VI

DISCUSSION

This dissertation describes the characterization of novel COL3A1 mutations resulting in Ehlers-Danlos syndrome type IV, the identification of a subset of carboxyl-terminal mutations that abrogate nucleation of the triple helix, the effect on fibril formation of mutation and investigations into the mechanism by which abnormal type III collagen is retained within patient cells. The mutations show position effects at the molecular level but the clinical complications of EDS IV do not appear to correlate with type or location of mutation.

EDS IV-causing mutations

I have identified 15 novel EDS IV-causing mutations all of which occur in triple-helical coding region of COL3A1. Eight point mutations substitute glycine whose small sidechain is required for accurate folding with amino acids having larger side chains. Single base changes result in glycine substitutions G373R, G769R, G829E, G904V, G1012R, G1015E, G1018V, and G1021E. Six exon skipping mutations are described at the cDNA level and one skipped exon resulting from a small genomic deletion. The cDNA-defined exon-skipping mutations remove exons 9, 14, 18, 24, 34, and 38. A 57 base pair genomic deletion where 13 nucleotides are missing from intron 21 and 44 from exon 22 result in skipping of exon 22 in cell strain 91-336. The concentration of point mutations toward the carboxyl-terminal end of the triple

helix and exon skipping mutations in the amino-terminal area reflects my mutation detection strategy not the distribution of mutations in nature. To date, EDS-IV-causing mutations have only been found in the triple helical portion of COL3A1, according to published reports and unpublished findings in our lab. This distribution may simply be a consequence of the low number of mutations detected; there are approximately twice as many characterized mutations of the genes encoding type I collagen and only five reports of mutations outside of the triple helix causing osteogenesis imperfecta (reviewed in Kuvivaniemi et al., 1997). The division of mutations types, exon skipping versus obligate glycine substitutions is quite different in the two diseases. In OI, reported glycine replacements represent 67% of total mutations in α1(I) and 61% in α2(I), while exon-skipping mutations make up less than 20% in α1(I), and approximately 30% in α2(I). EDS IV-causing mutations are almost equally divided between these two types of mutations (reviewed in Kuvivaniemi et al., 1997, unpublished results from the Byers lab).

Folding defects associated with EDS IV

Mutations close to the carboxyl-terminal end of the triple helix, G1012R, G1015E, G1018V, G1018D, G1021R and G1021E, contained cysteines at positions 1028 and 1029 required for disulfide bond at the end of the the triple helix but failed to form trimers. Experiments following the course of helix formation at 4°C, 20°C, and 37°C demonstrated that a protease-resistant helix does not form at any temperature in cells harboring a glutamic acid for glycine substitution at residue 1021 but helical

molecules are evident in cells harboring more amino-terminal glutamic acid substitutions (G385E and G061E). These data suggest that mutations near the carboxyl-terminal end of the helix destabilize a sequence required for nucleation of triple helical folding and the only molecules produced by these cells are composed entirely of α chains encoded by the normal allele. Previous studies on the cell strain with the G1018V mutation indicate that type III collagen melting temperature corresponds to that of normal cells and a reduction in temperature to 30°C, which often induces cells to secrete abnormal molecules, did not impel this cell line to do so (Superti-Furga, et al., 1989). Taken together, these findings support the notion that only one site can be used for nucleation of the triple helix even though other microdomains of high sequence stability are present elsewhere in the helix domain.

Collagen fibrils are altered in EDS IV patients

EDS IV-induced folding abnormalities affect collagen at the fibril level as well as the molecular level. Electron microscopic examination of dermal collagen fibrils from EDS IV patients characterized by myself (G373R, G769R, G1012R, G1018V and skipped exons 14, 18, 24, and 34) along with mutations G016C, G034R, G061E, G1021E and skipped exons 7, 8, 24 (a second patient), 27, 45, and 34-48 characterized by others in the lab indicate that both position and type of mutation affect fibrillogenesis. Both point mutations and exon skipping mutations near the amino-terminal end of the triple helix are larger overall than control fibrils and exhibit more heterogeneity. Exon-skipping mutations affecting the second third of the triple helix

also produce larger fibrils with a larger standard deviation than controls while fibrils from cells harboring point mutations are slightly smaller although they, too, are diverse in size. Both classes of mutations produce more uniform fibrils when the mutations are close to the carboxyl-terminal end of the helix; exon skipping mutations result in fibrils close in diameter to controls while those point mutations that interfere with helix initiation generate fibrils that are much smaller than controls. Small fibrils can be generated *in vitro* by increasing the amount of type III collagen relative to type I. While abnormal type III molecules might result in generalized secretion problems, it seems unlikely that the ratio of type I:type III collagen could be altered enough to account for my data. In fact, previous studies on the EDS IV individual with the G1018V mutation indicate that this patient secretes normal amounts of type I collagen (Byers et al., 1981). Another possibility is that type III collagen adds girth to fibrils whose interior is composed of type I collagen. Anti-type III collagen antibodies interact with molecules on the surface of fibrils in skin, and other tissues (Keene et al., 1987) even when fibrils are disrupted with chaotropic agents while type I collagen has been immunolocalized throughout the entire fibril (Holmes et al., 1993).

The heterogeneity of fibrils from patients with exon-skipping mutations is intriguing. Intutitively, it seems that molecules with 0, 1, 2 or 3 abnormal chains would interact differently with the components of fibrillogenesis and produce diverse fibrils. However, in one EDS IV individual with an exon 17-skipping mutation, mutant homotrimers accumulated in the medium in which the cells were cultured but these

molecules were not incorporated into the extracellular matrix (Chiodo et al., 1994). There is another report in the literature of a reduction in the amount of type III collagen in the patient's dermis to 11% of normal which is close to the 12.5% that would be observed if only normal homotrimers were incorporated (Cole et al., 1990). These findings suggest that only normal homotrimers associate into fibrils. However, if the absolute amount of type III collagen available for fibrillogenesis determined ultimate fibril morphology then these fibrils as well as those from the above mentioned point mutations would be smaller than normal.

Association of malfolded type III procollagen with ER resident proteins.

Type III procollagen contributes to the ER dilatation seen in the fibroblasts of some EDS IV patients (Byers, et al., 1981). I had hoped to determine which, if any, ER-resident proteins were involved in this intracellular retention. While there are many candidate proteins, my investigation was limited to those proteins for which antibodies were accessible to me. Unfortunately, I was not able to see any associations with the antibodies I used. I am convinced that calnexin does not play a role in retaining type III procollagen but I am not convinced that PDI is not involved. Finding more reliable antibodies would improve these studies as would changing SDS-PAGE systems to one capable of delivering sharp bands with more separation than min-gels.

My aim in undertaking the experiments detailed in this dissertation was to uncover new mutations resulting in EDS IV and make progress at understanding the disease at the molecular level so that it might be better understood in terms of genotype

and phenotype correlation. The carboxyl-terminal point mutations I characterized provide insight into how type III collagen folds. The characterization of other mutations in this area should lead to a definition of the first folding block in type III procollagen and whether nucleation is affected equally by local perturbations caused by replacement of obligate glycines and changes in helix registration caused by exon-skipping mutations. Additional studies are needed to answer questions as to the variability in fibril diameter seen in amino-terminal exon-skipping mutations which might also reflect differences in local perturbations versus registration changes.

END NOTE

Portions of Chapter II in this manuscript were adapted from the paper "Splicing defects in the COL3A1 gene: marked preference for 5' (donor) splice-site mutations in patients with exon-skipping mutations and Ehlers-Danlos syndrome type IV," published in the American Journal of Human Genetics 61, 1276-1286 by Ulrike Schwarze, Jayne A. Goldstein and Peter H. Byers, 1997.

Chapter V of this manuscript was adapted from the paper "Mutations in the COL3A1 gene result in the Ehlers-Danlos syndrome type IV and alterations in the size and distribution of the major collagen firbrils of the dermis" published in the Journal of Investigative Dermatology 108(3), 241-247 by Lynne T. Smith, Ulrike Schwarze, Jayne A. Goldstein, and Peter H. Byers, 1997.

Chapter III along with a portion of Chapter II has been submitted for publication.

REFERENCES

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. (1994) Molecular Biology of the Cell. 3rd ed. (New York: Garland Publishing) pp. 616-617.

Anderson, D.W., Thakker-Varia, S., Tromp, G., Kuivaniemi, H. and Stolle, C.A. (1997) A glycine (415)-to-serine substitution results in impaired secretion and decreased thermal stability of type III procollagen in a patient with Ehlers-Danlos Syndrome type IV. Hum. Mut. 9(1), 62-63.

Barsh, G.S, and Byers, P.H. (1981) Redued secretion of structurally abnormal type I procollagen in a form of osteogenesis imperfecta. Proc. Natl. Acad. Sci. USA 78(8), 5142-5146.

Bateman, J.F., Moeller , Hannagan, M., Chan, D., and Cole, W.G. (1992) Characterization of three osteogenesis imperfecta collagen α1(I) glycine to serine mutations demonstrating a position-dependent gradient of phenotypic severity. Biochem. J. 288, 131-135.

Bateman, J.F., Chan, D., Walker, I.D., Rogers, J.G., and Cole, W. G. (1987) Lethal perinatal osteogenesis imperfecta due to the substitution of arginine for glycine at residue 391 of the α1(I) chain of type I collagen. J. Biol. Chem. 262(15), 7021-7028.

Bateman, J.F., Mascara, T., Chan, D., and Cole, W. G. (1984) Abnormal type I collagen metabolism by cultured fibroblasts in lethal perinatal osteogenesis imperfecta. Biochem. J. 217, 103-115.

Beighton, P. The Ehlers-Danlos Syndromes. In McKusick's Heritable Disorders of Connective Tissue, Fifth Ed. (St. Louis: Mosby) pp. 189-251.

Bergeron, J.J.M., Brenner, M.B., Thomas, D.Y., and Williams, D.B. (1994) Calnexin: A Membrane-Bound Chaperone of the Endoplasmic Reticulum. TIBS 19, 124-128.

Birk, D.E., and Trelstad, R.L. (1986) Extracellular comparment in tendon morphogenesis: collagen fibril, bundle and macroaggregate formation. J. Cell. Biol. 103, 231-240.

Birk, D.E., Silver, F.H., and Trelstad, R.L. (1991) Matrix Assembly. In Cell Biology of Extracellular Matrix, 2nd ed. E. D. Hay, ed. (New York: Plenum Press) pp. 221-254.

Blond-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. T., Sambrook, J. F., and Gething, M. H. (1993) Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. Cell 75, 717-728.

Bonadio, J., and Byers, P.H. (1985) Subtle structural alterations in the chains of type I procollagen produce osteogenesis imperfecta type II. Nature 316, 363-366.

Bonadio, J., Holbrook, K.A., Gelinas, R.E., Jacob, J., and Byers, P.H. (1985) Altered triple helical structure of type I procollagen in lethal perinatal osteogenesis imperfecta. J. Biol. Chem. 260, 1734-1742.

Bulleid, N.J., Wilson, R., and Lees, J.F. (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, P.H., Wallis G.A., and Willing, M.C. (1991) Osteogenesis imperfecta: translation of mutation to phenotype. J. Med. Genet. 28, 433-442

Byers, P.H., Holbrook, K.A., Barsh, G.S., Smith, L.T., and Bornstein, P. (1981) Altered secretion of type III procollagen in a form of type IV Ehlers-Danlos Syndrome. Lab. Inves. 44(4), 336-341.

Byers, P.H., Holbrook, K.A., McGillivray, B., MacLeod, P.M., Lowry, R.B. (1979) Clinical and ultrastructural heterogeneity of type IV Ehlers-Danlos Syndrome. Hum. Genet. 47, 141-150.

Chessler, S.D., and Byers, P.H. (1992) Defective folding and stable association with protein disulfide isomerase/prolyl hydroxylase of type I procollagen with a deletion in the proα2(I) chain that preserves the Gly-X-Y repeat pattern. J. Biol. Chem. 267(11), 7751-7757.

Chessler, S.D., and Byers, P.H. (1993) Bip binds type I procollagen proα chains with mutations in the carboxyl-terminal propeptide synthesized by cells from patients with osteogenesis imperfecta. J. Biol. Chem. 268(24), 18226-18233.

Chomczynski, P., and Sacchi, N. (1987) Single-step method for RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156-159.

Cogan, J.D., McCarthy, E.M.S., Prince, M.A., Lekhakula, S., Futrakul, A., Bundey, S., Phillips III, J.A. (1996) A novel mechanism of aberrant splicing. Am. J. Hum. Genet. Suppl. 59, A10.

Cooper, D.N., Krawczak, M. and Antonarakis, S. E. (1993) The nature and mechanisms of human gene mutation. In Human Gene Mutation, D. N. Cooper and M. Krawczak. (Oxford, UK: Bios Scientific Publishing) pp. 259-291.

Cooper, D.N., and Krawczak, M. (1991) Mechanisms of insertional mutagenesis in human genes causing genetic disease. Hum. Genet. 87, 409.

Cooper, D.N., Youssoufian, H. (1988) The CpG dinucleotide and human genetic disease. Hum. Genet. 78, 151-159.

Copeland, C. S., Zimmer, K., Wagner, K. R., Healey, G. A., Mellman, I., and Helenius, A. (1988) Folding, trimerization, and transport are sequential events in the biogenesis of influenza virus hemagglutinin. Cell 53, 197-209.

Davis, J.M., and Bachinger, H.P. (1993) Hysteresis in the triple helix-coil transition of type III collagen. J. Biol. Chem. 268(34), 25965-25972.

Deak, S.F., Scholz, P.M., Amenta, P.S., Constantinou, C.D., Levi-Minzi, S.A., Gonzalez-Lavin, L., and Mackenzie, J.W. (1991) The substitution of arginine for glycine 85 of the α1(I) procollagen chain results in mild osteogenesis imperfecta. J. Biol. Chem. 266, 21827-21832.

Degen, E., and Williams, D. B. (1991) Participation of a novel 88-kd protein in the biogenesis of murine class I histocompatibility molecules. J. Cell Bio. 112(6), 1099-1115

Del Gatto, F., Gesnel M.-C., and Breathnach, R. (1996) The exon sequence TAGG can inhibit splicing. Nucleic Acids Res. 24, 2017-2021.

Dombrowski, K. E., Vogel, B.E., and Prockop, D. J. (1989) Mutations that alter the primary structure of type I procollagen have long-range effects on its cleavage by procollagen N-proteinase. Biochemistry 28, 7107-7012.

Engel, J., and Prockop, D. J. (1991) The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. Annu. Rev. Biophys. Biophys. Chem. 20, 137-152.

Estes, P.A., Cooke, N.E., and Liebhaber, S.A. (1992) A native RNA secondary structure controls alternative splice-site selection and generates two human growth hormone isoforms. J. Biol. Chem. 267, 14902-14908.

Fertala, A., Lieron, A. L., Ganguly, A., Li, S., Ala-Kokko, L., Anumula, K. R. and Prockop, D. J. (1994) Synthesis of recombinant human procollagen II in a stably transfected tumour cell line (HT1080). Biochem. J. 298, 31-37.

Fields, S., and Song, O. K. (1989) A novel genetic system to detect protein-protein interactions. Nature (London) 340, 245-246.

Feinberg, A. P. and Vogelstein, B. (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem 137(1), 266-267.

Fleischmajer, R., Perlish, J. S., Burgeson, R. E., Shaikh-Bahai, R., and Timpl, R. (1990) Type I and type III collagen interactions during fibrillogenesis. Ann. N. Y. Acad. Sci. 580, 161-175.

Ganguly, A., Smelt, S., Mewar, R., Fertala, A., Sieron, A. L., Overhauser, J., and Prockop, D. J. (1994) Targeted insertions of two exogenous collagen genes into both alleles of their endogenous loci in cultured human cells: the insertions are directed by relatively short fragments containing the promoters and the 5' ends of the genes. Proc. Natl. Acad. Sci. 91, 7365-7369.

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

Goldstein, J. A., Schwarze, U., Witz, A., and Byers, P.H. (1994) Marked heterogeneity in COL3A1 mutations that produce the EDS type IV phenotype. Matrix Biology 14(5), 392-393.

Haas, I. G. (1991) BiP: A heat shock protein involved in immunoglobulin chain assembly. In Current Topics in Microbiology and Immunology, vol. 167. (Berlin: Springer-Verlag) pp.71-82.

Hammond, C., Braakman, I., and Helenius A. (1994) Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. Proc. Natl. Acad. Sci. 91, 913-917.

Hausser, I., and Anton-Lamprecht, I. (1994) Differential ultrastructrual aberrations of collagen fibrils in Ehlers-Danlos syndromes type I-IV as a means of diagnostics and classification. Hum. Genet. 3, 394-407.

Holbrook, K. A., and Byers, P. H. (1981) Ehlers-Danlos syndrome type IV. Storage in the rough endoplasmic reticulum. Lab. Invest. 44, 342-350.

Holmes, D. F., Watson, R. B., Steinmann, B., and Kadler, K. E. (1993) Ehlers-Danlos syndrome type VIIB. Morphlogy of type 1 collagen fibrils formed in vivo and in vitro is determined by the conformation of the retained N-propeptide. J. Biol. Chem. 268, 15758-15765.

Johnson, P. H., Richards, A. J., Pope, F. M., and Hopkinson, D. A. (1992) A COL3A1 glycine 1006 to glutamic acid substitution in a patient with Ehlers-Danlos syndrome type IV detected by denaturing gradient gel electrophoresis. J. Inher. Metab. Dis. 15, 426-430.

Kontusaari S., Tromp, G., Kuivaniemi, H., Stolle C, Pope FM, Prockop DJ. (1992) Substitution of aspartate for glycine 1018 in the type III procollagen (COL3A1) gene causes type IV Ehlers-Danlos syndrome: the mutated allele is present in most blood leukocytes of the asymptomatic and mosaic mother. Am. J. Hum. Genet. 51, 497-507.

Kontusaari, S., Tromp, G., Kuivaniemi, H., Romanic, A.M., and Prockop, D.J. (1990) A mutation in the gene for type III procollagen (COL3A1) in a family with aortic aneurysms. J. Clin. Invest. 86(5), 1465-1473.

Kuhn, K. (1987) The classical collagens: types I, II, and III. In Structure and Function of Collagen Types. R. Mayne and R. E. Burgeson, eds. (Orlando, FL: Academic Press) pp. 1-37.

Kuivaniemi, H., Tromp, G., and Prockop, D. J. (1991) Mutations in collagen genes: Causes of rare and some common diseases in humans. FASEB-J. 5, 2052-2060.

Kuivaniemi, H., Kontusaari, S., Tromp, G., Zhao M., Sabol, C. and Prockop, D. J. (1990) Identical G^{+1} to A mutations in three different introns of the type III procollagen gene (COL3A1) produce different patterns of RNA splicing in three variants of Ehlers-Danlos syndrom IV. J. Biol. Chem. 265, 12067-12074.

Lamande, S. R., and Bateman, J. F. (1995) The type I collagen proα(I) COOH-terminal propeptide N-linked oligosaccharide. J. Biol. Chem. 270

Lamande, S. R., Chessler, S. D., Golub, S. B., Byers, P. H., Chan, D., Cole, W. G., Sillence, D. O., and Bateman, J. F. (1995) Endoplasmic reticulum-mediated quality control of type I collagen production by cells from osteogenesis imperfecta patients with mutations in the proα(I) chain carboxyl-terminal propeptide which impair subunit assembly. J. Biol. Chem. 270(15), 8642-8649.

Lapiere, C. M., Nusgens, B. and Pierard G. E. (1977) Interaction between collagen type I and type III in conditioning bundles organization. Connect. Tissue Res. 5, 21-29.

Laskey, R. A., Honda, B. M, Mills, A.D., and Finich, J. T. (1978) Nucleosomes are assembled by an acidic protein which binds histones and tranfers them to DNA. Nature 275, 416-20.

Laurent, R., and Agache, P. (1974) Is acrogeria a fibroblast disease? An ultrastrctural study. Dermatologica 148, 28-38.

Lees, J. F., and Bulleid, N. J. (1994) The role of cysteine residues in the folding and association of the COOH-terminal propeptide of types I and III procollagen. J. Biol. Chem. 269(39), 24354-24360.

Lightfoot, S. J., Holmes, D. F., Brass, A., Grant, M. E., Byers, P. H., and Kadler, K. E. (1992) Type I procollagens containing substitutions of aspartate, arginine, and cysteine for glycine in the pro alpha1(I) chain are cleaved slowly by N-proteinase, but only the cysteine substitution introduces a kink in the molecule. J. Biol. Chem. 267, 25521-25528.

Lin, H. Y., Masso-Welch, P., Di, Y. P., Cai, J. W., Shen, J. W., and Subjeck, J. R. (1993) The 170kDa glucose-regulated stress protein is an endoplasmic reticulum protein that binds immunoglobulin. Mol. Biol. Cell 4(11), 1109-19.

Linsenmayer, T. F. (1991) Collagen. In Cell Biology of Extracellular Matrix. (New York: Plenum Press) pp. 7-44.

Mackay, K., Raghunath, M., Superti-Furga, A., Steinmann, B., and Dalgleish, R. (1996) Ehlers-Danlos syndrome type IV caused by gly400glu, gly595cys and gly10003asp substitutions in collagen III: clinical features, biochemical screening, and molecular confirmation. Clin. Genet. 49, 286-295.

MacMillan, A.M., McCaw, P.S,. Crispino, J.D., and Sharp, P.A. (1997) SC35-mediated reconstitution of splicing in U2AF-depleted nuclear extract. Proc. Natl. Acad. Sci. USA 94, 133-136.

Maquat, L. E. (1996) Defects in RNA splicing and the consequence of shortened translational reading frames. Am. J. Hum. Genet. 59, 279-286.

Maquat, L. E. (1995) When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. RNA 1, 453-465.

Melnick, J., Aviel, S., and Argon, Y. (1992) The endoplasmic reticulum stress protein GRP94, in addition to BiP, associates with unassembled immunoglobulin chains. J. Biol Chem. 267(30), 21303-21306.

Minor, R. R., Sippola-Thiele, M., McKeon, J., Berger, J., and Prockop, D. J. (1986) Defects in the processing of procollagen to collagen are demonstrable in cultured fibroblasts from patients with the Ehlers-Danlos and osteogenesis imperfecta syndromes. J. Biol. Chem. 261, 1006-10014.

Munro, S., and Pelham H. R. B. (1987) A C-terminal signal prevents secretion of luminal ER proteins. Cell 48, 899-907.

Nagata, K. (1996) Hsp47: a collagen-specific molecular chaperone. TIBS 21:23-26.

Nagata, K., and Yamada, K. M. (1986) Phosphorylation and transformation sensitivity of a major collagen-binding protein of fibroblasts. J. Biol. Chem. 261(16), 7531-7536.

Nakai, A., Satoh, M., Hirayoshi, K., and Nagata K. (1992) Involvement of the stress protein hsp47 in procollagen processing in the endoplasmic reticulum. J. Cell Biol. 117(4), 903-914.

Narcisi, P., Wu, Y., Tromp, G., Earley, J. J., Richards, A. J, Pope, F. M., and 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, 278-283.

Natsume, T., Koide, T., Yokota, S., Hirayoshi, K., and Nagata, K. (1994) Interactions between collagen-binding stress protein hsp47 and collagen: analysis of kinetic parameters by surface plasmon resonance biosensor. J. Biol. Chem. 269 (49), 31224-31228.

Niemela, O., Risteli, L., Parkkinen, J., and Risteli, J. (1985) Purification and characterization of the N-terminal propeptide of human type III procollagen. Biochem. J. 232(1), 145-150.

Nuytinck, L., De Paepe, A., Renard, J.-P., Adriaens, F., and Leroy, J. (1994) Single-strand conformation polymorphism (SSCP) analysis of the COL3A1 gene detects a mutation that results in the substitution of glycine 1009 to valine and causes severe Ehlers-Danlos syndrome type IV. Hum Mut. 3, 268-274.

Olsen, B. R. (1991) Collagen Biosynthesis. In Cell Biology of Extracellular Matrix, 2nd ed. E. D. Hay ed. (New York: Plenum Press) pp. 177-220.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc. Natl. Acad. Sci. USA 86(8), 2766-2770.

Ou, W. J., Cameron, P. H., Thomas, D. Y. and Bergeron, J. J. M. (1994) Association of folding intermediates of glycoproteins with calnexin during protein maturation. Nature 364, 771-776.

Pack, M., Constantinou, C. D, Kalia, K., Nielsen, K. B., and Prockop, D.J. (1989) Substitution of serine for α1(I)-glycine 844 in a severe variant of osteogenesis imperfecta minimally destabilizes the triple helix of type I procollagen. J. Biol. Chem. 264, 19694-19699.

Piez, J.A. (1984) In Extracellular Matrix Biochemistry, K. A. Piez, and A. H. Reddi, eds. (New York: Elsevier Science Publishing Co., Inc.), pp.1-40.

Peterson, M.L., Bryman, M. B., Peiter, M., Cowan, C. (1994) Exon size affects competition between splicing and cleavage-polyadenylation in the immunoglobulin μ ene. Mol. Cell. Biol. 14, 77-86.

Pope, F. M., Narcisi, P. N., Nicholls, A. C., Germaine, D., Pals, G., and Richards, A. J. (1996) COL3A1 mutations cause variable clinical phenotypes including acrogeria and vascular rupture. Br. J. Dermatol. 135, 163-181.

Pope, F. M., Martin, G.R., Lichtenstein, J. R., Pentitinen, R., Gerson, B., Rowe, D.W., and McKusick, V. A. (1975) Patients with Ehlers-Danlos syndrome type IV lack type III collagen. Proc Natl Acad Sci USA 72, 1314-1316.

Privalov, P. L. (1982) Stability of proteins. Proteins which do not present a single cooperative system. Adv. in Prot. Chem. 35, 1-104.

Prockop, D. J. (1984) Osteogenesis imperfecta: phenotypic heterogeneity, protein suicide, short and long collagens. Am. J. Hum. Genet. 36, 499-505.

Prockop, D.J., and Kivirikko, K. I. (1984) Heritable disease of collagen. N. Eng. J. Med. 311, 276-386.

Raghunath, M., Bruckner, P. and Steinmann, B. (1994) Delayed triple helix formation of mutant collagen from patients with osteogenesis imperfecta. J. Mol. Biol. 236, 940-949.

Richards, A. J., Narcisi, P., Lloyd, J. C., Johnson, P. H., Hopkinson, D. A., and Pope, F. M. (1993) Substitution of glycines 1000, 1003 and 1006 in type III collagen all cause the acrogeric form of EDS IV, and destabilise the collagen triple helix. Matrix 13, 47.

Romanic, A.M., Adachi, E., Kadler, K. E., Hojima, Y., and Prockop, D. J. (1991) Copolymerization of pNcollagen III and collagen I: pNcollagen III decreases the rate of incorporation of collagen I into fibrils, the amount of collagen I incorporated, and the diameter of the fibrils formed. J. Biol. Chem. 266, 12703-12709.

Roth R. A., and Pierce, S.B. (1987) In vivo cross-linking of protein disulfide isomerase to immunoglobulins. Biochemistry 26(14), 4179-4182.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487-491

Sambrook, J., Fritsch, E. R., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger, F. S., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci., USA 74, 5463-5467.

Schwarze, U., Goldstein, J. A., and Byers, P. H. (1997) Splicing defects in the COL3A1 gene: marked preference for 5' (donor) splice-site mutations in patients with exon-skipping mutations and Ehlers-Danlos syndrome type IV. Am. J. Hum. Genet. 61, 1276-1286.

Singh, R., Valcarcel, J., and Green, M.R. (1995) Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science 268, 1173-1176.

Sirand-Pugnet, P., Durosay, P., Brody, E., and Marie, J. (1995) An intronic (A/U)GGG repeat enhances the splicing of an alternative intron of the chicken β-tropomyosin pre-mRNA. Nucleic Acids Res 23, 3501-3507.

Smith, L. T., Schwarze, U., Goldstein, J. A., and Byers, P. H. (1997) Mutations in the COL3A1 gene result in the Ehlers-Danlos syndrome type IV and alterations in the size and distribution of the major collagen firbrils of the dermis. J. Inv. Derm. 108(3), 241-247.

Smith, L. T., Wertelecki, W., Milstone, L. M., Petty, E. M., Seashore, M. R., Braverman, I., Jenkins, T., and Byers, P. H. (1992) Human dermatosparaxis: A form of Ehlers-Danlos syndrome that results from failure to remove the amino-terminal propeptide of type I collagen. Am. J. Hum. Genet. 51, 235-244.

Smith, T., Ferreira, L. R., Hebert, C., Norris, K., and Sauk, J. J. (1995) Hsp47 and cyclophilin B traverse the endoplasmic reticulum with procollagen into pre-golgi intermediate vesicles: A role for hsp47 and cyclophilin B in the export of procollagen from the endoplasmic reticulum. J. Biol. Chem. 270(31), 18232-328.

Sterner, D. A., and Berget, S. M. (1993) In vivo recognition of a vertebrate mini-exon as an exon-intron-exon unit. Mol.Cell. Biol. 13, 2677-2687.

Stover, M. L. Primora, D., Liu, S. C., McKinstry, M. B., and Rowe, D. W. (1993) Defective splicing of mRNA from one COL1A1 allele of type I collagen in non-deforming (type I) osteogeneis imperfecta. J. Clin. Invest. 92, 1994-2002.

Valcarcel, J. Gaur, R.K., Singh, R., and Green, M.R. (1996) Interaction of $U2AF^{65}RS$ region with pre-mRNA of branch point and promotion base pairing with U2 snRNA. Science 273, 1706-1709.

van Oers, C. C. M., Adema, G. J., Zandberg, H., Moen, T. C., and Baas, P. D. (1994) Two different sequence elements within exon 4 are necessary for calcitonin-specific splicing of the human calcitonin/calcitonin gene-related peptide I pre-mRNA. Mol. Cell. Biol. 14, 951-960.

Vogel, B.E., Doelz, R., Kadler, K.E., Hojima, Y., Engel, J., and Procklop, D.J. (1988) A substitution of cysteine for glycine 748 of the alpha1 chain produces a kink at this site in the procllagen I molecule and an altered N-proteinase cleavage site over 225 nm away. J. Biol. Chem. 263(35), 19249-19255.

Wang, J., Takagaki, Y., Manley, J. L. (1996) Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability. Genes & Dev. 10, 2588-2599.

Westerhausen, A., Kishi, J., and Prockop, D. J. (1990) Mutations that substitute serine for glycine α1-598 and glycine α1-631 in type I procollagen. J. Biol. Chem. 265(23), 13995-14000.

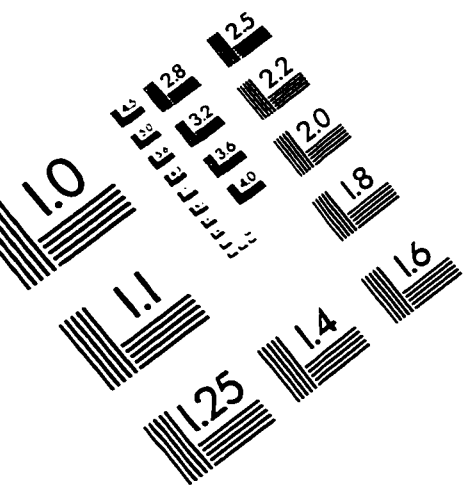
Williams, C. J., and Prockop, D. J. (1983) Synthesis and processing of a type I procollagen containing shortened pro-alpha 1(I) chains by fibroblasts from a patient with osteogenesis imperfecta. J. Biol. Chem. 258, 5915-5921.

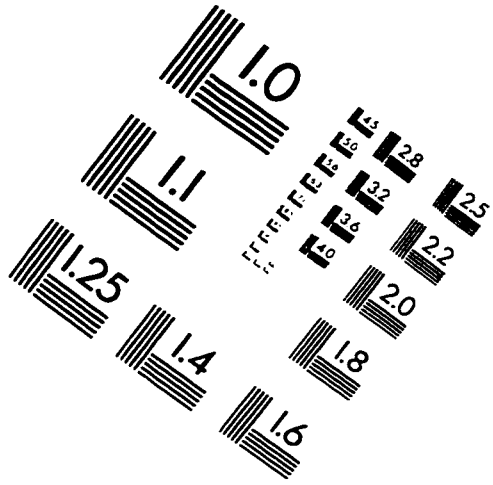
Willing, M. C., Cohn, D. H., Starman, B., Holbrook, K. A., Greenberg, C. R., and Byers, P. H. (1988) Heterozygosity for a large deletion in the α2(I) collagen gene has a dramatic effect on type I collagen secretion and produces perinatal lethal osteogenesis imperfecta. J. Biol. Chem. 263(17), 8398-8404.

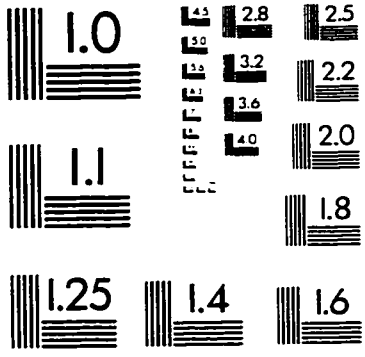
VITA

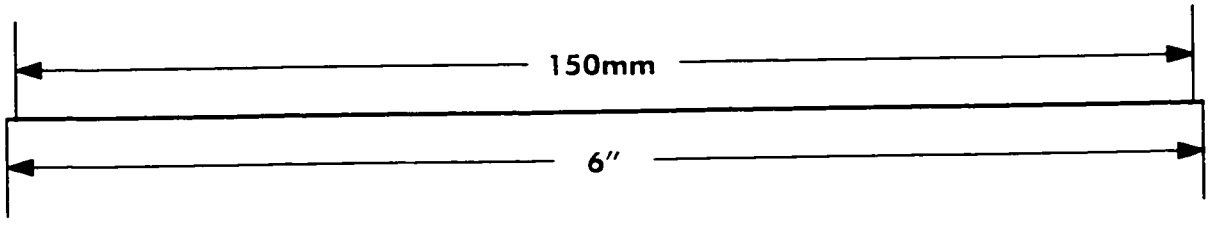
Jayne A. Goldstein was born in Toledo, Ohio in 1956. She received a B.S. in Cell and Molecular Biology from San Francisco State University in 1986. She received her Ph.D. in Pathology from the University of Washington in 1998.

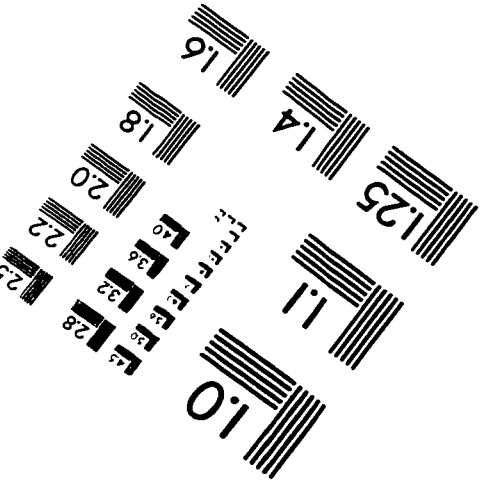
IMAGE EVALUATION
TEST TARGET (QA-3)

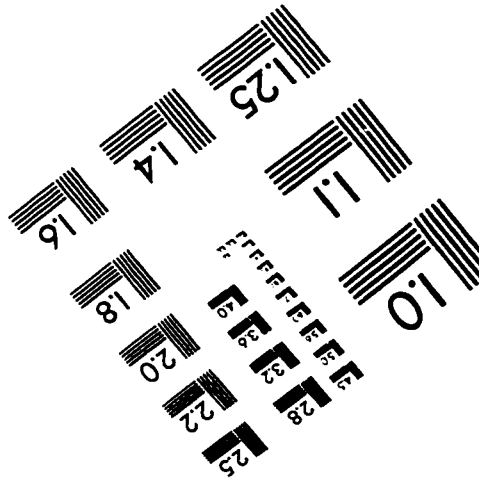












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989