Dissecting the Pathology of Pseudoachondroplasia through Structural and Biosynthesis Studies on Cartilage Oligomeric Matrix Protein

By
Asawari Mokashi

A DISSERTATION

Presented to the Department of Biochemistry and Molecular Biology
And the Oregon Health and Sciences University
School of Medicine
in partial fulfillment of
the requirements for the degree of
Doctor in Philosophy

May 2003

School of Medicine
Oregon Health & Science University

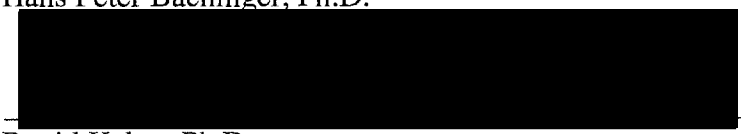
CERTIFICATE OF APPROVAL

This is to certify that the Ph.D. thesis of

Asawari Mokashi

has been approved

Hans Peter Bächinger, Ph.D.



David Kabat, Ph.D.

Gail M. Clinton, Ph.D.

Richard G. Brennan, Ph.D.

William R. Skach, MD

TABLE OF CONTENTS

Acknowledgements

I would like to thank my advisor, Dr. Hans Peter Bächinger for the excellent guidance given to me. He gave me plenty of opportunities to explore and find what I really liked to do and fostered independent thinking. I really enjoyed working with him because of his "Science is fun" philosophy and for giving me the rare opportunity of learning various biophysical techniques.

I am grateful to the members of my research advisory committee; Dr. David Kabat, Dr. Richard Brennan, Dr. Gail Clinton and Dr. William Skach for their invaluable advice and helping me gain a focus on my experiments. I thank them for agreeing to evaluate my research. I am especially thankful to Dr. Kabat for all his help and support.

My friends at the Bächinger laboratory, Dr. Kerry Maddox, Dr. Janice Vranka, Bruce Boswell, Dr. James Bann, Dr. Robin, Dr. Kazunori Mizuno, Eric Steele and late Diane Mechling provided me with hours of helpful discussion as well as a great support, for which I am extremely grateful. I wish to thank Dr. Michael Sussman of Shriners' Hospital for Children for taking active interest in this project and providing me with the invaluable patient samples. I wish to thank the entire faculty and staff of Shriners' research, especially Dr. Lynn Sakai and Dr. William Horton for taking active interest in the COMP project and sharing ideas and resources.

Balancing research with motherhood is not a simple task. I am extremely grateful to Dr. Bächinger for being very supportive. A special thanks to my family, neighbors and an excellent day care provider who rallied round to look after my daughter, Aesha. Aesha's cheerful and adjusting nature has made going to school easier.

I am indebted to my husband Samir for being so supportive and encouraging through the ups and downs of my life as a student and a mother.

ABSTRACT

Mutations in the cartilage oligomeric matrix protein (COMP), an extracellular matrix glycoprotein secreted abundantly by the cartilage, cause a developmental disorder of the skeleton called pseudoachondroplasia (PSACH). The electron microscopy of the patient cartilage reveals the presence of the diagnostic fingerprint-like structures in the rough endoplasmic reticulum (rER) of the chondrocytes. Though several skeletal tissues secrete COMP, none of these, with exception to the cartilage show the archetypical fingerprints in their rER. Most of the mutations are mapped to the thrombospondin type III sequence (COMP-3) repeats of COMP.

The goal of this project was to identify the molecular mechanisms cause PSACH. To do this, first, I needed to determine the effect of PSACH mutation on the structure of the protein. Second, I needed to understand the implication of the structural anomaly in COMP at a cellular level, i.e., I needed to identify the cellular processes that prevented export of COMP out of the rER and the consequences of the COMP retention on cartilage integrity.

The effect of PSACH mutation on the structure of COMP was determined by biophysical characterizations of the recombinant peptides. These studies showed that the mutation affected the structure of the COMP-3 domain as well as its calcium-binding properties, suggesting that the non-native structure of the COMP-3 domain was responsible for the retention of COMP molecules in the rER. However, the mutation did not influence the structure of the globular C-terminal domain, implying that this domain was capable of forming proper associations with its ligands in spite of the mutation. These associations could occur prematurely within the cells and play a major role in PSACH pathology.

The processing of COMP in the rER of chondrocytes under normal and pathological conditions was studied. The effect of the non-native COMP-3 domain on the processing of COMP was observed by immunofluorescence and affinity chromatography techniques. These studies showed the mutation resulted in selective interactions of mutant COMP with the rER chaperones and nascent ECM proteins, causing their retention within the chondrocytic rER.

COMP secretion by primary chondrocytes was studied by metabolic labeling of COMP. These studies showed that COMP was secreted at slower rates in the chondrocytes than the tissues that do not retain mutant COMP, providing an insight on the tissue-specificity of this disorder as well as the process of aggregate formation.

Thus various techniques were used to obtain an understanding of the disorder at molecular and cellular levels, which allowed me to propose a model for PSACH pathogenesis.

Chapter 1.
Introduction

Pseudoachondroaplasia (PSACH) is a disorder where mutations in the genes encoding the extracellular matrix (ECM) proteins-COMP or type IX collagen affect the normal development of cartilage and long bone. PSACH can also be categorized as an endoplasmic reticulum (ER) storage disease in which COMP and type IX collagen are retained in the ER of chondrocytes in complexes along with several other proteins. PSACH caused by mutations in COMP is the focus of my thesis. In order to understand this disorder, the goal of my research project has been to study the effects of COMP mutations at the molecular level (i.e. on the structure and function of COMP) and the cellular level (protein processing in the endoplasmic reticulum). Here I discuss in detail protein biosynthesis in the ER, COMP structure, functions and give a histortorical perspective of COMP. I also briefly discuss the role of cartilage in bone development to give a clear picture of PSACH and the complex issues that make this project a difficult but an interesting undertaking.

A: COMP is Associated with Pseudoachodroplasia

1.1 Chondrodysplasias - Cartilage Development Gone Awry:

(Reviewed in (Horton and Hecht 2000) and (Horton 1993)). Cartilage is a highly specialized connective tissue that is made up of cells called chondrocytes embedded in an amorphous, gel-like extracellular matrix. The matrix consists of proteins that give it structure. There is an exchange of information between the chondrocytes and their surrounding matrix. As a result, groups of chondrocytes can have different phenotypes. The composition of the cartilage matrix and the presence of different growth factors and

morphogens in it can determine the morphology of the chondrocytes. Conversely, chondrocytes can control their external environment by secreting different ECM proteins and growth factors. Depending on the morphology of the chondrocyte, and the composition as well as the amount of the ECM, cartilage can have different properties, and perform different functions. This property allows the cartilage to be able to undergo a series of differentiation steps to form bone. Cartilage is classified as hyaline cartilage elastic cartilage and fibrous cartilage.

Of the three cartilage types, hyaline cartilage is the most abundant and found in long bones like the femur and humerus. It plays a pivotal role in the development and elongation of the long bones because of its ability to create an environment essential for successive steps of bone formation. Through a series of steps, mesenchymal cells destined to form bones differentiate into a compact "cartilage model." The cartilage undergoes a process called *endochondral ossification* where the chondrocytes proliferate, differentiate, hypertrophy and are finally replaced by the bone. In vertebrates, most of the skeleton is formed and elongated by endochondral ossification. Thus a mutation in a gene that affects cartilage formation or differentiation, effectively also affects the development of long bones. Such mutations lead to a group of diseases known as chondrodysplasias (*chondro*-cartilage, *dys*-abnormal, *plasia*-growth). Chondrodysplasias are characterized by dwarfism and osteoarthritis in addition to other symptoms. Many known skeletal dysplasias are caused by mutations in fibroblast growth factor receptor 3 that is critical for cell signaling in cartilage development. Other dysplasias are caused by mutations in collagens like type II and type IX collagens and glycoproteins like cartilage

oligomeric matrix protein. These are ECM proteins that are either cartilage specific or secreted in abundance by the chondrocytes.

1.2 Mutations in COMP Cause PSACH:

Pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (MED), two autosomal dominant disorders are examples of chondrodysplasias. PSACH is dominantly inherited form of short-limbed dwarfism (Figure 1). In contrast to other growth disorders, patients appear normal at birth and the disease does not manifest till the child starts walking. Children have waddling gait, reduced flexing at elbows, shoulders, hip joints, and excessive mobility of wrists, fingers and joints of feet and ankles in all directions. Most epiphyses including those of hips, knees, elbows ossify later than normal and are markedly deformed at the insertions of ligament and joint capsules. Severe deformation occurs at sites that bear excessive stress like knee joints and ankles resulting in early onset osteoarthritis, and abnormalities in the epiphyses of hands, long bones and hips that characterize this disorder. However cartilage and bones protected from excessive stress, like the base of the skull, face and sacrum are unaffected. Electron microscopy of cartilage from a patient with PSACH reveals highly structured, fingerprint like, alternating electron dense and electron lucent bands in the rough endoplasmic reticulum of the chondrocytes (Figure 2). These highly symmetrical structures are archetypical of this disorder (Cooper, Ponseti *et al.* 1973) and used to differentiate PSACH from other dwarfing disorders.

MED is an autosomal dominant disorder that affects the epiphyseal cartilage of peripheral joints. It is also characterized by early-onset degenerative joint disease. Its

symptoms on the whole, are not as severe as PSACH and MED is not associated with dwarfism. Besides COMP, mutations in any one of the three chains of type IX collagen, cartilage specific extracellular protein, matrilin-1 and solute carrier member 26, (Briggs 2002) also lead to MED.

Briggs *et al.* linked PSACH to the pericentromeric region of chromosome 19 in 1993 (Briggs, Rasmussen *et al.* 1993). Within a year, Newton *et al.* located the gene to cartilage oligomeric matrix protein (COMP) on chromosome 19 in the band p13.1 (Newton, Weremowicz *et al.* 1994). Since this glycoprotein was abundantly expressed in cartilage, it was an ideal candidate to screen for PSACH and MED. In 1995, Briggs *et al.* found mutations in the COMP gene located on chromosome 19p13.1. in the patients with PSACH and MED (Briggs, Hoffman *et al.* 1995). The mutations include deletions, substitutions or insertions in the COMP gene but no mutations have been discovered that would result in premature stop codons, such as out-of-frame deletions and insertions. Some mutations associated with mild forms of PSACH are found in type IX collagen. However, there is disagreement whether the phenotypes associated with type IX mutations are categorized as MED or as very mild forms of PSACH (Briggs and Chapman 2002).

1.3 The History of COMP:

Fife *et al.* first isolated a non-collagenous, high molecular weight protein (>400 kD) from mammalian (human, porcine, equine, bovine and canine) hyaline cartilage in 1984 (Fife, Hook *et al.* 1985). This protein was reduced to ~160 KD subunits by β mercaptoethanol and made up 5-10% of the total non-collagenous proteins in the articular

cartilage. This protein was distributed in the cartilage matrix and around the chondrocytes in the territorial matrix (Franzen, Heinegard *et al.* 1987) showed that this protein appeared at a specific time during chondrogenesis (cartilage development), distinct from other cartilage specific proteins like type II collagen and was found at high levels in the articular cartilage. Since this protein appeared early in bone development, it was thought that it played a role in cartilage development and bone morphogenesis. Interest in this protein among various groups increased when it was shown that it was a potential serum and synovial fluid marker for osteoarthritis (Saxne, 1992).

Mörgelin *et al.* (Mörgelin, Engel *et al.* 1992) were the first to refer to this oligomeric protein by the descriptive term-cartilage oligomeric matrix protein (COMP). They isolated COMP in its native state from Swarm rat chondrosarcoma and looked at its structure by rotary shadowing electron microscopy. It had a bouquet-like shape with five 28 nm long flexible arms terminating in a globular domain. The globular domain had a diameter of 4.8 nm. The five strands joined in a single cylindrical domain that was 7.7 nm long and had a diameter of 3.3 nm.

Later Hedbom *et al.* (Hedbom, Antonsson *et al.* 1992) characterized this acidic protein. The molecular weight of the intact protein was established to be 524 kD by sedimentation equilibrium centrifugation. On electrophoresis under reducing conditions, COMP migrated as a single band of 100 kD.

Thus tying the electron microscopy data of Mörgelin *et al.*, which showed COMP as a five-armed structure together with the electrophoresis data, it was established that COMP was a pentamer of similarly sized subunits. It was rich in acidic residues (20% Asp and 12.5% Glu/Gln). Absence of hydroxyprolines established that the protein was non-

collagenous. The presence of 25 cysteine residues/1000 residues indicated an extensive disulfide bonding. It contained N-linked oligosaccarides, but no O-linked oligosaccharides or keratan sulfates indicated by the absence of galactosaminitol.

Oldberg *et al.* (Oldberg, Antonsson *et al.* 1992) were the first to clone the COMP cDNA isolated from a bovine tracheal chondrocyte library and a Swarm rat chondrosarcoma library. Sequence analysis revealed the COMP monomer to be a protein of 755 amino acid residues with a calculated molecular mass of 82.7 kD. COMP cDNA expressed in the COS cells showed that COMP protein was a homopolymer of five disulfide linked identical subunits.

Newton *et al.* cloned and sequenced human COMP in 1994 (Newton, Weremowicz *et al.* 1994). They located the COMP gene to band p13.1 on chromosome 19. This localization helped Briggs *et al* to pin down COMP as the protein responsible for PSACH as described previously.

The primary sequence of COMP is homologous to that of a family of extracellular matrix glycoproteins, the thrombospondins (TSP) (Mörgelin, Engel *et al.* 1992; Oldberg, Antonsson *et al.* 1992; Bornstein and Sage 1994) and (Adams and Lawler 1993). There are four TSP members, i.e. TSP1, TSP2, TSP3, TSP4 in addition to COMP (also known as TSP5). Like most ECM proteins, the thrombospondin glycoproteins are modular. Each subunit of the multimeric protein contains multiple copies of sequence motifs that are homologous to domains in other proteins. Each monomer is composed of:

1) N terminal oligomerization domain that form an α helical coiled coil

2) 4 EGF like repeats

3) Thrombospondin type 3 domain (COMP-3) where most PSACH mutations occur and

4) Globular C terminal domain

Figure 3 shows the ultrastructure of a COMP pentamer. All TSP family members have similar type 3 repeats and globular C terminal domains with a 60% identity among them. However COMP lacks the N-terminal heparin binding domain that is present in the other TSPs and the first 83 residues in COMP are unique. Although several functions have been suggested for COMP, only recently studies have shed light on some of its possible functions which will be discussed later.

The COMP gene consists of 19 exons (Cohn, Briggs *et al.* 1996). Exons 1-3 encode the amino terminal oligomerization domain, exons 4-8 encode four EGF-like repeats, exons 8-14 encode the eight thrombospondin type III sequence repeats and exons 14-19 encode the globular carboxyl-terminal (CT) domain.

1.4 Structure of COMP:

The first 46 residues of the monomer are involved in the pentamerization and the formation of **the N terminal domain.** The crystal structure indicates that this domain ia composed of five individual chains that form a parallel coiled-coil structure (Malashkevich, Kammerer *et al.* 1996). The chains are folded predominantly as right-handed α helices that wrap around the other chains forming a left-handed superhelix. Complimentary hydrophobic interactions and disulfide bridges between the α helices stabilize the superhelix. This structure is extremely thermostable. The mid-point of the thermal transition (T_m) is >100°C. The pentameric bundle encloses a hydrophobic pore whose diameter varies between 2Å and 6Å that can enclose small apolar groups.

The four EGF-like type II repeats of which two are presumed to be Ca^{2+} binding repeats (Handford, Mayhew *et al.* 1991) may be involved in protein-protein interactions. This domain is one of the smallest, but the most widely distributed modules found in ECM proteins with biologically diverse functions. EGF binds to cell surface receptors and some EGF-like domains are involved in protein-protein interactions, but there is no direct evidence of any such function in the EGF-like domains in COMP (Maurer and Hohenester 1997), nor are there any known disorders linked to this domain in COMP.

The type III domain is made of eight sequence-repeats and is conserved in all the TSP family members (Figure 4). It is also known as calmodulin-like domain as it contains eight contiguous calmodulin like sequence repeats with consensus sequence N- (D)Q-D-D-DG-GDAC(D)-D-D-D--------DNCP--- (Briggs, Hoffman *et al.* 1995). In calmodulin, Asp residues line the calcium-binding pocket and bind to Ca^{2+} ions by charge-charge interactions. This domain is known to bind to Ca^{2+} ions in TSP4 (Lawler, McHenry *et al.* 1995) but up till now there were no data showing whether COMP type III domain also bound to calcium. It has 17 cysteines but the disulfide bonds within the protein have yet to be established. Human and bovine type 3 domains have a RGD cell-binding motif. However we cannot say for certain if this motif has a functional cell-binding role as rat COMP lacks this motif. Type 3 domain is the predominant locus for the PSACH mutations and to date, 37-38 out of the 40 known COMP mutations linked to PSACH have been localized to this domain (Briggs and Chapman 2002).

The C-terminal globular domain is known mainly to be associated with MED, a disorder similar to PSACH, but with less severe characteristics. It is known to bind to types I, II and IX collagens in a zinc dependant manner (Rosenberg, Olsson *et al.* 1998;

Holden, Meadows *et al.* 2001 and Thur, Rosenberg *et al.* 2001). It also interacts with fibronectin, an ECM protein in a divalent cation dependant manner (Di Cesare, Chen *et al.* 2002). There are two cysteines in this domain. It is not known if they form an intradomain disulfide bond or an interdomain disulfide bond with one of the 17 cysteines in the type III sequence repeats. Three PSACH mutations have been found in this domain (Briggs and Chapman 2002).

Post–translational modifications: ***N-Linked glycosylation:*** COMP has no O-linked glycosylation. N-linked glycosylation through distinct asparagine residues adds 3571 ± 291 Da to the mass of a COMP subunit. Asn-101 is substituted with oligosaccharide of mass 1847.2 ± 6.6 Da. Asn-171, located in the CT domain is substituted with an oligosaccharide of mass 1724 ± 226 Da (Zaia, Boynton *et al.* 1997). Fetal COMP has a heterogeneous pattern of substitution at Asn-101. The differences in glycosylation between adult and fetal COMP may cause different interactions leading to different patterns of distribution of COMP seen in adult and fetal cartilage suggesting a role of COMP in cartilage development.

Disulfide bonding: The type III domain of COMP has 17 cysteine residues and the globular domain has 2 cysteine residues. Disulfide mapping of COMP is in progress in at Shriners Hospital for Children, Portland, OR. Disulfide mapping in TSP2 showed consecutive bonding order in the type III domain as well as connection of type III domain with the C-terminal domain through a disulfide bond (Misenheimer, Hahr *et al.* 2001). But it is not known whether such connectivity between the domains exists in COMP.

1.5 Functions of COMP:

Cartilage, tendons and ligaments secrete COMP in abundance. It is also secreted in small quantities in the trachea, eye, skeletal muscles, heart, placenta, skin and blood vessels and smooth vascular muscles (Fang, Carlson *et al.* 2000). COMP fragments in synovial fluid and blood plasma serve as early diagnostic markers for rheumatoid arthritis, osteoarthritis. In spite of its abundance, COMP has no known functions. Based on its distribution in the ECM, homology to thrombospondins and structural features, several functions have been hypothesized in the cartilage.

- ***Structural stability of ECM through protein-protein interactions:*** *In vitro* studies show that COMP can bind to several ECM proteins like fibronectin and types I, II and IX collagens through its peripheral globular C-terminal domain. Being a pentamer, COMP has the capability to have pentavalent associations with these proteins forming a lattice that stabilizes the ECM. It may mediate fibril formation or mediate interactions between the collagens and other fibril forming proteins. Thus, if a mutation causes structural defects in COMP, it may lead to either secretion of the defective molecules or their retention in the ER. The presence of defective COMP molecules or the low levels or the absence of normal ones would prevent meaningful protein interactions in the cartilage matrix and destabilize the matrix.
- ***Storage and delivery of morphogenic hormones:*** Some ECM molecules regulate the delivery of various growth factors and hormones to the cells by binding to them and releasing them in controlled fashion. For example, fibrillin, the major component of the matrix microfibrils binds the to transforming growth factor β (TGF β) and regulates their release (Kaartinen and Warburton 2003). Similar role has been implicated for COMP. The oligomerization domain of COMP is a 7.5 nm long hydrophobic pore with a

diameter of 0.2-0.6 nm. Vitamin D and retinol bind in this pore. Retinol binds in the form of physiologically active all-*trans* retinoic acid with K_d comparable to retinol-binding protein (Guo, Kammerer *et al.* 2000). The physiologically active forms of vitamin D promote cartilage formation, bone mineralization and regulates several proteins in these tissues. All-*trans* retinoic acid can activate homeobox gene clusters, influence limb formation and stimulate matrix calcification and collagen synthesis in growth plate chondrocytes of the bone. COMP may be involved in storing hydrophobic hormones and morphogens and delivering them to their cellular targets, therefore playing an important role in regulating formation and repair of cartilage and bone.

- ***Cellular proliferation:*** Studies on the head of rat femur bone by Shen *et al.* showed that the COMP distribution and the sites of COMP synthesis changed with the age of the developing bone (Shen, Heinegard *et al.* 1995). They hypothesized a putative role for COMP in cartilage development. The presence of COMP around the actively dividing cells in the proliferative zone of the developing cartilage and the low levels or marked absence in non-dividing cells (hypertrophic and the resting zone respectively) led to the speculation that it may regulate chondrocyte proliferation and differentiation. Thus tissue distribution of COMP as well as its ability to bind to morphogens make it an ideal candidate to regulate cellular proliferation.

- ***Signaling between external load and cellular response:*** Smith *et al.* (Smith, Zunino *et al.* 1997) suggest that in the tendon, COMP synthesis is upregulated in response to external load on the limbs. They further propose that COMP may be involved in resisting the external load by acting as a candidate molecule in the chain of signaling between external load and cellular response to this load. The evidence that at birth, COMP is

present in the tendons at very low levels and increases rapidly during growth, when new biomechanical demands are placed on the tissues further strengthens this hypothesis. Giannoni *et al.* showed that the expression of COMP in articular cartilage is sensitive to long term cyclic compression and that this sensitivity is blocked by blocking anti β-1 integrins. This strongly suggests that in the joints cell/matrix interactions are important events in the mechanotransduction of COMP and that COMP plays a role in resisting external load (Giannoni, Siegrist *et al.* 2003). Also, PSACH associated with COMP mutations does not manifest till the affected child is able walk, thus applying load on joints, tendons and ligaments.

In spite of various functions implicated for COMP, COMP null transgenic mice, with heterozygous and homozygous null mutations do no show any marked phenotype, suggesting that other proteins may compensate for the absence of COMP (Svensson, Aszodi *et al.* 2002).

1.6 PSACH is an ER storage disease:

ER storage disorders (ERSDs) are diseases caused by retention in the ER of a critically important protein that is unable to reach its target site and perform its function (Kim and Arvan 1998). The retained proteins have mutations that cause structural changes and hence retained by ER mechanisms that monitor the fidelity of protein biosynthesis. Cells affected by ERSDs routinely show dilated ER partly due to accumulated abnormal protein and also due to supranormal accumulation of several ER resident proteins that participate in removing misfolded proteins from the ER. The accumulated proteins may either be in a soluble form or present as fibrillar or granular aggregates. The diagnostic

marker of PSACH is highly dilated ER which contains fingerprint like complexes which are known to contain COMP (Maddox, Keene *et al.* 1997). This suggests that PSACH is an ERSD. (It is to be noted that in PSACH, unlike other ERSDs, the proteins are retained in the ER as highly symmetrical complexes.) The retention of COMP in the chondrocytes gives rise to several questions about how PSACH mutations affect COMP biosynthesis. An overview of protein processing in the ER will give a broader perspective of the events that occur at the cellular level that eventually lead to PSACH.

B: Protein Processing in the Endoplasmic reticulum

Newly synthesized soluble, secretory proteins, proteins destined for the organelles like the Golgi body, lysosomes, vacuoles as well as transmembrane proteins are carried to their target destination by microsomal transport pathway. This pathway consists of rough endoplasmic reticulum (rER), Golgi body and transport vesicles, tubules and cisternae. The proteins are synthesized on the ribosomes. They undergo biosynthesis, folding and assembly as they sequentially move from the ribosomes, to endoplasmic reticulum and Golgi complex before they are finally targeted to their destination on a membrane or out of the cell. Transport vesicles, tubules and cisternae ferry the proteins through the different compartments (Figure 5).

The rER is the entry point for these molecules to go into the transport pathway. Its membrane constitutes more than half of the total membrane of an average animal cell. It is organized into a netlike labyrinth of branching tubules and flattened sacs extending throughout the cytosol. The space enclosed by the rER membranes is called ER lumen or cisternal space. The rER membrane separates the rER lumen from the cytosol and mediates selective transfer of molecules between the two compartments. rER also plays a major part in lipid and protein biosynthesis. Most of the proteins are folded in this compartment, though a few proteins undergo folding in the Golgi body (Chevet, Jakob *et al.* 1999).

1.1Translocation:

This is the first step in protein transport in which the proteins synthesized on the ribosomes are inserted into the ER (reviewed (Matlack, Mothes *et al.* 1998). Multicomponent protein complexes, termed translocons form a pore in the rER membrane and facilitate the transport of nascent proteins across or insertion into the ER membrane (Corsi and Schekman 1996). In order to be translocated, the proteins need to be selected away from the other proteins present in the cytoplasm. They must transfer across the membranes that are normally barriers to such movements. The soluble proteins must be distinguished from the transmembrane proteins so that the latter can be moved laterally into the membrane bilayer after correct orientation with respect to the plane of membrane. The permeability of the bilayer must be maintained at all times. There are two pathways for targeting and translocation:

Co-translational translocation: It is the primary translocation pathway in the mammals. It also occurs in yeast. The protein precursor remains attached to the ribosome while it is targeted to and passes through the translocon on the ER membrane.

Post-translation translocation: Mainly occurs in Yeast (*Saccharomyces cerevisiae*). Here the polypeptide is fully translated and released from the ribosome before engaging the translocon.

The proteins destined for translocation are identified by their N-terminal signal sequence specified by a stretch of 7-20 amino acid residues and recognized by the signal recognition particle (SRP). In yeast, the extent of hydrophobicity determines the targeting pathway of the proteins. More hydrophobic core dictates co-translational import into the ER. Another set of proteins called the nascent polypeptide-associated

complex (NAC) prevents the nascent polypeptides that do not contain the signal sequence from being targeted to the ER membrane. The SRP in the ribosome-nascent polypeptide-SRP complex binds to the SRP receptor located on the rER membrane in a GTP-dependent manner. The receptor forms a part of the translocon. The other protein components of the translocon in mammals are TRAM (translocating chain-associated membrane protein) and the pore-forming heterotrimeric Sec61p. The nascent polypeptide passes through the aqueous pore (diameter ~20 Å) of the translocation apparatus. In co-translational translocation, as amino acids are being added to the growing chain, the elongating peptide emerges from the ribosome and continues through the Sec61 channel. For soluble proteins, the final step of translocation is full transfer from the pore into the ER. Once inserted within the ER, signal peptidase enzyme cleaves off the signal peptide co-translationally.

1.2 Summary of Protein Folding:

I) Protein Folding within the rER: Though the genes decide the primary structure of the proteins, co- and post-translational modifications also govern the folding, biological function and final destination of an exportable protein. For the proteins in the secretory pathway, the rER is the first and usually the major site of folding. It is also the site of modifications required for providing the correct information for proper folding. In order to ensure proper folding, the ER lumen provides the essential modifying enzymes and co-factors called chaperones to assist protein folding and enhance the efficiency of folding. Quality control (QC) mechanisms recognize and retain misfolded, incompletely folded or partially assembled copies of the proteins, ensuring that only completely and correctly

folded proteins progress out of the ER. QC mechanisms also provide mechanisms to empty the lumen of irreversibly misfolded proteins by either promoting the refolding of misfolded proteins or degrading the irreversibly misfolded proteins through ER associated degradation pathways (ERAD).

Many proteins undergo modification in the rER lumen while they are being translated. Multi-domain proteins start folding into individual domains at the N-terminus while the protein is still being synthesized. *Folding catalysts* directly or indirectly affect the protein conformation by covalently modifying the proteins (glycosylation, hydroxylation etc.) or by stabilizing a secondary structure through disulfide bond formation. Some of modifying proteins such as signal peptidase, which cleaves off the signal peptide, are located close to the translocon. Other modifications include glycosylation of the nascent polypeptide by oligosaccharyl transferase, hydroxylation of certain residues by their specific hydroxylases (e.g. prolyl hydroxylase, lysyl hydroxylase), formation of correct disulfide bonds by protein disulfide isomerase (PDI), *cis-trans* isomerization of proline by peptidyl-prolyl *cis-trans* isomerase. Glutamyl carboxylase carboxylates glutamyl residues of several blood clotting factors and calcium-binding protein through a vitamin K-dependant mechanism (Kim, Kwon *et al.* 1996). The transmembrane carboxyl-terminus of some membrane proteins destined for the plasma membrane is exchanged for covalent attachment to the sugar residue of glycosylphosphatidylinositol (GPI) anchor. *Classical chaperones* however do not participate in the covalent modifications but enhance the correct folding by providing secluded environment for proper folding and refolding of proteins and minimize aggregation. Details of the roles of chaperones in protein folding and quality control mechanisms will be discussed later.

II) ER to Golgi Transport: Once the folded proteins are deemed transport competent by the quality control, they enter ER exit sites. Here, specialized proteins called COPII associate with the cytosolic face of the ER membrane and make it bud off as vesicles and tubules. The inclusion of the cargo is either by bulk flow or by interaction with specific cargo receptors (discussed later). Once the vesicles detach from the ER, they form an intermediate compartment called VTC (vesiculotubular clusters). The carriers fuse with *cis*-Golgi compartment, releasing the soluble proteins in the cisternae. As the proteins are carried through the different Golgi compartments, namely *cis*-Golgi, intermediate golgi and *trans*-golgi they undergo further maturation, such as complex N-linked glycosylation (Rabouille, Hui *et al.* 1995). They are then packaged and targeted to their final destination- the plasma membrane or out of the cells.

Large proteins, like types I and II collagens, which cannot be packed into cargo vesicles, undergo further maturation and anterograde transport through a process called cisternal maturation in specialized compartments, composition of which changes as the proteins mature (Bonfanti, Mironov *et al.* 1998).

1.3 Role of Chaperones in Protein folding:

Protein folding is a thermodynamic process that requires a change in free energy and large changes in conformation between the folded and unfolded states. However, in the absence of cofactors and helper enzymes, folding would occur too slowly to work in the cells under the conditions of physiological pH, protein concentrations and temperature. Though small, single domain polypeptides could fold correctly by themselves, more complex proteins like large multi-domain proteins or oligomeric proteins require

additional proteins to fold and assemble correctly. Laskey first used the term "molecular chaperones" for such helper proteins (Laskey, Honda *et al.* 1978). Chaperones are defined as a wide variety of factors that facilitate the formation of native structures of proteins and nucleic acids without themselves forming a part of the protein or specifying the protein structure. According to Ruddon *et al.* (Ruddon and Bedows 1997) protein folding (the "on" pathway) is in competition with misfolding and aggregation (the "off" pathways). Chaperones foster correct folding by circumventing the "off" pathways. Thus, they assist in enhancing the overall efficiency and speed of protein folding and assembly. They are often present as a complex of proteins that act sequentially in protein folding by binding to intermediates that are in various stages of folding and passing them on to the next chaperone in the cascade, eventually releasing the native protein. They physically interact with the substrate proteins and lower the activation energy needed for a discreet conformational change within a protein (Kim, Kwon *et al.* 1996; Horwich 2002). Chaperones occur in both the sites of protein synthesis and folding, namely, rough endoplasminc reticulum (rER) and cytosol. Cytosolic chaperones play a key role in folding, transport and biological activity of cytosolic proteins as well as those targeted for transport to cell organelles like the mitochondria and the nucleus. The chaperones in the rER help the folding and assembly of exportable and transmembrane proteins (Kim and Arvan 1998)

1.4 Molecular Chaperones Involved in Protein Folding in the rER:

HSP70 Chaperones: (reviewed in (Bukau and Horwich 1998)) *HSP70* chaperones along with their co-chaperones assist in folding of several proteins in almost all the cellular

compartments. They were first identified as proteins that were induced under stress conditions and played crucial compensatory roles that allowed cell survival under stress situations by limiting and potentially reversing the aggregation of misfolded proteins. Later they were also found to play an essential role under normal physiological conditions (Bukau and Horwich, 1998). 1) They assist in folding of newly translated proteins. 2) They guide translocating proteins across organellar membrane through action at both - *cis* and *trans* sides. 3) They disassemble protein oligomers. 4) They facilitate proteolytic degradation of the unstable proteins by keeping them in soluble, protease–sensitive forms.

BiP, which is also known as GRP78 (glucose-regulated protein 78), is an example of a HSP70 chaperone. It is a soluble protein that is retained in the lumen of the rER through the KDEL ER retention sequence. It is ubiquitous and highly conserved through evolution. BiP binds to the nascent peptides as they are being inserted in the rER lumen. This binding has a two-fold use: 1) rER has a very unique environment; it has a very high concentrations of nascent and unfolded polypeptides as well as an oxidizing environment created by the presence of oxidized glutathione in the lumen (Kim and Arvan 1998). This environment is conducive to improper inter- and intramolecular associations. By binding to the nascent polypeptides, BiP prevents such improper associations between them (Corsi and Schekman 1996). 2) Since BiP is also an ATPase, it also provides energy for the movement of the polypeptide across the rER membrane (translocation) by hydrolyzing ATP.

BiP consists of an ATPase domain and substrate binding domain. Both the domains are strongly coupled. ATP binds to BiP with high affinity, but the BiP-ATP complex has

very low affinity for the substrate. ADP binds to BiP with low affinity, but BiP-ADP complex has high affinity for the substrate. The substrate–binding domain consists of a substrate-binding pocket and a latch made of an α-helix that controls the entry and exit of the substrate. Co-chaperones like BAG-1, Hip and Hop are essential for the chaperoning activity of HSP70 molecules. (Figure 6 summarizes the HSP70 cycle in the bacteria)

The HSP70 reaction cycle Reviewed in (Bukau and Horwich 1998; Fink 1999). The ATP-HSP70 complex recognizes heptameric motifs in non-native proteins. These motifs are clusters of hydrophobic residues flanked by basic residues and a distinct absence of acidic residues. These patches are usually buried inside a correctly folded protein and a) are exposed on nascent polypeptides during translation or b) become available after perturbations in their structure or c) exposed on unassembled sub units of proteins. Nascent substrate makes hydrophobic contacts and hydrogen bonds with the substrate-binding channel of HSP70. Co-chaperones-HSP40 and BAG-1 bind to the ATPase domain and stimulate ATP hydrolysis to ADP by increasing the rates of release of ATP from HSP70. This locks the substrate in the substrate-binding pocket. HSP40 dissociates rapidly replaced by HIP (HSC interacting protein). HIP stabilizes the ADP bound state of HSP70 complex. HSP70 undergoes adenine nucleotide exchange, replacing ADP with ATP under the stimulation of another co-chaperone HOP. The substrate is released from the complex. Every time the polypeptide is released, it has an opportunity to progress along the folding pathway, but will continue to interact with the chaperone till it no longer presents the HSP70 binding motifs (Gething 1999).

HSP70 chaperones and quality control: Mutations that perturb the structure of the nascent proteins and destabilize them promote aggregation. These aggregates can be

highly toxic to the cells or impair normal cellular functions (Bucciantini, Giannoni *et al.* 2002); (Horwich 2002). HSP70 molecular chaperones bind to the misfolded proteins and not only promote their folding, but prevent them from forming toxic aggregates. GRP78 plays an important role in initiating the unfolded protein response under stress conditions (discussed later) and targets irreversibly misfolded proteins for degradation (also discussed later). Thus HSP70 chaperones play a crucial role in protein folding and cell viability.

Protein disulfide isomerase (PDI): Many secretory proteins contain disulfide bonds (-S-S-) that are required for their correct folding and/or activity. These bonds are formed by the oxidation of sulfhydryl groups (-SH) of two cysteines residues. The ER provides an oxidizing environment that favors disulfide bond formation, though non-native conformations often occur under uncatalyzed conditions. The cytoplasm, which has considerably less oxidizing environment (redox potential of –150 mV versus –230 mV in ER), has virtually no disulfide bond formation (Kim and Arvan 1998). PDI binds to the backbones of the nascent protein substrates with broad specificity and catalyzes the rearrangement of disulfide pairings in a nascent protein until its native structure, which is also most thermodynamically stable, is achieved.

The RS^- group of PDI attacks the -S-S- group in the protein and forms a covalent disulfide linked intermediate with the substrate. The free thiol generated on the substrate attacks another disulfide bond and generates another pairing (Stryer 1995).

PDI has two thioredoxin-like regions, each containing a cys-gly-his-cys sequence that may represent two independently acting catalytic sites. As shown in figure 7, in *E. coli*, ER oxidoreductase (ERO1), in conjunction with co-factor FAD provides oxidizing

equivalents to PDI. (ERO1-L beta) is the human analog. In the periplasm of E. coli, excess electrons generated are disposed through the respiratory chain, passing to oxygen. In eukaryotes, besides the $FAD/FADH_2$ system, another unknown protein is thought to be involved in the oxidation of the reduced PDI molecules (Fewell, Travers *et al.* 2001).

PDI and ER quality control: PDI also has the KDEL ER-retention signal in its carboxyl terminal, which binds to the KDEL receptor in the microsomal membrane. It is a multifunctional polypeptide identical to the β subunit of prolyl-4-hydroxylase and a cellular thyroid-binding enzyme. Since PDI has strong affinity for unfolded peptides and hydrophobic molecules like steroids and thyroid hormones, it may also be involved in recognizing incompletely folded proteins and preventing their aggregation while fostering correct disulfide bond formation (Noiva 1999).

ERp72: This chaperone belongs to PDI family of thioreductases and contains three copies of CXXC- active site motif found in PDI. It is found to be associated with misfolded mutant proteins and has yet to be seen in complex with normal proteins. It is induced during unfolded protein response by a cell under stress conditions and may be involved in catalyzing correct disulfide bond formation in the misfolded proteins. It has also been implicated in ER-associated degradation of misfolded proteins (Kim and Arvan, 1998).

Calnexin and calreticulin: Calnexin and calreticulin are involved in quality control machinery of N-linked glycoproteins. Calnexin is an ER transmembrane phosphoprotein that belongs to the family of lectins. Its luminal domain, which is similar to calreticulin, contains a KDEL ER retention signal sequence. It contains a proline-rich P-domain that binds to calcium with high affinity. Carbohydrates bind to this domain with high

specificity. Homologs of calnexin include calmegin and calnexin-t, which are found in testis. It is believed that all the calnexin family members assist similarly in protein folding. Calnexin/calreticulin play an important role in ER quality control by binding to, and retaining unfolded/incompletely folded/misfolded proteins within ER. Calnexin also binds to nascent peptides containing N-glycosylation sites as they are being translocated into the ER lumen and aids their movement across the ER membrane.

The Calnexin/calreticulin cycle: All nascent glycoproteins are glycosylated with a pre-assembled oligosaccharide (Glc3Man9GlcNAc2) at the N residue in the glycosylation site (N-X-S/T-) shown in figure 8. Figure 9 summarizes the calnexin/calreticulin cycle. Glucosidase I and II sequentially remove the outer most glucose and the two remaining glucose moieties respectively. As the core oligosaccharide is being trimmed, chaperones like PDI and ER57 help the nascent proteins to fold correctly. Another ER resident protein, UDP-glucose: glycoprotein galactosyl transferase (UGGT) acts like a folding sensor and specifically re-glycosylates misfolded glycoproteins. Calnexin-family members bind to these monoglucosylated proteins and retain them in the ER. Glucosidase II cleaves the glucose residue, thus releasing the protein from calnexin. If the protein has been able to achieve the native structure, it can leave the ER and move down the transport pathway. However, if the protein is still misfolded, UGGT re-glucosylates it and it re-enters the calnexin cycle. Although calnexin cannot discriminate between folded and misfolded proteins, it retains the misfolded protein in

the ER, through series of bind-and-release cycles and eventually targets the misfolded protein for degradation or storage in the ER. (Chevet, Jakob *et al.* 1999). Thus glycoprotein processing and folding occur simultaneously and the **modification of the**

basic glycosylation structure in the ER is a primary mechanism that distinguishes between the folded and unfolded proteins. How does QC distinguish between slow folding proteins and misfolded proteins? Mannosidases mediate trimming of mannose residues from proteins engaged in calnexin/calreticulin-UGGT cycle. Mannosidase I and II remove the two terminal mannose residues of the oligosaccharide forming Mannose $_{5\text{-}7}$ glycans (Parodi 2000). Action of Mannosidase I tags the misfolded protein for degradation by proteosome mediated pathway (Kopito and Ron 2000). Data also suggests that inhibition of Mannosidase I suppresses degradation of misfolded N-linked glycoproteins (Parodi 2000).

Calreticulin is a soluble homolog of calnexin. It functions as a chaperone in a manner similar to calnexin. It was originally identified in the sarcoplasmic reticulum membranes and plays a minor role in calcium storage in the skeletal and cardiac muscle SR. It has the highly conserved KDEL ER-retention sequence. Calreticulin can also act as a modulator of the regulation of gene transcription by nuclear hormone receptors. It binds to the DNA binding domains of steroid receptors and transcription factors containing the KXFF(K/R)R motif and prevents their interactions with DNA *in vitro*. It functions in conjunction with other chaperones and may be a signal **to recruit** other chaperones to a nascent protein. When a misfolded glycoprotein is bound to calnexin, PDI and Erp57 rearrange its disulfide bonds (Michalak, Corbett *et al.* 1999).

UGGT: This is a soluble 170 kD glycoprotein present in the ER and possibly in the ER exit sites. As mentioned earlier, it acts as a **folding sensor** in glycoprotein biosynthesis. It recognizes the oligosaccharide and the protein moiety of the misfolded glycoproteins. It specifically requires the innermost *N*-acetylglucosamine (GlcNac) residue of the high-

mannose oligosaccharide for its sensory function. The current hypothesis is that the inner GlcNac is immobile in a folded protein because of its interaction with the polypeptide. But in the misfolded, partially folded or unfolded state, it may be more dynamic. The dynamic properties of the oligosaccharide or its influence on the polypeptide chain may determine the state of local folding of the glycoprotein. UGGT may also recognize hydrophobic patches exposed on proteins in a non-native state, similar to classical chaperones. Instead of indiscriminately reglucosylating the entire multidomain glycoprotein, UGGT reglucosylates glycans directly linked to the misfolded domain. This would allow calnexin/calreticulin to ignore the already folded portions of the protein (Ritter and Helenius 2000).

GRP94: (Reviewed in (Argon and Simen 1999). This ER resident protein belongs to HSP90 family. However, it comprises a distinct subfamily among other HSP90 proteins. Though it makes up 5-10% of the luminal proteins, not much is known regarding the structure or the modes of function of the protein. It is a dimer and is the most abundant calcium binding protein in the ER. It functions as chaperone and delivers peptides for T-cell presentation. For unknown reasons, it is also detected on the cell surface. It is expressed only in the ER of multicellular organisms. It is expressed both, constitutively, and under stress conditions through the unfolded protein response (UPR) which is discussed later. Under either condition, it is upregulated in parallel with BiP, and thought to facilitate protein folding and prevent aggregation. It is a part of the chaperone network consisting of BiP, ERp72 and calreticulin that interact with maturing proteins (Linnik, and Herscovitz, 1998). Unlike BiP, it appears to have a more limited specificity, binding only to a few substrates like IgG chains, cholesterol esterase, thyroglobulin, MHC class

II, erB2, a herpes virus glycoprotein, apolipoprotein B, collagen and protein C. Since it has been observed to bind to advanced folding intermediates and incompletely assembled oligomers (and fully folded cholesterol esterase), downstream of, or in conjunction with other chaperones, **it is likely that GRP94 may function as the "finishing chaperone" in the progressive folding of proteins by a relay of ER chaperones.**

Cyclophilins and FK Binding proteins: Immunophilins are a family of cytosolic receptors capable of binding to one of the two major immunosuppressant agents-cyclosporin A (CsA) or FK 506. Proteins that bind FK 506 are termed as FK506 binding proteins (FKBPs) and those that bind cyclosporin A are called cyclophilins (CyP). **Immunophilins function as peptidyl prolyl *cis-trans* isomerases** (PPIases) whose activity is inhibited by their respective immunosuppressant compounds. As PPIases, the immunophilins accelerate folding of some proteins both in *vivo* and *in vitro* by catalyzing the initial steps in the folding and rearrangement of proline containing proteins. Within the CyP family there are several different proteins that show a high degree of homology including CyPA (CyP-18), CyPB and CyPC. CyPA is the most abundant and ubiquitous **cytosolic** cyclophilin found in all vertebrate tissue and is present in T-cells. CyP B is a 19 kDa immunophilin with similar functions as CypA and is localized to the rER. CyP B is upregulated under stress conditions that cause accumulation of unfolded/misfolded proteins in the rER.

Chaperone networks that solubilize and refold stable protein networks: A biologically relevant fate of misfolded proteins in the ER is to escape the protective activity of chaperones and form stable aggregates. Folding/unfolding intermediates are trapped in

the non-native forms with exposed hydrophobic patches. This causes them to associate with other molecules through hydrophobic interactions. The aggregates become larger and more stable. They are enriched in anti-parallel β-strands as compared to the native state. They are seen as heat-shock granules, inclusion bodies or highly ordered fibrillar structures. Unless these aggregates are dissolved, they will impair cell function and prevent normal processing of proteins through the ER. In fact, Bucciantini *et al* have found that self-associating aggregates have inherent toxic property (Bucciantini, Giannoni *et al.* 2002). A network of heat inducible proteins consisting of HSP100 (ClpB, ClpA, ClpX in *E.coli*), HSP90 (*E.coli* HtpG), HSP70(*E.coli* DnaK) and HSP60 (*E.coli* GroEL) and α-crystalline-like small heat-shock proteins function in several ways to aid proper folding of misfolded proteins. HSP100 hexamers solubilize the aggregated proteins by acting as molecular ratchets and crowbars to pull apart aggregates and disassemble protein complexes in an ATP-dependant manner (Glover and Tkach 2001) (Mogk, Schlieker *et al.* 2003). HSP100 can disaggregate "soft" aggregates, up to 600kDa (Horwich 2002). HSP70 helps in disaggregating the complexes and mediate their refolding. The HSP100 family is also involved in proteolysis of misfolded proteins.

HSP47: This collagen-specific chaperone was one of first few chaperones to be discovered. It is expressed in parallel with collagen expression and is also known as colligin. It is associates with procollagens at every step in their folding in the ER. It binds to collagen in monomeric and trimeric form as well as procollagens in various states of folding. It dissociates from collagens in the Golgi. Under stress conditions, it prevents secretion of abnormal collagens. Studies by various labs have given insight in the precise function of HSP47 in the ER pathway (Thomson and Ananthanarayanan

2000), (Tasab, Batten *et al.* 2000) and (Macdonald and Bächinger 2001). Unlike the typical chaperones that bind to incompletely or improperly folded chaperones, HSP47 binds cooperatively and preferentially to folded triple helices without influencing the stability and folding of triple helix. HSP47 may protect the collagen triple helix that has incomplete covalent modifications from intracellular degradation. HSP47 prevents premature fibril formation of collagens in the ER possibly by masking reactive groups that are required to form complex, higher-order structures. HSP47, by preventing transport of individual molecules from ER, may concentrate procollagen molecules in specific regions of the ER and *segregate them* to the cisternal maturation pathway.

1.5 Protein Sorting in the ER:

There is a need for mechanism(s) to sort the correctly folded proteins from the incorrectly folded proteins. The mechanisms that recognize misfolded or misassembled proteins have not been clearly established. Such mechanisms would have to distinguish whether the exposed regions, that are typically buried within a correctly folded/assembled protein are from transient folding intermediates or are permanent features of a misfolded protein. The following hypotheses have been postulated to explain why misfolded proteins are retained in ER and will be discussed below:

A) Sustained interaction with molecular chaperones (Retention and retrieval of misfolded proteins)

B) Cargo receptors for anterograde trafficking of correctly folded proteins

A) Retention and retrieval of misfolded protein by ER chaperones: Chaperones like GRP78, PDI, calnexin, ERp72 interact weakly with each other in a calcium dependant manner and form a calcium matrix. This divides the ER lumen into an immobile matrix and a mobile fluid-phase. The fluid phase *concentrates* the folding intermediates, *promotes their folding* by facilitating their association with the foldases, the immobile matrix *retains* the incorrectly/incompletely proteins in the ER (Fewell, Travers *et al.* 2001), (Chevet, Jakob *et al.* 1999). All secretable folding-protein-intermediates interact with the molecular chaperones in bind-and-release cycles till they reach a certain level of conformational maturity and are competent for further intracellular transport. Since misfolded proteins are unable to bury the chaperone binding epitopes, they are associated with the folding chaperones longer than normal proteins. They cannot "escape the clutches" of the ER resident proteins. BiP, calnexin, calreticulin and GRP94 are often seen associated with misfolded proteins for extended periods of time. BiP and GRP94 association can target the misfolded proteins for degradation.

ER retention is a specific process in some cases. For instance, Eps-1, a PDI-related protein, is essential for the retention and degradation in the ER, of mutant Pma1 p, a plasma membrane ATPase.

In some cases, misfolded protein has to be *retrieved* from the Golgi to be degraded by the ERAD (Vashist, Kim *et al.* 2001). For instance, KHN, a soluble version of membrane protein in simian virus 5 hemagglutin neuraminidase, when expressed in *S. cerevisia*, fails to form disulfide-linked dimers and is rapidly degraded by proteosomes. However, to be efficiently degraded, it is essential for them to actually leave the ER, reach the Golgi and be retrived to the ER. The gene for *BST-1* is required for this

transport, and in absence of a functional *BST-1* KHN is retained in the rER and stabilized. Some the hypothesis for the necessity for the retrieval mechanisms have been suggested. Modifications in the Golgi, such as the modification of the O-mannosylated residues may signal the degradation of a protein on retrieval to the rER. However, not all retrieved proteins need such a modification to get degraded, hence carbohydrate moieties on the degradation substrates may not serve as a general mechanism for targeting a protein for degradation. Alternatively, the retrieval pathway may provide the means to partition the misfolded proteins destined for ERAD from those in the process of folding. A third hypothesis suggests that the misfolded proteins are returned to the regions of Er specialized for ERAD.

B) Cargo receptors for the anterograde trafficking: Transport competent proteins in the ER are exported from the ER to the Golgi complex through clusters of closely opposed vesicles and convoluted tubules termed as ER-Golgi intermediate compartment (ERGIC) (Bannykh 1997). Vesicles coated with COP protein mediate the export. Though most of the soluble secretory proteins are incorporated in the vesicles by default (*bulk-flow transport*), there is growing evidence that certain cargo proteins are localized to ER exit sites and are *selectively extracted* by specific receptors (or escorts) into their COPII coated transport vesicles.

P58/ERGIC-53, a calcium dependant animal lectin that cycles between ER and the Golgi complex acts as a calcium dependant cargo receptor for soluble glycoproteins (Velloso, Svensson *et al.* 2002). ERGIC-53 is a transmembrane protein. The cytoplasmic domain contains the ER exit and Golgi retrieval motif that allows it to ferry between the two compartments. The ER luminal domain has a carbohydrate recognition domain (CRD)

that recognizes and binds to subset of glycoproteins through recognition of mannose residues in the sugar moieties (asparagines-linked core glycan from which terminal three glucose residues have been removed (Man_9). Its structure resembles the CRD of calnexin. The selectivity of the sorting process is maintained by specific interactions between the cargo proteins and another conserved surface patch on the side on ERGIC53/p58 that is opposite to the CRD domain. ERGIC53 is essential for the efficient secretion of coagulation factors V and VIII. Mutations in ERGIC-53 cause a rare, hereditary bleeding disorder through the deficiency of the two factors. Cathepsin C and cathepsin Z-related proteins, two lysosomal proteins need ERCIC-53 to facilitate their exit out of the ER.

Receptor associated protein (RAP) is another example of an escort protein. It interacts with transmembrane proteins such as the LDL family of receptors and prevents their t aggregation or premature association with their ligand in the ER by escorting them out of the ER, into the Golgi. BAP31 and cathepsin A are escorts of cellubrevin and neuraminidase respectively (Annaert, Becker *et al.* 1997; Kim and Arvan 1998). They ensure that only transport competent target molecules are loaded into the cargo vesicles for transport to Golgi.

1.6 Fate of Misfolded Proteins:

Misfolded proteins are retained in the ER. It is essential that the ER remove them because they may accumulate and form cytotoxic aggregates that would hinder the normal functioning of the cell. The ER quality control consists of three different mechanisms to clear the rER of misfolded proteins: (1) Expression of proteins involved

in polypeptide folding is strongly enhanced by a process called the *Unfolded Protein Response (UPR).* (2) In the cells in which unfolded polypeptides accumulate, *translation initiation is inhibited* to prevent further accumulation of unfolded proteins. (3) Misfolded/misassembled proteins are degraded by the *ER-Associated protein Degradation* (ERAD) mechanism. Recent studies in *S. cerevisiae* have shown that the processes of UPR and ERAD are functionally linked to each other (van Laar, van der Eb *et al.* 2001). (Figure 10 summarizes the fates of the misfolded proteins in the rER.)

1) **Initiating the Unfolded Protein Response (UPR):** When misfolded proteins accumulate in the ER, the QC mechanism responds by inducing the synthesis of specific chaperones (Fewell, Travers *et al.* 2001; Patil and Walter 2001). The ER chaperones clear the misfolded proteins from ER by either directly *refolding* the proteins or *degrading* them (ER associated degradation). In the mammals, proteins like XBP, ATF6α (Activating transcription factor 6α) and ATF6 β are transcriptional upregulators of ER proteins in response to stress. ATF6α and ATF6 β are constitutively expressed as transmembrane proteins in the ER. Under stress, they are proteolytically cleaved and their soluble cytoplasmic domains translocate to the nucleus. These domains constitute the bZIP transcription factor that binds to ER stress element (ERSE) and activates the transcription of ER chaperones that enhance folding or stimulate degradation as well as XBP gene. Unspliced XBP mRNA cannot act as an efficient activator of UPR.

Another UPR mechanism involves ER transmembrane endonucleases called IREα and IRE β as well as the transcription factor XBP. IREα and IREβ sense high levels of unfolded proteins in the ER through their luminal sensory domain through their interactions with BiP. In absence of stress, BiP binds to the IREs and prevents their

oligomerization. Under stress situations like high temperatures, low pH, etc the levels of incorrectly folded proteins increase. As a result, BiP associates with misfolded proteins and is unavailable for binding to the IREs. Free IREs dimerize and form an active complex that splices XBP mRNA which can then be translated into an active transcription factor. XBP translocates to the nucleus, binds to unfolded protein response element (UPRE) that is found in the promoters of several target genes and activates their transcription. An active UPR causes higher expression of genes of chaperones like BiP, PDI and GRP94, factors needed for protein translocation, lipid metabolism, ER-associated protein degradation, ER-to-Golgi traffic, glycosylation and protein targeting from vacuole to cell surface. UPR can also induce cell cycle arrest and apoptosis.

2) Translational Control: When proteins accumulate in the ER, the cell protects itself from toxic protein accumulation by decreasing overall protein synthesis (Patil and Walter 2001). An ER kinase PERK phosphorylates general translation initiation factor eIF 2α, thus down-regulating overall protein synthesis.

Figure 11 summarizes the diverse responses induced by the UPR.

3) ER Associated Degradation: If the misfolded proteins cannot form mature structures that allow for further processing down the microsomal pathway, they may be degraded either by ER resident proteases or by cytosolic ubiquitin-proteosomal degradation system. Jointly, these two proteolytic mechanisms comprise the ER associated degradation (ERAD). Many factors needed for ERAD are induced by the UPR.

3a) ER proteases: Misassembled or misfolded proteins are the targets of degradation of ER resident proteases. Unassembled immunoglobulin light chains are degraded by a serine protease in the ER after a prolonged association with BiP (Gardner, Aviel *et al.*

1993). If tropoelastin is retained in the ER by blocking ER to Golgi transport, it is degraded by cysteine proteases (Davis and Mecham 1996). However, a critical threshold of tropoelastin must be reached before degradation can occur. Davis *et al.* speculate that the ability of tropoelastin to undergo phase transition at high concentrations and form coacervates may target it for degradation. ER-60, an ER cysteine protease is known to degrade unassembled or partially assembled oligomeric proteins like T cell antigen receptor and HMG-CoA (Urade and Kito 1992).

3b) Ubiquitin-Proteosomes mediated protein degradation: In eukaryotic cells, most cell proteins are degraded by the multicatalytic threonine protease called proteosome that is located in the cytosol. The degradation is comprised of three distinct steps: 1*)* *recognition* of the degradation substrates, 2) efflux of the soluble proteins out of the ER, into the cytosol (*retrotranslocation*) and 3) *degradation* by proteosomes (Zhang, Nijbroek *et al.* 2001).

Recognition: Glycosylated proteins are associated with calnexin before their degradation. There are multiple mechanisms that distinguish between slowly folding and irreversibly misfolded proteins. Removal of mannose moieties from the core glycans of the N-linked glycoproteins mark irreversibly misfolded proteins for degradation. Fully folded glycoproteins containing trimmed mannoses are free to leave the rER because they do not contain glucose moieties that mediate their binding to calnexin. Mechanisms for non-glycosylated proteins have yet to be identified. It has been observed that proteins are associated with certain chaperones like BiP, GRP94 and PDI for extended periods of time before they are degraded (Fewell, Travers *et al.* 2001). HSP70 co-chaperone BAG-1acts

as a link between the molecular chaperones and the ubiquitin –proteosomal pathways. It may assist the release of substrate from HSP70 and may assist substrate transfer to the proteosomes (Hohfeld, Cyr *et al.* 2001).

Retrotranslocation: Once the substrates to the proteosomal pathway have been identified, they are taken to the Sec61 translocon pore, possibly by BiP and retrotranslocated into the lumen (Bonifacino 1996). BiP may also provide energy for its retrotranslocation through the ATP-dependent process that is similar to translocation. Since the pore has a diameter of 20Å (Bonifacino 1996), *aggregated complexes as well as complexes with a diameter larger than 20Å cannot be exported out of the translocon* (Wickner, Maurizi *et al.* 1999).

Ubiquitination: Proteins have first to be modified by ubiquitin before they are degraded by the proteosomes (Reviewed in Hochstrasser, 1996). Ubiquitin, a 76 amino-acid peptide is transferred to the aberrant proteins through a cycle mediated by a group of enzymes. It is joined reversibly to its substrate by covalent isopeptide linkage between the carboxyl-terminus of ubiquitin and the lysine ε-amino groups of substrate. The C-terminus has to be activated before it can form the isopeptide bond. Briefly, ubiquitin-activating enzyme (E1) adenylates ubiquitin. The ubiquitin-AMP intermediate of the ternary complex is then attacked by a sulfhydryl group of E1, yielding an ubiquitin-E1 thiolester. The activated ubiquitin is passed to one of the several distinct ubiquitin-conjugating enzymes, namely E2. E2 proteins catalyze the substrate ubiquitination, either alone, or in conjunction with ubiquitin-protein ligase, E3. Typically, the substrate has an assembly of ubiquitin chains. The substrate is then transferred to proteosomes.

Degradation: A proteosome particle is a part of ATP-dependant proteolytic system that catalyzes degradation of rate-limiting enzymes, transcriptional regulators, critical regulatory proteins and misfolded proteins. The proteosome is a 26S cylindrical particle which is made up of a 20S catalytic particle and two 19S regulatory subunits. 19 S particle is made up of several subunits. The 20S is a cylinder made of four seven-subunit rings stacked on top of each other. The outer two rings are made of α subunits and the inner two are made of β subunits. The catalytic activity is located on the β subunits on the inner surface of the enclosed chamber. A set of very narrow channels (13 Å) lead to the two ends of the particle, thus allowing only unfolded peptide to enter and exit the particle.

The MBP1 subunit of 19S particle has a high affinity for poly-ubiquitin chains and serves as an anchor that holds the polypeptide on the 26 S complex. Prior to degradation of the substrate protein, the ubiquitin chain is depolymerized by a large, diverse family of Ubquitin-isopeptidases. A co-factor of 26S proteosome known as Factor-H is essential for degradation of proteins with an acetylated N-terminus. The peptidase activities of the 20S particle include a chymotrypsin-like activity which cleaves after large hydrophobic residues, a trypsin-like activity which cleaves after basic residues and a post-glutamyl hydrolase that cleaves after acidic residues. These activities are catalyzed by active sites on different β subunits in a processive fashion. Thus the proteosomes cleave the protein into peptide fragments. (Coux, Tanaka *et al.* 1996). Ubiquitin-proteosomal pathway is a major pathway by which misfolded proteins get degraded.

1.7 ER Storage Diseases:

In spite of the QC mechanisms in place within the ER, the presence of misfolded proteins often lead to a disorder because:

a) Intracellular retention/degradation of an important protein prevents it from *reaching its target site* and performing an important function

b) Misfolded proteins accumulated in the cells may *interfere with normal trafficking* of other proteins. Abnormal accumulations in this case may cause the rER to be distended.

c) Aggregated proteins may form *cytotoxic complexes* within the cell.

d) If some misfolded proteins manage to escape ER retention, they can impair some vital cellular functions.

Such diseases, caused by abnormal trafficking of misfolded proteins are categorized as ER storage diseases (ERSDs) (Kim and Arvan, 1998; Brooks, 1997). Some well-documented ERSDs include congenital hypothyroid goiter, osteogenesis imperfecta (OI), diabetes insipidus and cystic fibrosis (CF). ERSD can be characterized as ERSDs due to loss-of -function mutations or ERSD due to gain-of-negative-function mutations (conformational mutations) (Kaufman 2002)

a) **Loss-of-function mutations:** Many inherited recessive disorders are caused by mutations that disturb the productive folding of proteins. Misfolded proteins are either non-functional or are efficiently degraded. In some cases, the mutated proteins are still functional however the QC cannot distinguish these proteins from non-functional proteins and are retained nonetheless. In the *"loss of function diseases"*, the mutant proteins are degraded in the ER or are retained in soluble forms that do not affect the normal secretion of other proteins by the affected cells, therefore the absence or low levels of a protein in its physiologically relevant site would cause the disease. If the condition is heterozygous,

the normal allele is also expressed and its expression may be able to make up for the degraded mutant. These are typically inherited as recessive disorders (Kopito and Ron 2000; Kaufman 2002). Such diseases could be treated by administering therapeutic doses of the deficient protein or by gene therapy.

Cystic fibrosis is an autosomal recessive disorder caused by a mutation in a plasma membrane chloride conductance channel called CF transmembrane conductance regulator (CFTR). The most common variant is CFTR508, a Phe deletion in position 508. The structural mutant, is biologically functional, but is retained within the ER, mainly by being anchored to calnexin, and degraded by the proteosomes in the cytoplasm. (If the cells expressing the mutant are grown at lower temperatures, the CFTR508 is functionally competent, presumably because the slower rate of polypeptide folding allows the chain to obtain a more correct tertiary structure). Since CFTR is absent in its target destination, it cannot perform its function as a chloride channel. This causes respiratory and gastrointestinal symptoms in the patients.

Congenital hypothyroid goiter that is caused by thyroglobulin (Tg) deficiency is another example of the "loss-of–function" subtype of ERSD. Defective thyroglobin molecules cannot homodimerize and are retained in the ER of the thyrocytes. Since Tg plays a key role in thyroid hormone synthesis and storage, mutations in the Tg genes cause deficiency of thyroid hormone, and lead to hypothyroidism.

b) **Gain-of-negative function mutations:** On the other end of the ERSD spectrum are the, *conformational diseases* that are inherited as dominant disorders. Conformational disorders arise when a specific protein undergoes conformational rearrangement that allows it to aggregate and become deposited within tissues or cellular compartments. The

aggregates interfere with normal cellular processing, resulting in a gain of "negative function". These diseases can also be induced, as in case of prions. Unlike the loss-of-function disorders, they are not treatable by gene therapy to deliver a wild-type copy of the mutant gene.

Protein aggregation: Protein aggregation has been reviewed by (Horwich 2002). Small proteins less than150 amino acids undergo two-state folding pathway, existing only in native and unfolded states. They undergo specific kinetic steps on the way to the native state, and there may be multiple choices of intermediate conformations at any given point. Unlike the two-state proteins, folding does not proceed efficiently for proteins of more than 100-150 amino acids. There are kinetic barriers to reaching the native state. It is necessary for the proteins to acquire the conformation of "obligatory intermediate" structure to reach their native states. Some intermediates are "kinetic traps" that lead the protein away from the productive pathways. Such partly folded intermediate conformations are prone to associate and aggregate. Chaperones may prevent a polypeptide from descending into a kinetic trap or may rescue it from one. Regardless of the nature of the intermediates, when such states are significantly populated, the intermediate species interact with each other to form aggregates.

Goldberg *et al.* (London, Skrzynia *et al.* 1974) studied this phenomenon in the refolding studies on tryptophanase diluted from urea. They found that aggregation is driven by the same stereospecific interactions that are required for folding of the polypeptide chain to the native form, but these interactions occur within separate chains, i.e., intermolecular contacts replace intramolecular contacts. There is concentration dependant competition between unimolecular intrachain interactions and multimolecular interchain interactions.

The rates of the latter increase rapidly with protein concentration. Secondary structure studies on phosphoglycerate kinase (PGK) show that aggregates lack substantial α-helical structures and apparently expose core β-strands. The combination between hydrophobic and β-strand interactions constitutes the interchain aggregations in PGK. Conformational mutations in proteins lead to disordered structures with a propensity to aggregate. These unstructured proteins overwhelm the QC mechanisms that prevent aggregation, reach critical concentrations required to assemble the aggregates and cause the disease pathogenesis.

Juvenile pulmonary emphysema caused by α-1-antitrypsin-Z (ATZ) deficiency is an excellent model for conformational disorder that is also an ERSD. ATZ is a plasma serine protease secreted by the liver cells (hepatocytes). It is the major physiologic inhibitor for neutrophil elastase. It regulates the elastase-mediated degradation of elastic tissue in the lung. In cases of ATZ deficiency, there is uncontrolled elastase activity in the lungs, which results in increased elastin breakdown and development of emphysema (Khan, Salman *et al.* 2002). ATZ deficiency is caused by folding mutants that are retained within the ER of the hepatocytes and degraded by proteosomes. Depending on the efficiency of the degradation, there are two sub-types. In the first sub-type, ATZ is efficiently degraded in the hepatocellular ER; hence only the lungs are affected because of uncontrolled lung elastase activity (loss-of –function disorder). In the second sub-type, there is a lag in the degradation as the mutant is unable to interact efficiently with calnexin. This leads to retention of hepatotoxic mutant ATZ in the hepatocytes, causing chronic liver injury and hepatocellular carcinoma, in addition to emphysema (Qu, Teckman *et al.* 1996).

Type I diabetes, is a well-characterized protein folding defect. A gain-of-function C96→Y substitution in the pro-insulin 2 gene disrupts pro-insulin folding. The misfolded protein aggregates are retained in the pancreatic ER and result in diabetes phenotype. The absence of downstream signaling components in glucose metabolism ameliorates the disease pathogenesis.

Pseudoachondroaplasia, caused mainly by mutations in COMP is also an example of conformational disorder. Mutations in COMP lead to association of COMP with itself and other proteins. The aggregates form a highly regular fingerprint like structures in the chondrocytic ER. The pathology of this disorder is not clearly understood yet.

Lethal hypochondrogenesis, which is cause by mutant type II collagen, is an example of dominant inheritance of ERSDs caused by inefficiently degraded mutants. Gly853→Glu substitution, located in the sequence encoding the triple helix in the pro-α2 chain impairs the assembly and transport of the collagen molecules. Misfolded collagen molecules are retained as granular aggregates within the distended rER of the affected chondrocytes (Bogaert, Tiller *et al.* 1992). This is also an example of an ERSD affecting skeletal tissues.

Thus the severity of the ERSDs and the likelihood of reversing the effects of the disease depend on the category in which they fall, as well as the state of oligomerization of the protein involved.

Role of oligomerization in the severity of the ERDS: If a heterologus mutation occurred in a gene expressing an oligomeric protein, the normal protein expressed by the wild-type allele would assemble with protein expressed by mutant allele, forming non-functional oligomers. The normal protein would be trapped in the ER and degraded along with the

mutant. The mutation would be inherited as dominant negative if the misfolded proteins form aggregates that impair cellular functions. Typically, mutations affecting oligomeric ECM proteins are inherited as dominant mutations. The picture is further complicated if a mutant oligomer escapes ER retention; its presence would be more harmful to the organism than a null mutation or a mutant that is efficiently retained and destroyed in the ER. This is true especially in case of ECM proteins because a mutant would interfere with the formation of stable matrix interactions with other proteins even if present in small proportion. This would result in a dominantly inherited disorder.

Low levels of proteins (expressed by the wild type allele) or absence of protein due to null mutation would be more tolerable because other proteins may be able to compensate for lesser quantities/absence of the normal protein.

1.8 ERSDs caused by mutations in ECM components:

Many ECM proteins are oligomers and are involved in protein-protein interactions that provide structure and stability to the matrix. The severity of the ERSD phenotype as well as the mode of inheritance would largely depend on the fate of the retained proteins as well as the efficiency of the retention. Either all the retained proteins are completely degraded or the mutant is not degraded but is retained in a soluble form or as cytotoxic insoluble aggregates within the cell. Some of the mutant forms may also escape ER retention and be deposited in the matrix.

Osteogenesis imperfecta (OI), a heterogeneous disorder, caused by defective type I collagen, is the most studied the ERSD involving mutations in matrix proteins. It is

characterized by clinical anomalies in type I collagen containing tissues like bone, skin, ligaments, tendon, sclera and dentin. It is also known as brittle bone disease (Cole 2002). OI is the most common inherited skeletal disorder (Le Parc, Molcard *et al.* 2000). Type I OI is the most common and the mildest form of OI, characterized by bone fragility, gray-blue sclera and deafness. The most severe form is the lethal perinatal OI-II. OI-III is a severe non-lethal form of progressively deforming OI.

OI may be caused by mutations in any one of the genes expressing type I collagen chains (namely COL1A1 and COL1A2 encoding two pro-α1 (I) and one proα2 (II) pro-peptides respectively) or in any of the enzymes responsible for the post-translational processing of collagens (Cole 1997). OI is extensively studied because several variants are possible depending the stage of collagen biosynthesis that is affected by the mutations. If the mutation occurs in collagen, which is usually the case, the severity of the disorder is determined by the location of mutation within the pro-collagen. Since the triple helix formation occurs in a zipper-like manner from C-terminus to N-terminus, mutations near the C-terminal affect the structure of the triple helix more severely than the mutations near the N-terminus of the pro-peptide would.

Genotype-phenotype relations have shown that the severity of OI phenotype is worse if the mutation is expressed than if the mutation is not expressed (Cole 2002).

In OI, 50% to 75% pro-collagen molecules contain mutant pro-peptides. ER quality control retains and degrades some mutant molecules reducing the overall concentration of collagen molecules in the tissues. However, variable populations of abnormal molecules escape ER retention and are incorporated in the ECM. Here they interfere with the matrix formation of the tissue, impairing tissue structure and function. The *dual* effects of low

levels of normal collagens in the tissue and the interference in matrix formation by the secreted mutants lead to various OI phenotypes.

C: PSACH is an ERSD

As mentioned in the previous section, PSACH is also an ERSD. Mutations in a single COMP allele lead to an autosomal dominantly inherited disorder. Because COMP is a pentamer, heterozygous mutations in the COMP gene lead to any one of the possible 32 (2^5) combinations in an oligomer. Of these, only one of the combinations would contain all five normal chains, i.e., 97% of the COMP molecules contain one or more defective chains. In spite of this, it is interesting to note that COMP-null transgenic mice are normal and show no phenotype (Svensson, Aszodi *et al.* 2002). This indicates that other proteins can compensate for the functions of COMP. However, a single mutation leads to aggregation of COMP molecules, not only through self-associations, but through abnormal interactions with other proteins as well. It is know that at least one other protein- type IX collagen is retained in the chondrocyte rER (Maddox, Keene *et al.* 1997). Our hypothesis is that misfolded COMP molecules are essential to form aggregates. These aggregates impair normal cartilage development by a) disrupting the trafficking of other proteins processed in the chondrocyte rER; b) retaining in the rER other ECM proteins that are the building-blocks of the cartilage matrix; c) if some destabilized COMP molecules manage to escape rER retention, they may form abnormal matrix interactions, compromising the matrix stability. Since the integrity of the cartilage is essential for development of the long bones, COMP mutations may ultimately affect the skeletal development of a growing child.

1.1 How does a mutation in COMP affect the structure of the protein?

To date thirty-eight mutations have been identified in the type III repeats and three in the C-terminal (CT domain). Many mutations involve in-frame deletions or expansions of the aspartic acid codons or cause substitutions of aspartic acid residues. Tandem repeats of aspartic acid residues are highly conserved in the type III repeats and are thought to be involved in calcium binding. Calcium binding studies done on the type III repeats in the TSP1 protein show that calcium binds to this region in a cooperative manner and protects it from proteolysis (Lawler and Simons 1983) (Misenheimer and Mosher 1995). However very little structural information is available on the type III sequence repeats and the C-terminal domains. Nor has the importance of calcium binding on the structure or the biological function of COMP been demonstrated.

Calcium is known to help mediate matrix-matrix as well as cell-matrix interactions. It can stabilize isolated domains as well as superstructures of macromolecular assemblies. Calcium binding often leads to the formation of biologically stabilized structures that are resistant to proteolysis. A mutation that interferes with the calcium-binding properties of a protein domain will affect its stability, structure and function adversely. Mutations in the calcium-binding domains of many ECM proteins often lead to genetic disorders (Maurer and Hohenester 1997). For instance, Marfan's syndrome is linked to mutations in calcium binding regions of fibrillins (Dietz, Cutting *et al.* 1991; Maslen, Corson *et al.* 1991). Marfan's disorder is characterized by overgrowth of long bones and aortic dilation and rupture. Fibrillins are ECM proteins secreted in skeletal tissues and blood vessels and are the major components of supramolecular fiber structures called microfibrils. It was

shown that the mutations in the calcium-binding EGF modules (cbEGF) of fibrillin-1 reduce their affinity to calcium and render them susceptible to proteolysis. Some of the cleavage sites were clustered around the mutation, indicating local structural changes within the cbEGF domain whereas other cleavage sites were at a distance from the mutation and in domains other than the mutated domain. (Reinhardt, Ono *et al.* 1997; Reinhardt, Ono *et al.* 2000). This suggested a single mutation can cause long-range structural changes in the primary surface of the molecule.

We hypothesized that the structural conformation of the type III domain was sensitive to the presence of calcium. Mutations in the type III domain destabilized the structure and the calcium-binding properties of this domain and possibly those of the adjacent C-terminal domains and initiated PSACH pathology. Hence both these domains were recombinantly expressed and biophysical studies were carried out to determine their structure and their calcium binding properties.

In order to understand why the quality control mechanisms that either degrade or prevent the aggregation of misfolded proteins malfunction in the PSACH cartilage, various aspects of COMP biosynthesis were studied. A major obstacle in these studies is that the primary chondrocytes dedifferentiate in serial monolayer with respect to their morphology and phenotype. They change from round to flattened fibroblast–like cells and secrete type I collagen instead of types II and IX collagens which are the markers of the chondrocytic phenotype (Zaucke, Dinser *et al.* 2001). These cells are therefore, not ideal for studying in the *in vivo* conditions observed in the cartilage. Some studies described here were done on the patient biopsy material. But this material is rare, and available in insufficient quantities to perform all the necessary tests required to study this

disorder. Hence, I had to design systems that closely resemble the milieu that the mutant COMP molecules are exposed to in the PSACH cartilage.

First, it was necessary to compare rates of COMP processing in chondrocytes with other tissues to obtain clues on the tissue-specificity of this disorder. Studies on COMP processing have been performed on cultured chondrocytes, ligaments and tendons (Hecht, Deer *et al.* 1998; Delot, Brodie *et al.* 1998). Primary chondrocytes, isolated from bovine fetal cartilage retain their chondrocytic phenotype if kept in suspension and were used for metabolic labeling of COMP.

Second, it was necessary to identify the proteins that interacted with mutant COMP. *Immunofluorescence studies and electron microscopy* on biopsy material obtained from a PSACH patient were performed to identify the chaperones localized to the largely distended rER within the chondrocytes (The patient was first screened for PSACH mutation). In order to understand whether the PSACH pathology is due to absence of important proteins in the ECM, the tissues were also screened for other ECM molecules. Since these studies were limited to the antibodies used to screen the biopsy material they could not give information on the early stages of aggregate formations nor detect all protein-protein interactions in the aggregates.

Hence affinity chromatography was performed to further identify the proteins that interacted with mutant COMP. Since most mutations occur in type III domain of COMP, wild type and D446→N mutant peptides comprising the type III domain were recombinantly expressed and affinity columns were used to detect interactions between them and other chondrocytic rER proteins. The goal of the affinity study was to identify not only known chaperones that retain mutant COMP molecules within the cell, but also

to identify novel chondrocyte-specific proteins that are involved in tissue-specific retention of COMP.

Significance: In brief summary, these studies provide information about a new protein domain and also describe the influence of a mutation on the structure of this domain. Additionally, the effect of this mutation on the adjacent C-terminal domain provides more information on the pathological consequences of PSACH. Biosynthesis studies enhanced our understanding as to why this multidomain protein is retained in the chondrocytes and not degraded. Insight was also provided on the tissue–specificity of the COMP retention. Overall, these studies allowed us to understand the pathology of PSACH disease process.

Footnotes

List of Abbreviations:
Asn: Asparagine
Asp: Aspartic acid
ATF: Activating transcription factor
CbEGF: Calcium binding EGF
CFTR: Cystic fibrosis transmembrane conductance regulator
COMP: Cartilage oligomeric matrix protein
CRD: Carbohydrate recognition domain;
CyP: Cyclophilin
Cys: cysteine
ECM: Extracellular matrix
EGF: Epidermal growth factor
ER: Endoplasmic reticulum
ERAD: ER associated degradation
ERGIC: ER-Golgi intermediate compartment
ERSD: ER storage disease
FKBP 65: FK506 binding protein 65
Glc: Glucose
GlcNAC: *N*-acetylglucosamine
Gly: glycine
GPI: Glycophosphatidyl inositol
GRP: Glucose-regulated protein
HIP: HSC interacting protein
HSC; Heat shock chaperone

HSP: Heat shock protein
Man: Mannose
MED: Multiple epiphyseal dysplasia
NAC: nascent polypeptide-associated complex
OI: Osteogenesis imperfecta
PDI: Protein disulfide isomerase
Pma 1: Plasma membrane ATPase 1
PSACH: Pseudoachondroplasia
RAP:
rER: Rough ER
Ser: Serine
SR: Sarcoplasmic reticulum
SRP: Signal recognition particle
TGF β: Transforming growth factor β
TRAM: translocating chain-associated membrane protein
Thr: Threonine
TSP: Thrombospondin
UGGT: UDP-glucose: glycoprotein galactosyl transferase
UPR: Unfolded protein response
UPRE: Unfolded protein response element
VTC: Vesiculotubular clusters

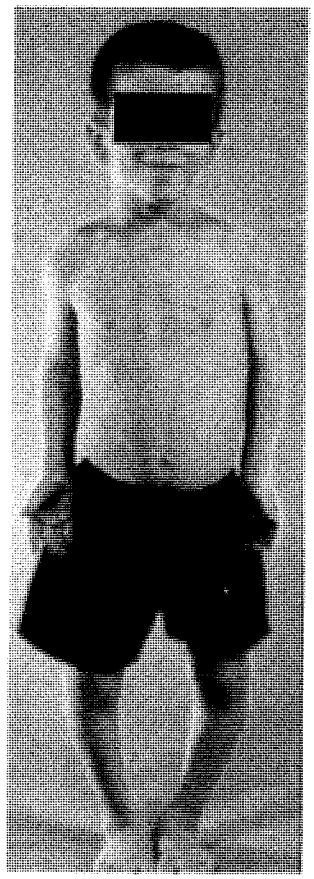
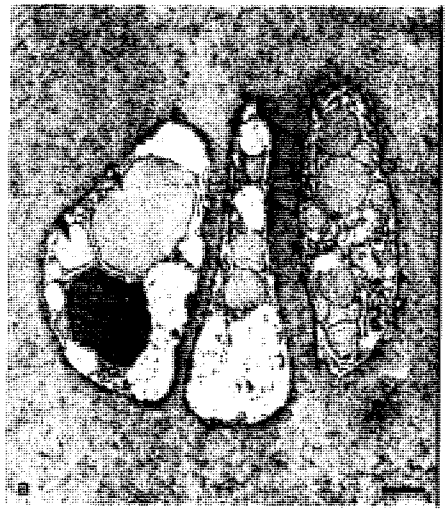


Fig1: A nine year old child with classic PSACH symptoms- short structure, bow legs and disproportionately long hands.



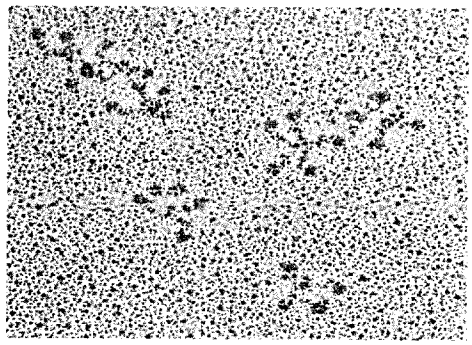


Fig.2 (top): Enlarged chondrocytes in cartilage of PSACH patient. The rER is enlarged and occupies a disproportionately large volume of the cell. Bottom (left): Fingerprint like organized structures retained in the abnormal rER. Bottom (right): Pentameric bouquet-shaped COMP molecules that are localized to the dark bands of the aggregates. Reproduced form (Maddox, Keene et al. 1997).

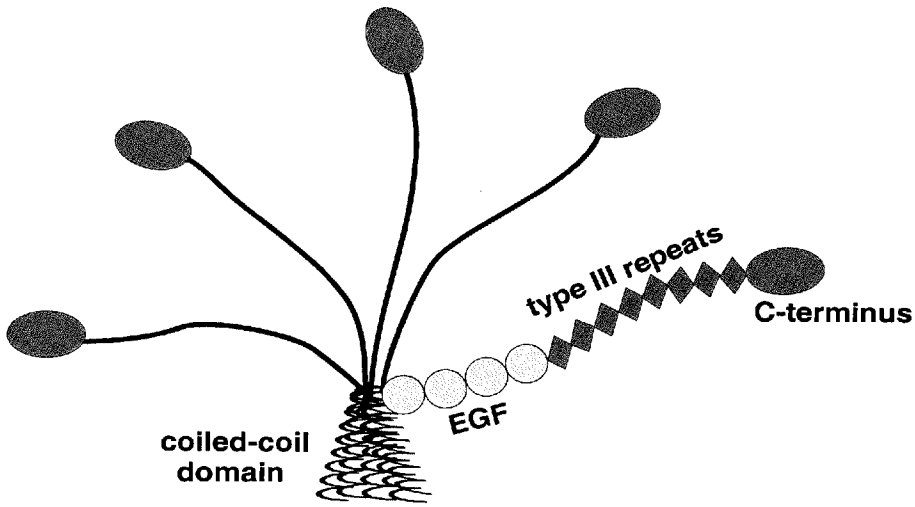


Figure 3: Modular nature of COMP pentamer: COMP pentamer consists of five polypeptides chains, each made of one N-terminal oligomerization domain, four EGF-like repeats, eight thrombospondin type III sequence repeats and the globular C-terminal domain.

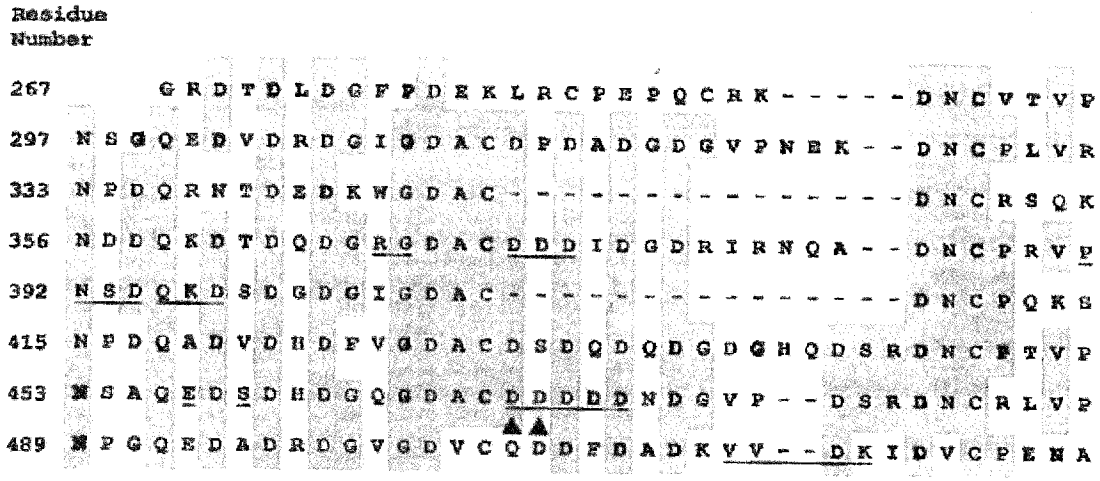


Figure 4: The alignment of the type III sequence repeats in COMP: Conserved residues are blocked in gray and the residue numbers at the start of the repeats are shown on the left. Amino acids residues that have been substituted (missense mutations) are shown in bold and the regions where there have been in-frame deletions have been underlined. Reproduced from (Briggs and Chapman, 2002).

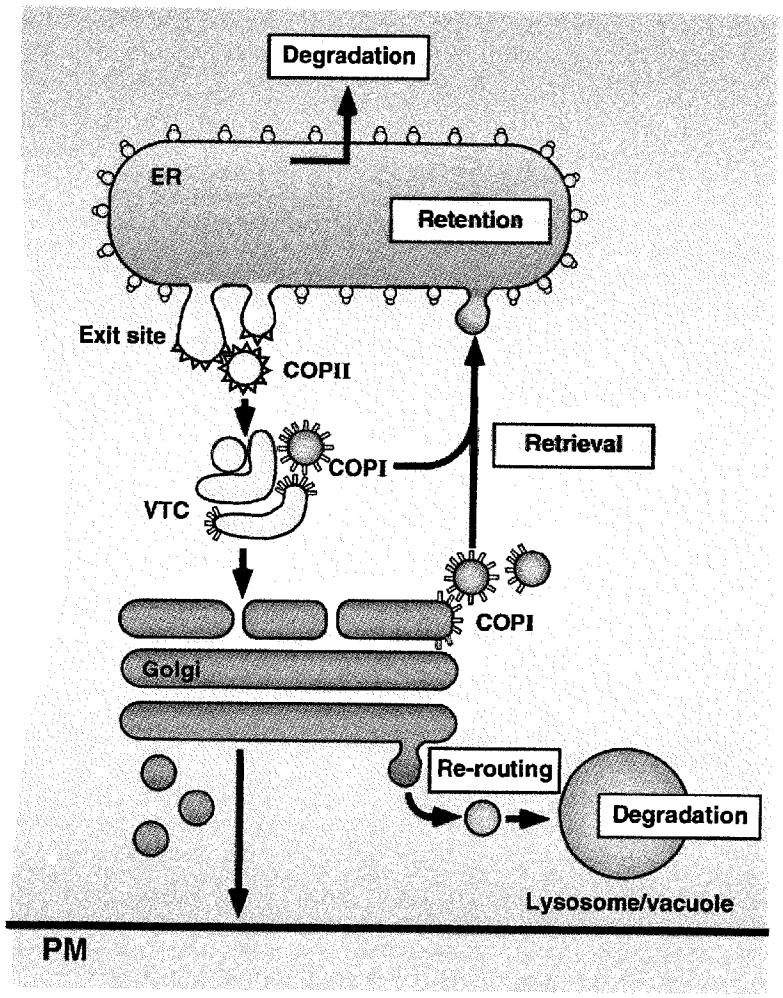


Figure 5: Intracellular routes followed by proteins in the secretory pathways: Newly synthesized proteins enter the secretory pathway in the endoplasminc reticulum (ER) by translocation. They undergo post-translational modifications in the ER. Correctly folded proteins move to the Golgi complex and the trans-Golgi network through VTC (vesiculotubular compartment) and COP coated vesicles. They are targeted to their final destination i.e., lysomes, vaculoes or to the plasma membrane or out of the cell through secretory granules. Misfolded proteins are retained in the rER or retrieved from the Golgi complex and targeted for degradation. Reproduced from (Wickner, Maurizi *et al.* 1999).

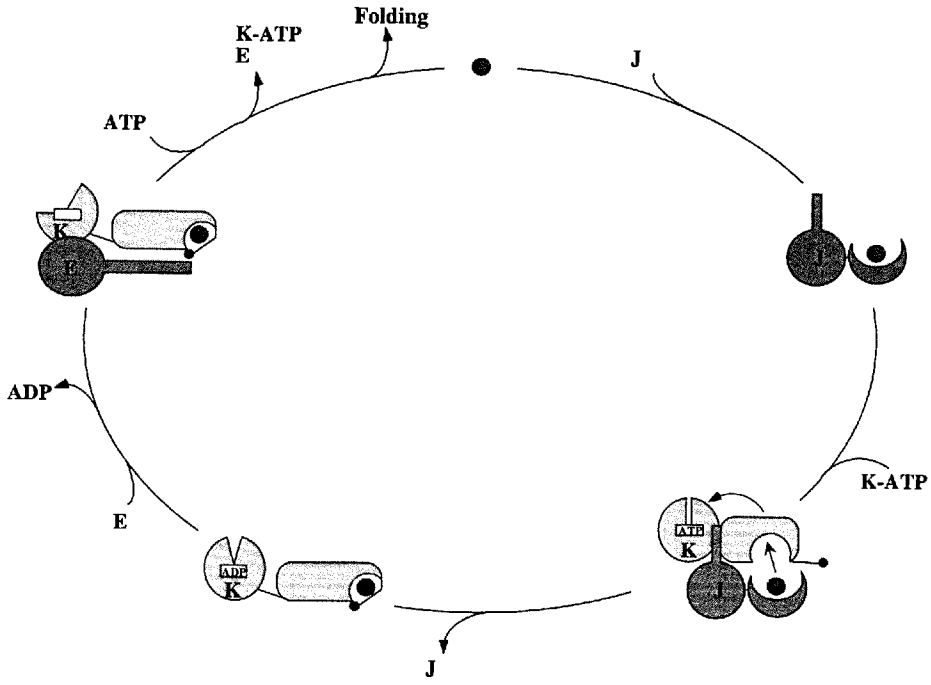


Figure 6: Model of the chaperone cycle from the Dna K system (bacterial cytoplasmic homolog of HSP70): The cycle starts with the association of DnaJ (bacterial anolog of mammalian HSP40 co-chaperone) denoted as J with the substrate (closed circle), followed by the transfer of the substrate to the ATP form of DnaK (K). This transfer is coupled to the locking-in of the substrate in the substrate-binding pocket of the DnaK by ATP hydrolysis. Following the substrate transfer, DnaJ leaves the complex and GrpE (E) associates with the DnaK –substrate complex to trigger ADP release from the DnaK. This allows binding of ATP and subsequent release of the substrate and GrpE from DnaK. In mammals, the action of mammalian BAG-1 is similar to GrpE. Reproduced from (Bukau and Horwich, 1998).

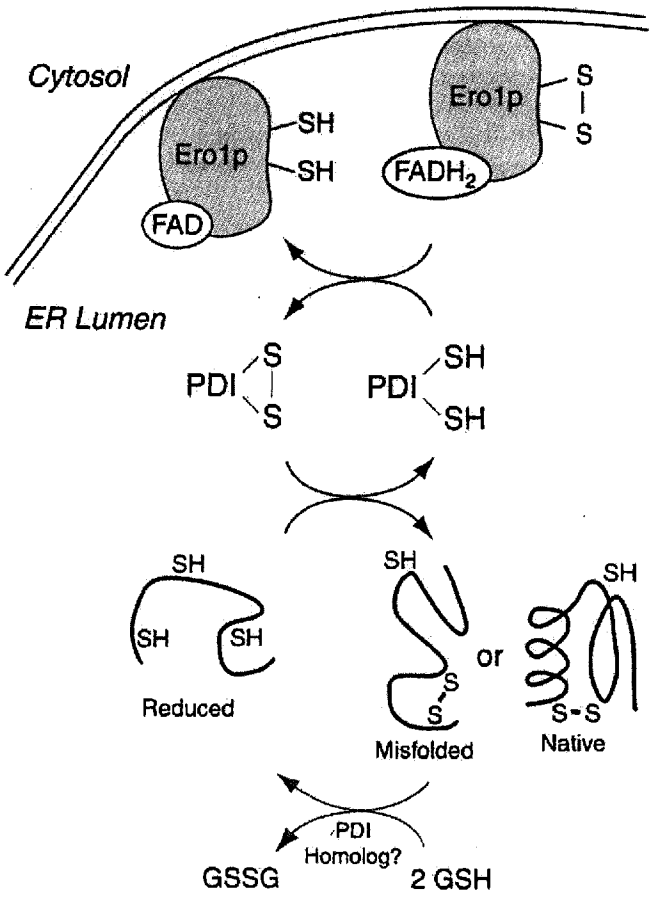


Figure 7: Role of Protein disulfide iomerase (PDI) in protein folding: PDI introduces disulfide bonds in nascent proteins with the help of Ero1 p. FAD acts as a co-factor in generating reducing equivalents on Ero1 p. These oxidizing equivalents are passed to PDI, which can then oxidize the substrate. An unknown PDI homolog may be involved in reducing the incorrect disulfide bonds. Reproduced from (Fewell, Travers *et al.* 2001).

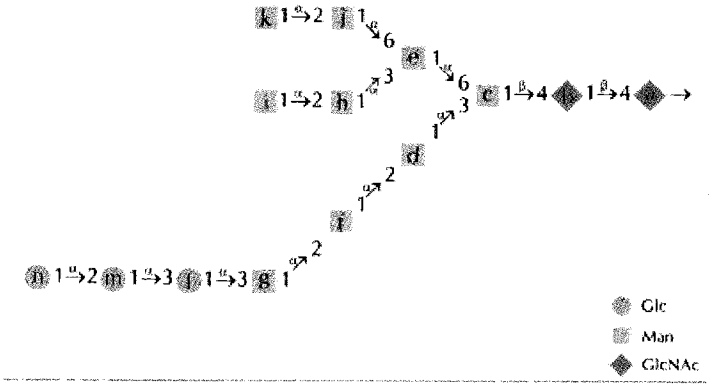


Figure 8: The Structure of the oligosaccharide transferred to the proteins in the mammalian cells. a, b, c etc. indicate the order of addition of monosaccharide units for *in vivo* synthesis of $Glc_3Man_9GlcNAc_2$-P-P-dolichol. Glc, Man and GlcNAc are abbreviations for glucose, mannose and *N*-acetylglucosamine respectively. Reproduced from (Parodi 2000).

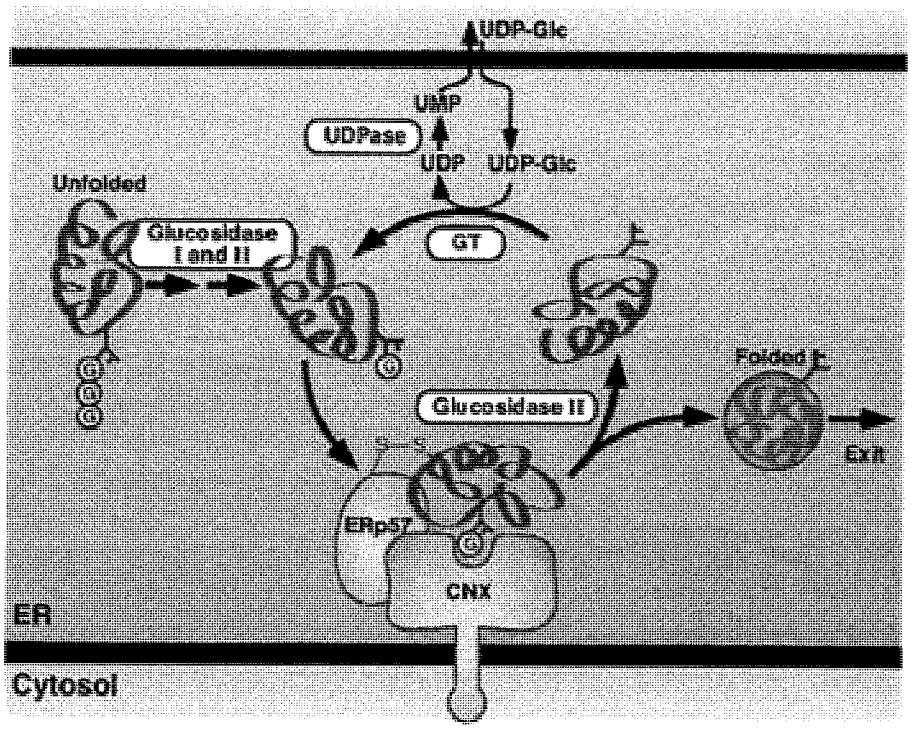


Figure 9: Calnexin cycle in the rER: Nascent glycoproteins contain a branched N-linked oligosaccharide, with one branch ending in three glucose residues. Proteins with a single glucose residue attached to the glycan bind to calnexin (CNX) or calreticulin. UGGT (GT) specifically recognizes misfolded proteins lacking the last glucose residue and reglucosylate them. Calnexin rebinds to the monoglucosylated proteins and retains them in the ER. Reproduced from (Ellgaard, Molinari *et al.* 1999).

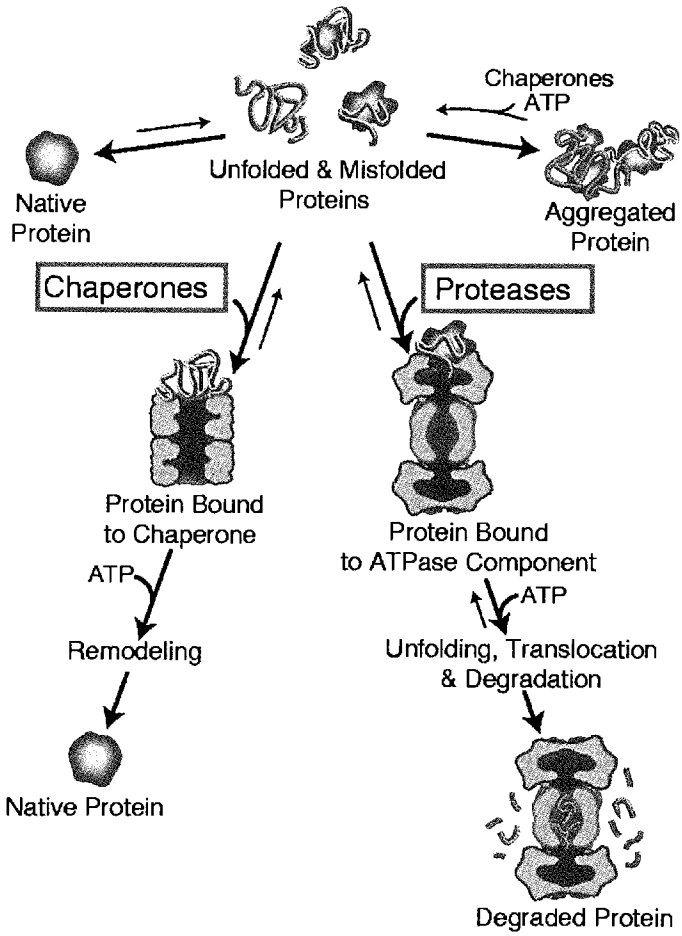


Figure 10: Protein triage model for quality control: rER chaperones try to refold the misfolded proteins to their native form. Some misfolded proteins may be degraded by the ubiquitin-proteosome pathway or by rER proteases. However, some proteins are irreversibly folded and form insoluble aggregates that are deposited win the cell. Reproduced from (Bonifacino 1996).

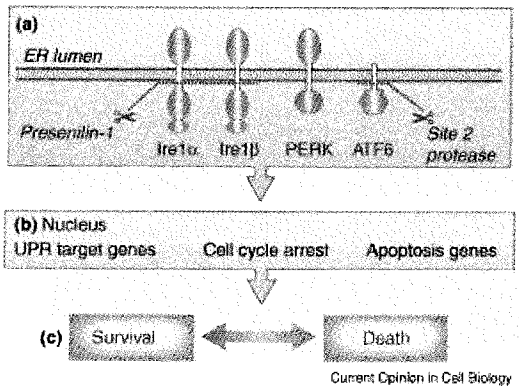


Figure11: Diversity in the mammalian unfolded protein response: (a) Response to ER stress in mammalian cells is mediated by several transmembrane kinases like Ire1α, Ire1β and PERK as well as transcription factors like ATF6. (b) UPR induces diverse effects including induction of folding chaperones, cell cycle arrest and apoptosis (Patil and Walter, 2001).

Chapter 2

J Biol Chem, Vol. 275, Issue 15, 11412-11417, April 14, 2000

A Cartilage Oligomeric Matrix Protein Mutation Associated with Pseudoachondroplasia Changes the Structural and Functional Properties of the Type 3 Domain*

B. Kerry Maddox[§], Asawari Mokashi[¶], Douglas R. Keene, and Hans Peter Bächinger

From the Research Department, Shriners Hospital for Children, the [§] Department of Oral and Molecular Biology, and the [¶] Department of Biochemistry and Molecular Biology, Oregon Health Sciences University, Portland, Oregon, 97201

The experimental procedures were carried out as follows: AM: Peptide purification and secondary structure analysis of the peptides.
BKM- recombinant expression of COMP-3 peptides, calcium binding by equilibrium dialysis, peptide purification.
DK: Rotary shadowing of the COMP peptides.

ABSTRACT

Cartilage oligomeric matrix protein (COMP) is a member of the thrombospondin family of extracellular matrix glycoproteins. All members of the family contain a highly conserved region of thrombospondin type 3 sequence repeats that bind calcium. A mutation in COMP previously identified in a patient with pseudoachondroplasia resulted in abnormal sequestration of COMP in distinctive rER vesicles. The mutation, Asp-446 Asn, is located in the type 3 repeats of the molecule. This region was expressed in a mammalian culture with and without the mutation to study the structural or functional properties associated with the mutation. The biophysical parameters of the mutant peptide were compared with those of the wild type and revealed the following difference: secondary structural analysis by circular dichroism showed more α-helix content in the wild-type peptides. The calcium binding properties of the two peptides were significantly different; there were 17 calcium ions bound/wild-type COMP3 peptide compared with 8/mutant peptide. In addition, wild-type COMP3 had a higher affinity for calcium and bound calcium more cooperatively. Calcium bound by the wild-type peptide was reflected in a structural change as indicted by velocity sedimentation. Thus, the effect of the COMP mutation appears to profoundly alter the calcium binding properties and may account for the difference observed in the structure of the type 3 domain. Furthermore, the highly cooperative binding of calcium to COMP3 suggests that these type 3 sequence repeats form a single protein domain, the thrombospondin type 3 domain.

INTRODUCTION

Cartilage oligomeric matrix protein (COMP)[1] is a member of the thrombospondin (TSP) family of extracellular glycoproteins (Oldberg, Antonsson *et al.* 1992). Like the other four members of the gene family, COMP is an oligomer composed of identical subunits (Bornstein and Sage 1994). The chains associate at the N terminus through a coiled-coil domain. Assembly of the oligomer requires the formation of an -helical coiled-coil domain that is later stabilized by interchain disulfide bonds (Malashkevich, Kammerer *et al.* 1996). Electron microscopy shows the COMP molecule to be composed of five flexible arms with a large globular domain at the end of each arm and often described as being bouquet-like in its appearance (Morgelin, Engel *et al.* 1992). TSP 3 and 4 also form pentamers (Morgelin, 1992; Qabar, 1995 and Lawler, McHenry *et al.* 1995)), whereas TSP 1 and 2 are trimeric molecules (Bornstein and Sage 1994).

The TSP family members have in common the domain structure typified by the oligomerization domain, three or four epidermal growth factor domains (EGF), and an invariant number of thrombospondin type 3 sequence repeats considered as seven or eight domains, the number depending on how the amino acid sequence is aligned (Bornstein and Sage 1994). A comparison of TSP 1, 2, 3, and 4 with COMP sequence reveals a high degree of conservation in the type 3 repeats. The temporal and spatial expression pattern of the TSP family members is distinctive, although somewhat complementary and occasionally overlapping (Tucker, Adams *et al.* 1995). COMP was originally discovered in cartilage extracts and has been immunolocalized to developing as well as mature

cartilage (Fife and Brandt 1984; Hedbom, Antonsson *et al.* 1992). COMP has also been found within and around the tendon (DiCesare, Morgelin *et al.* 1994).

There are no known human pathologies associated with the thrombospondin family members with the exception of COMP. In the last few years COMP mutations have been identified in two dominantly inherited skeletal disorders: pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (MED). These osteochondrodysplasias result from base substitutions, deletions, or additions in the COMP gene. PSACH is characterized by a disproportionate short stature and joint laxity. Manifestations of PSACH usually first appear at the onset of walking with a distinctive waddling gait. MED patients generally present with debilitating hip pain in childhood. Short stature maybe observed later in childhood. The radiological abnormalities of MED are restricted to the epiphyses, whereas there is additional involvement of the metaphyses and the spinal column in patients with PSACH. Patients with both disorders manifest symptoms of precocious osteoarthritis.

Another extracellular matrix component of cartilage, type IX collagen, has also been implicated in MED. This form of MED (EDM2) was mapped to chromosome 1p32, and the locus was identified as the COL9A2. Mutations have been identified in COL9A2 (Muragaki, Mariman *et al.* 1996). A third MED locus has recently been mapped to COL93A, and a mutation was identified (Paassilta, Lohiniva *et al.* 1999). There seems to be a growing consensus that the phenotypes of the MED patients with known type IX collagen mutations are less severe than those with COMP mutations (Paassilta, Lohiniva *et al.* 1999). However, there is a variation in the severity of symptoms within families of a known single mutation creating a complex genetic or physiological explanation. Unlike

MED that can result from mutations in one of three genes, PSACH is only associated with mutations in the COMP gene.

The mutations in COMP associated with PSACH and MED predominantly occur within the thrombospondin type 3 repeats, the region proposed to bind calcium (Cohn, Briggs *et al.* 1996; Deere, Sanford *et al.* 1998). The presumed importance of calcium binding to the biological function of COMP is largely based on studies of TSP 1 demonstrating that calcium binding occurs in a cooperative manner and protects against proteolysis (Lawler and Simons 1983; Misenheimer and Mosher 1995). No structural information is known about these type 3 repeats. This region contains 17 cysteines and a large number of acidic residues; it also contains 13 putative EF hands. The structural conformation of these repeats therefore is expected to be highly sensitive to the presence of calcium (Lawler, Chao *et al.* 1982). To determine the effect of the Asp-446 Asn mutation on these sequence repeats, we have expressed wild-type and mutant peptides in a mammalian cell culture system.

Earlier biochemical methods to isolate COMP from cartilage have demonstrated the importance of cationic interactions. COMP was selectively extracted using the cationic chelator EDTA (Morgelin, Engel *et al.* 1992). This suggests that COMP is secured in the cartilage matrix through cationic interactions and further demonstrates the potential importance of calcium binding and the effect of a mutation at one of these sites. A second cation, zinc, has been proposed to promote COMP interactions with collagen types I and II (Rosenberg, Olsson *et al.* 1998). This interaction occurs between the collagen triple helical domains and the globular C terminus of COMP. The interactions of these two cartilage matrix components, COMP and type II collagen, suggest a crucial biological

function of COMP. In comparison, there is very little information on ligands that form molecular interactions with the type 3 repeats of COMP.

The fate of a COMP molecule in PSACH has been established for the Asp-446 Asn mutation: COMP is sequestered inside the rER vesicles (Maddox, Keene *et al.* 1997). It was also determined in this study that type IX collagen was found in these vesicles as well. The stored material is organized into a lamellar pattern of alternating electron dense and lighter layers with COMP and type IX localizing to the more dense layers (Maddox, Keene *et al.* 1997). Earlier studies of aggrecan have also shown it to be present in the enlarged rER vesicles (Stanescu, Stanescu *et al.* 1982; Stanescu, Stanescu *et al.* 1984). In contrast to these cartilage macromolecules, type II collagen is secreted and assembled into the extracellular matrix of PSACH cartilage (Maddox, Keene *et al.* 1997). The secretory pathways for aggrecan and type II collagen differ (Vertel, Velasco *et al.* 1989), and presumably the pathway for aggrecan is shared by COMP and type IX collagen.

These investigations present a complex scenario where mutations in COMP cause a gross accumulation of cartilage-specific macromolecules. Assuming that the wild-type allele is transcribed and translated at approximately the same rate as the one containing the mutation, then probability suggests that 97% of the molecules will contain one or more mutant polypeptide chains. However, it is not known if the mutant chains fold properly and are subsequently assembled into COMP pentamers. Also unknown is whether the mutant molecules interact with other cartilage components retained in the rER (*i.e.* aggrecan and type IX collagen) in a specific manner or if the secretion of the aberrant molecule is retarded, potentially causing nonspecific aggregation. In this report, we describe the specific structural consequences of the mutation in the thrombospondin type

3 domain.

Experimental Proceedures

Transgene Expression of COMP Peptides in 293 Cells-- Primary chondrocytes were isolated as described previously from PSACH patient cartilage (Maddox, Keene *et al.* 1997) and cultured for up to eight passages. The region containing the TSP type 3 repeats was amplified following reverse transcription of total RNA (Life Technologies, Inc., Superscript Preamplification Kit) using a forward primer that contained an *Nhe*I cleavage site (5'-agtagctagctggtcgcgacactgacctagac-3'), whereas the reverse primer contained an *Not*I site (3'-ccpgagaacgctgaagtcacgtaggcggccgcatttgata-5').

The resulting polymerase chain reaction bands were gel-purified and ligated into the pCRII vector (Invitrogen). One wild-type clone and one carrying the PSACH mutation were selected following cDNA sequence analysis of the plasmids. These COMP3 cDNA fragments were excised from the pCRII sequencing vector with *Nhe*I and *Not*I and ligated into the respective cloning site in the expression vector pCEPSP-rF17H (Reinhardt, Ono *et al.* 1997). This vector was derived from plasmid pCEP4 (Invitrogen) and engineered to contain the signal sequence of BM40 for secretion of the peptide into the extracellular medium (Mayer, Poschl *et al.* 1995). The pCEP-COMP3 expression vectors were cultured in DH52 cells and DNA-isolated under sterile conditions. For stable episomal expression, human embryonic kidney cells (293-EBNA cells, Invitrogen) were transfected with the pCEP-COMP3 constructs in the presence of calcium chloride (Reinhardt, Keene *et al.* 1996). Selection antibiotics Geneticin (Life Technologies, Inc.,

G418) and hygromycin-B (Calbiochem) were added to the cell culture medium 48 h after transfection at a concentration of 500 μg/ml. The concentration of antibiotics was reduced to 250 μg/ml following an additional 48 h. Serum-free medium was collected and tested for the presence of the COMP3 peptides using SDS-polyacrylamide gel electrophoresis. The gel band was identified as COMP3 by N-terminal sequencing following transfer to a polyvinylidene difluoride membrane (the first ten residues were APLAGRDTDLD). The COMP3 peptide starts with APLA, the C terminus of the signal sequence; the remainder of which is removed following processing and secretion.

Purification of the COMP Peptides-- Medium from the cultured cells was collected and fractionated with ammonium sulfate. The COMP3 peptides remained in the 60% ammonium sulfate supernatant following precipitation and centrifugation. After dialysis into 20 mM Tris/HCl buffer, pH 8.0, the supernatant was applied to a Q-Sepharose ion-exchange column (Amersham Pharmacia Biotech) and eluted with a NaCl gradient from 0 to 1 M in the same buffer. Pooled fractions containing the COMP3 peptides were further separated by molecular sieve chromatography on a Superose-12 column (Amersham Pharmacia Biotech) in 50 mM Tris/HCl buffer, pH 7.4, containing 0.15 M NaCl.

Structural Analysis of the COMP3 Peptides-- Circular dichroism (CD) spectra of the COMP3 peptide in 10 mM MOPSO, pH 6.8, was analyzed in the presence and absence of 5 mM calcium chloride at 5 °C. The spectra was recorded between 260 and 180 nm on a Jasco J-500A or Aviv 202 spectropolarimeter using thermostatted cells for temperature. Secondary structural analysis was determined using a variable selection method (Compton and Johnson 1986; Compton, Mathews *et al.* 1987). Tertiary structure was

analyzed by velocity sedimentation in a sucrose gradient (50 mM Tris, pH 7.5, containing 150 mM NaCl and 5-20% sucrose). Peptide solutions at concentrations of 10 μM were prepared in the presence of either 5 mM $CaCl_2$ or 5 mM EDTA tetrasodium. The centrifugation was performed for 20 h at 56,000 rpm at 4 °C in a Beckman L80M ultracentrifuge using a SW 60Ti rotor. Sedimentation coefficients were calculated as described previously (Fessler and Fessler 1974). The molecular weight of the COMP3 peptide was assessed by equilibrium sedimentation using a Spinco Model E centrifuge (Beckman) at 17,000 rpm at 20 °C in double sector cells (Zeng, MacDonald *et al.* 1998 and Reinhardt, 1997 #29).

RESULTS

The COMP molecule is composed of five identical polypeptide chains and within each chain is a linear arrangement of domains or repeated domains (Fig.1A). The Asp-446 Asn mutation previously characterized in a PSACH patient (Maddox, Keene *et al.* 1997) occurred in the type 3 repeats where the majority of COMP mutations associated with skeletal pathologies also occur (Deere, Sanford *et al.* 1998; Loughlin, Irven *et al.* 1998; Deere, Sanford *et al.* 1999). Thus, this region was selected to be expressed in a recombinant cell culture system and analyzed.

Expression of Recombinant COMP3 Peptides-- The COMP3-Wt and -Mu constructs were derived from RNA isolated from a PSACH patient with a g a base substitution at position 1361 of the coding sequence that resulted in the substitution of asparagine for aspartic acid (Maddox, Keene *et al.* 1997). Utilizing an expression system previously

described for other secreted extracellular matrix molecules (Mayer, Poschl *et al.* 1995; Reinhardt, Keene *et al.* 1996), COMP3 peptides were expressed in human embryonic kidney cells (293-EBNA cells). The first step of protein purification was ammonium sulfate fractionation of the cell culture medium followed by column chromatography. The COMP3 peptides remained in solution at 60% saturation (Fig. 1*B*). Following dialysis of the ammonium sulfate supernatant, the peptides were further purified by cation exchange and size exclusion chromatography (Fig.2, *A* and *B*).

Rotary Shadowed Images of COMP3 Peptides-- Rotary shadowed images of COMP molecules indicate that the arms are flexible rods that connect the "stalk" of the bouquet (the coiled-coil domain) to the globular domain (Morgelin, Engel *et al.* 1992). These arms are composed of 4 EGF-like and type 3 repeats. It is presumed that EGF-like repeats form small globular domains that are aligned linearly from the solution structural determination of a pair of EGF-like domains (Downing, Knott *et al.* 1996). It is not known what the type 3 repeat structures are or how they are organized relative to one another. From the data presented here (Fig. 3*A*), COMP3-Wt peptides appear to be short rods with an average length of 14.2 nm ± 1.68 ($n = 28$). This measurement is consistent with that based on the rotary shadowed images of native COMP molecules. The reported distance between the coiled-coil and the middle of the globular domain of a full-length COMP subunit is 28.6 nm (Morgelin, Engel *et al.* 1992). The 4 EGF-like domains would be ~12 nm based on the structural measurements of one EGF-like domain in human factor IX (Baron, Norman *et al.* 1992). This would leave 16.4 nm for the TSP type 3 repeats and half of the terminal globule based on these reported molecular measurements. Thus, the

length of the COMP3-Wt peptide at 14.2 nm is structurally consistent with the native wild-type COMP domains.

The COMP3-Mu peptides were shorter with an average length of 10.3 nm ± 1.68 (n = 35; Fig. 3*B*). Whereas the size of these peptides is small and near the limit of resolution of the rotary shadow technique, the measured differences between COMP3-Wt and -Mu are supported by additional measurable variations between the two species of peptides. It is likely that the mutation affected the overall structure of COMP3-Mu.

Structural Analyses of the COMP3 Peptides-- Differences in secondary structure were observed in circular dichroism spectra between peptide populations in the presence and absence of calcium (Fig. 4*A*). Analysis of the spectra indicates that the differences were notable in the loss of helical structure and an increase in the antiparallel -sheets in COMP3-Mu compared with COMP3-Wt peptides (see TableI). The change in the secondary structure was directly coupled to the presence of calcium. Titration experiments with calcium showed a cooperative structural change for both peptides (Fig. 4*B*). However these experiments did not allow the determination of the number of binding sites; therefore we measured calcium binding by equilibrium dialysis (see below).

Molecular Weight Determination of COMP3 Peptides-- The molecular weights of COMP3-Wt and -Mu peptides were measured by sedimentation equilibrium in a model E analytical ultracentrifuge. The solutions showed monodisperse species with molecular masses of 31,260 ± 170 Da for COMP3-Wt and 30,780 ± 100 Da COMP3-Mu (data not shown). The partial specific volume was calculated from the amino acid composition and

found to be 0.678 cm^3/g (Perkins 1986), and the degree of hydration was estimated to be 0.3 g/g. A comparison of the measured molecular weights with the molecular mass determined from the amino acid sequence (28,632 Da) indicates that the COMP3 peptides have little or no post-translational modifications (such as glycosylation). The slightly higher molecular weight of the COMP3-Wt peptides could be a reflection of more bound calcium (see below).

Sedimentation Velocity Experiments in Sucrose Gradients-- The COMP3 peptides were analyzed for shape by centrifugation in 5-20% sucrose gradients in the presence and absence of calcium (Fig.5). Sedimentation coefficients $s_{20,w}$ in the absence of calcium were 2.5 ± 0.1 S and 2.7 ± 0.1 S for the wild-type and mutant peptide, respectively. In the presence of calcium the sedimentation coefficient of the wild-type peptide changed to 2.8 ± 0.1 S, whereas the mutant peptide showed no change in $s_{20,w}$ value. The increase in $s_{20,w}$ of 2.5 to 2.8 for COMP3-Wt peptides indicates a significant change in the frictional ratio upon binding of calcium. This would be consistent with an earlier study of thrombospondin 1 that demonstrated an elongation of the subunit "arms" in the presence of EDTA (Lawler, Derick *et al.* 1985). Using the determined sedimentation coefficient and molecular weight of COMP3-Wt and assuming the degree of hydration is 0.3 ml/g, the frictional ratio is 1.38 in the presence of calcium. This frictional ratio indicates that the hydrodynamic dimensions of the wild-type COMP3 peptide are 15.0 × 2.1 nm for a prolate ellipsoid or 13.4 × 1.8 nm for a rod. The measured length of the rotary shadowed image of the wild-type peptide, 14.2 nm, is consistent with these calculations.

Calcium Binding by Equilibrium Dialysis-- One important but speculative function of the type 3 domains in the thrombospondin family is to bind calcium. COMP3 peptide solutions were dialyzed against concentrations of calcium chloride ranging from 50 μM to 2 mM. A small amount of radiolabeled calcium was added for quantitative purposes. The free and protein-bound calcium concentrations were calculated for several experiments using protein concentrations that varied from 1.6 to 10.3 μM (Fig.6). The resulting plot indicated that the number of bound calcium molecules/molecule was much lower for the COMP3-Mu peptides (8 compared with 17 for COMP3-Wt). The Hill coefficient, a measure of the cooperativity of the binding curve, showed COMP3-Wt bound calcium in a more cooperative manner (3.74 *versus* 1.57 for mutant COMP3). Also, the concentration at half saturation was 102 and 218 μM for wild-type and mutant peptide, respectively.

Discussion

Data presented here show the first measurable consequence that relates a mutation in COMP to a molecular mechanism involved in generating the PSACH phenotype. This consequence is an altered potential for mutant COMP peptides to bind calcium. The secondary and tertiary structures are quantitatively affected by the reduced calcium binding. Most mutations in COMP are clustered within the type 3 repeats in the skeletal dysplasias PSACH and MED; presumably other COMP mutations alter molecular structure in a similar way.

The calcium binding data indicate a very complex scenario with a large number of potential binding sites and a highly cooperative mechanism for binding. The primary structure of the type 3 repeats conforms to the consensus sequence of the EF-hand motif,

a common structure associated with calcium-binding proteins. Sites contain a stretch of amino acids represented by X, Y, Z, -X, -Y, -Z that provide coordinating oxygens for calcium-ligand interactions in a distinctive geometry. There are up to 13 such sites in the type 3 repeats that share homology with the EF-hand structure (Lawler and Simons 1983) compared with 17 calcium ions shown to bind the COMP3 peptides here. The difference suggests that the COMP3 peptide forms a more complex structure than that represented by linear repeated sequences.

The 17 potential binding sites determined for the COMP3 peptides are also higher than that previously calculated for TSP I where 35 ± 3 ions of calcium bound one trimeric TSP I molecule (Misenheimer and Mosher 1995), which is approximately 12 calcium ions per polypeptide. One of these 12 is assumed to be bound to the EGF-like repeat with the consensus sequence of a calcium binding EGF domain; the COMP domain structure predicts two calcium binding EGF repeats based on residues crucial for ligand binding (Handford, Mayhew *et al.* 1991). These differences may be because of slight variations in design of binding assays such as ligand concentration (Misenheimer and Mosher 1995) or the small sequence differences between COMP and TSP 1.

In addition to the 13 EF-hand-repeating motif structure, the type 3 repeats have also been organized into 8 repeating units called 3A and 3B with the asparatic acid and cysteine residues aligned (Lawler and Hynes 1986). This alignment suggests that the cysteine residues may form intradomain disulfide bonds, thus stabilizing the eight repeated units. However, the disulfide bonds have not been established and it is difficult to see how more than eight calcium ions would bind with this domain structure. The structural organization of the type 3 repeats is commonly viewed as a series of separately folded

motifs along the polypeptide chain, although no structural data have been presented to support or dispute this presumption. The highly cooperative calcium binding data presented here strongly suggest that these sequence repeats form a single protein domain. This is also supported by the sequence conservation within all five thrombospondin family members. If the peptide was composed of independently folded domains, one could expect more variations within these repeated sequences. Therefore we propose that the type 3 sequence repeats form a single protein domain: the TSP type 3 domain.

It is interesting to note that the mutant peptide bound less than half the normal amount of calcium and yet theoretically only one potential binding site would be directly affected. This suggests that other factors are involved. Calcium binding to TSP I occurs by a cooperative mechanism suggesting that there is a conformational change upon an initial binding event that effects the conformation of other sites. Thus, the impact of an amino acid mutation at one calcium binding site could greatly alter binding at other sites.

The reduced calcium binding in the mutant COMP peptide has significant biological relevance in view of the effect of COMP mutations on skeletal development. Following translation of a COMP polypeptide, the chains normally assemble into a mature pentameric COMP molecule and are trafficked through the Golgi and secreted. The sequestration of COMP by PSACH chondrocytes may potentially be the result of quality control mechanisms within the rER that retain incorrectly folded proteins. Unsecreted COMP would thus accumulate in the rER and possibly form aggregates that include COMP and other cartilage components such as aggrecan and type IX collagen. Both of these molecules have been shown to be retained intracellularly in PSACH chondrocytes (Stanescu, Stanescu *et al.* 1982; Maddox, Keene *et al.* 1997). In another calcium binding

molecule, fibrillin, calcium binding maintains the tertiary structure of two adjacent EGF-like domains (Downing, Knott *et al.* 1996). The results suggested that the calcium bound at the interface between the domains fixes the domains in a linear orientation. It is important to note that there is a significant difference in the potential effect of loss of calcium binding in fibrillin compared with COMP. In contrast to the calcium bound to fibrillin, calcium binding to COMP occurs in a cooperative manner. Perhaps calcium stabilizes a particular structure in COMP that promotes proper folding, secretion, and matrix assembly.

The thrombospondin family members contain three regions that have been shown to or have potential to bind divalent cations: the two calcium binding EGF-like repeats, the type 3 domain with some homology to other calcium binding molecules, and the C terminus that has been shown to bind zinc (Rosenberg, Olsson *et al.* 1998). Of these, the only ionic association with biological relevance is that of zinc binding, which permits interaction with collagens I and II at the C terminus. This interaction is enormously significant in that COMP is a component of tendon and cartilage, two tissues that are affected in PSACH by COMP mutations. From the data presented here that show altered conformation of COMP3 peptides with the Asp-446 Asn mutation, we speculate that the native molecular structure is also affected. It is likely that the interaction with fibrillar collagens and potentially other extracellular matrix components may be disrupted.

In conclusion, data presented here demonstrate a direct, specific, measurable effect of PSACH-associated COMP mutations on molecular properties. The mutation, Asp-446 Asn, reduces the potential of the COMP3 peptide to bind calcium ions resulting in a highly altered conformation. The secondary and tertiary structure was disturbed. With

these results, it is proposed that the pathogenesis of PSACH includes the profound alteration of the type 3 domain and possibly the overall molecular structure. Future studies will be required to understand how the altered structure causes storage of selected cartilage components.

Acknowledgements: We acknowledge the excellent technical assistance of Jay E. Gambee, Bruce A. Boswell, Sara F. Tufa, and Virginia L. Scott. In addition, we thank Drs. Lynn Sakai and Dieter Reinhardt for resources and technical support. We also thank Dr. Jürgen Engel for advice and helpful discussions.

Footnotes*

This work was supported by grants from Shriners Hospital for Children (to H. P. B. and D. R. K.), the R. Blaine Bramble Medical Research Foundation, the Fred Meyer Charitable Trust, and by Grant AR45582 from the Department of Health and Human Services.The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
To whom correspondence should be addressed. Tel.: 503-221-3433; Fax: 503-221-3451; E-mail: hpb@shcc.org.

Abbreviations

The abbreviations used are: COMP, cartilage oligomeric matrix protein; TSP, thrombospondin; EGF, epidermal growth factor; PSACH, pseudoachondroplasia; MED, multiple epiphyseal dysplasia; rER, rough endoplasmic reticulum; Wt, wild type; MOPSO, 3-[*N*-morpholino]2-hydroxypropanesulfonic acid.

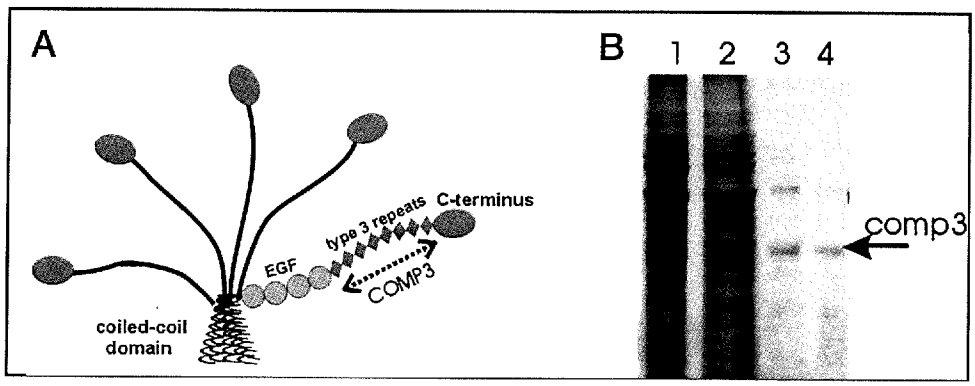


Fig. 1. Expression of recombinant COMP3 peptides. The region of COMP representing the type 3 repeats was expressed as COMP3 peptides (*A*). Cell culture medium from transfected 293-EBNA cells was fractionated at 60% ammonium sulfate saturation and separated by SDS-polyacrylamide gel electrophoresis on a 10% polyacrylamide gel under reducing conditions (*B*). *Lanes 1* and *2* are the Wt and Mu precipitates, respectively, and *lanes 3* and *4* are the supernatants. The Coomassie Blue band indicated by the *arrow* was transferred onto a nylon membrane and identified to be the COMP3 peptide by sequence analysis.

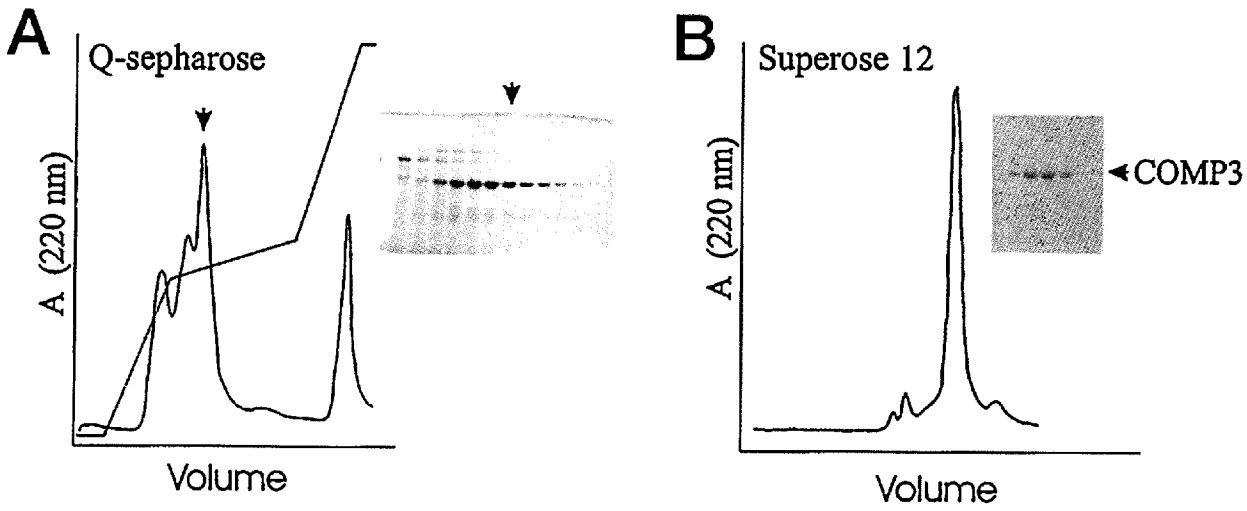


Fig. 2. Purification of the COMP3 peptides. The 60% ammonium sulfate supernatant was dialyzed and separated by Q-Sepharose (*A*) followed by Superose 12 column chromatography (*B*). The fractions were analyzed by SDS-polyacrylamide gel electrophoresis

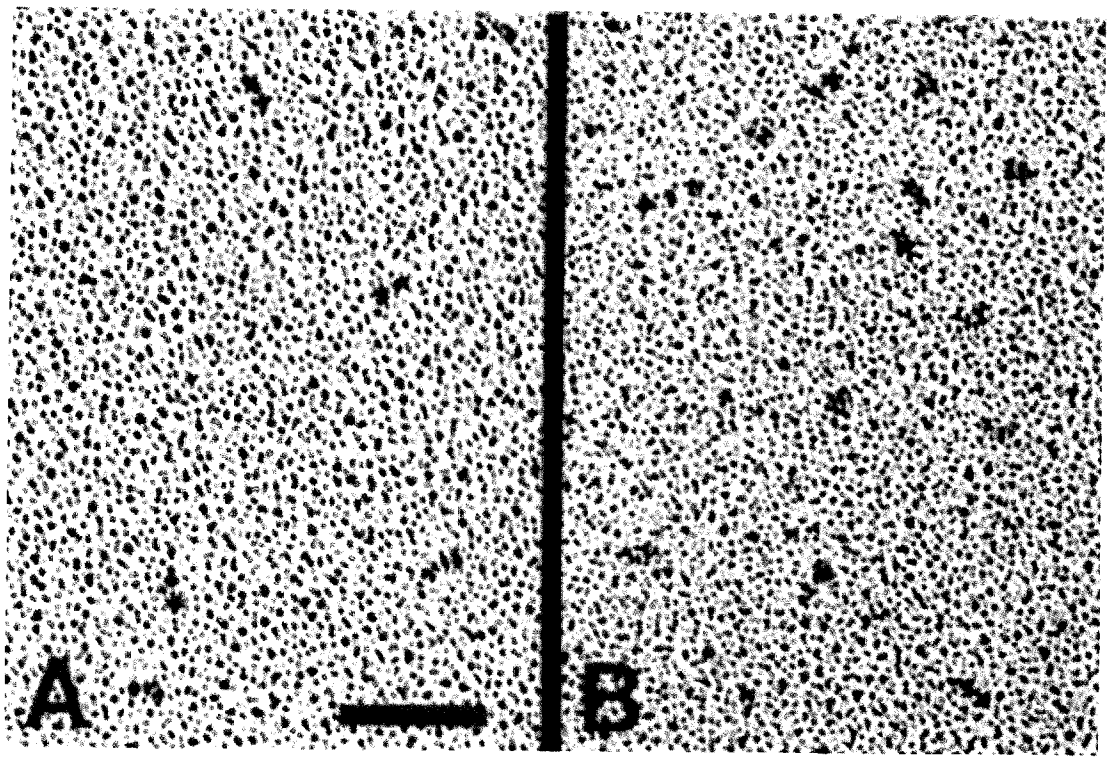


Fig. 3. Rotary shadowed images of COMP3 peptides. Wild-type peptides appear to be short extended rods and are slightly longer than the mutant peptides (*A* and *B*, respectively). *Bar*, 50 nm.

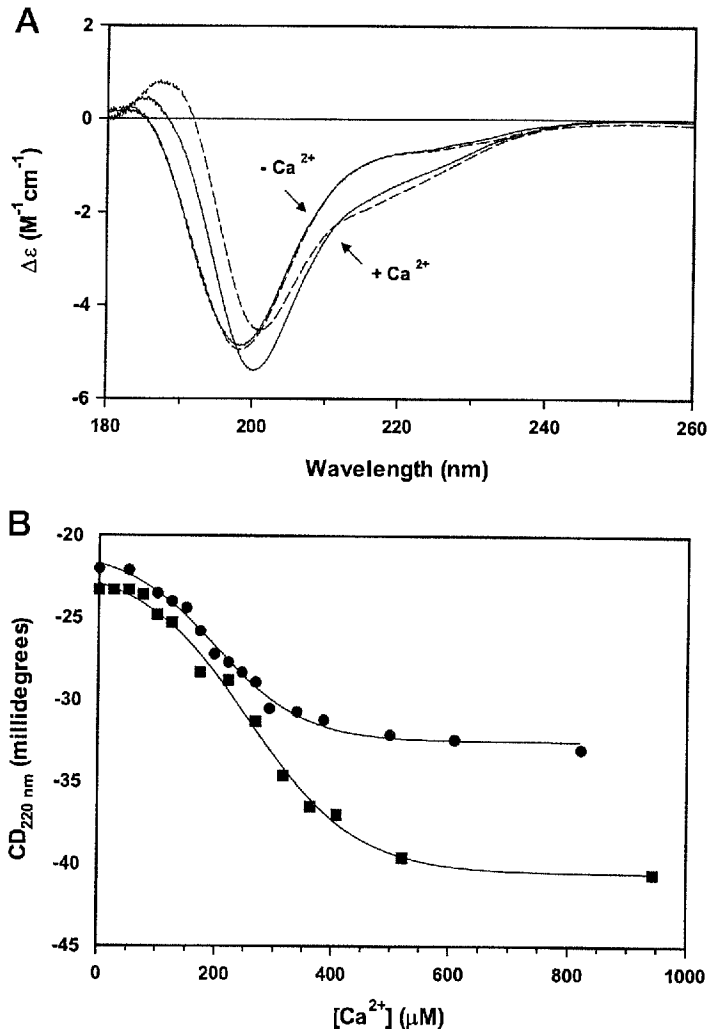


Fig. 4. Secondary structure analysis of the COMP3 peptides. The circular dichroism spectra of wild-type (*dashed line*) and mutant (*solid line*) peptides were analyzed in the presence and absence of 5 mM Ca^{2+} as indicated (*A*). The CD signal at 220 nm was analyzed at increasing levels of calcium for COMP3-Wt *squares* and -Mu *circles* (*B*).

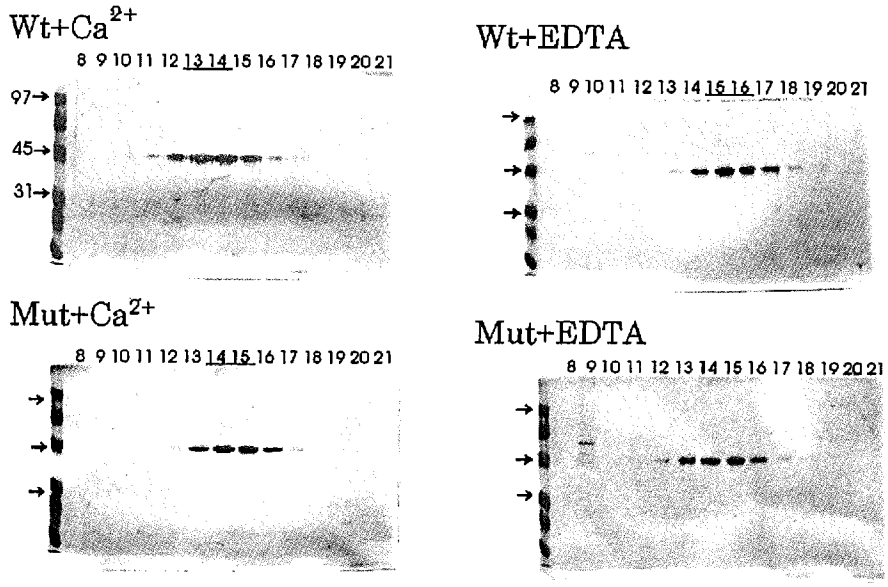


Fig. 5. Conformational analysis of COMP3-Wt and -Mu peptides. Velocity sedimentation in a sucrose gradient shows the effect of calcium on the conformation of the peptides. The sedimentation properties of COMP3-Wt peptides are more affected by calcium than COMP3-Mu.

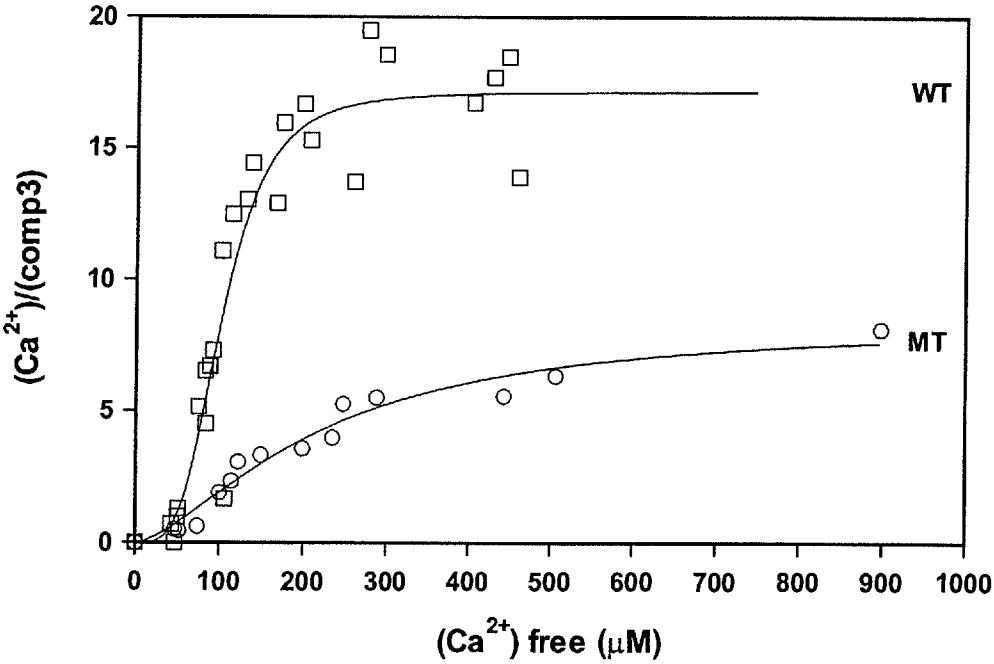


Fig. 6. Calcium binding of the COMP3 peptides. Equilibrium dialysis of the peptides against calcium chloride concentrations of 50 μM to 2 mM indicated differences in the number of available binding sites for COMP3-Wt compared with the mutant peptides.

Structure	H%	A%	P%	T%	O%	Total%	RMSE
COMP	9	28	0	23	41	100	<0.20
COMP+Ca	10	20	2	28	39	100	<0.34
Mutant	9	27	0	23	41	100	<0.20
Mutant +Ca	6	23	2	27	41	100	<0.20

Table I
Secondary structure analysis of COMP3 peptides

CD analysis of COMP3 peptide at 5 °C in MOPSO. H, helix; A, anti-parallel βsheet; P, parallel βsheet; T, turn; RMSE, root mean square error; O, other aperiodic structure.

Chapter 3

Matrix Biology Vol. 20, 439-450, 2001

Selective Intracellular Retention of Extracellular Matrix Proteins and Chaperones Associated with Pseudoachondroplasia.

Janice Vranka[1], Asawari Mokashi[1,2], Douglas R. Keene[1], Sara Tufa[1], Glen Corson[1], Michael Sussman[1], William A. Horton [1,3], Kerry Maddox[1,2], Lynn Sakai[1,2], Hans Peter Bächinger[1,2]*

[1]Research Department, Shriners Hospital for Children, Portland, Oregon, 97201, USA
[2]Department of Biochemistry and Molecular Biology, and [3]Department of Molecular and Medical Genetics, Oregon Health Sciences University, Portland, Oregon, 97201, USA

The experimental proceedures were performed as follows:
AM: Mutational analyses
JV: Immunofluorescence studies
DRK and ST: Electron microscopy

*Correspondence should be addressed to:
Hans Peter Bächinger, Ph.D.
Shriners Hospital for Children
3101 SW Sam Jackson Park Road
Portland, OR 97201
Phone: 503-221-3433
Fax: 503-221-3451
email: hpb@shcc.org

ABSTRACT

Mutations in the cartilage oligomeric matrix protein (COMP) gene result in pseudoachondroplasia (PSACH), which is a chondrodysplasia characterized by early-onset osteoarthritis and short stature. COMP is a secreted pentameric glycoprotein that belongs to the thrombospondin family of proteins. We have identified a novel missense mutation which substitutes a glycine for an aspartic acid residue in the thrombospondin (TSP) type 3 calcium-binding domain of COMP in a patient diagnosed with PSACH. Immunohistochemistry and immunoelectron microscopy both show abnormal retention of COMP within characteristically enlarged rER inclusions of PSACH chondrocytes, as well as retention of fibromodulin, decorin and types IX, XI and XII collagen. Aggrecan and types II and VI collagen were not retained intracellularly within the same cells. In addition to selective extracellular matrix components, the chaperones HSP47, protein disulfide isomerase (PDI) and calnexin were localized at elevated levels within the rER vesicles of PSACH chondrocytes, suggesting that they may play a role in the cellular retention of mutant COMP molecules. Whether the aberrant rER inclusions in PSACH chondrocytes are a direct consequence of chaperone-mediated retention of mutant COMP or are otherwise due to selective intracellular protein interactions, which may in turn lead to aggregation within the rER, is unclear. However, our data demonstrate that retention of mutant COMP molecules results in the selective retention of ECM molecules and molecular chaperones, indicating the existence of distinct secretory pathways or ER-

sorting mechanisms for matrix molecules, a process mediated by their association with various molecular chaperones.

Keywords: Cartilage oligomeric matrix protein (COMP); Pseudoachondroplasia (PSACH); Extracellular matrix proteins; Chaperones

INTRODUCTION

Pseudoachondroplasia (PSACH), an autosomal dominant skeletal dysplasia characterized by short-limbed dwarfism and early onset osteoarthritis, is caused by mutations in the gene encoding COMP (reviewed in (Horton and Hecht 2000). COMP, a pentameric glycoprotein and member of the thrombospondin family of proteins, is found in the extracellular matrix of developing and adult cartilage, and is present in ligament, tendon, and synovium(Hedbom, Antonsson *et al.* 1992; Smith, Zunino *et al.* 1997; Delot, Brodie *et al.* 1998; Di Cesare, Fang *et al.* 1999). The COMP monomer is composed of several modular domains each of which are thought to fold independently of one other. The amino terminal α-helical coiled-coil domain is important in pentamerization of the COMP molecule (Efimov, Lustig *et al.* 1994; Malashkevich, Kammerer *et al.* 1996), followed by four epidermal growth factor (EGF)-like domains and the thrombospondin (TSP) type 3 domain. The TSP type 3 domain has recently been shown to bind calcium (Chen, Deere *et al.* 2000; Maddox, Mokashi *et al.* 2000) and is where the majority of the mutations associated with PSACH are found to occur (Briggs, Hoffman *et al.* 1995; Hecht, Nelson *et al.* 1995; Ballo, Briggs *et al.* 1997; Briggs, Mortier *et al.* 1998; Loughlin, Irven *et al.* 1998). Finally, each COMP monomer contains a C-terminal globular domain which has been shown to bind to collagen I and II with high affinity, an interaction dependent on the presence of divalent cations (Adams and Lawler 1993) (Rosenberg, Olsson *et al.* 1998). This suggests a role for COMP in the organization of collagen fibrils contributing to tissue structure and function.

The hallmark of PSACH is the presence of enlarged, distinctive lamellar rER vesicles within chondrocytes (Maynard, Cooper *et al.* 1972; Stanescu, Stanescu *et al.* 1982). These have previously been shown to contain COMP and type IX collagen (Maddox, Keene *et al.* 1997; Delot, Brodie *et al.* 1998), two components normally expressed and secreted by chondrocytes. Type II collagen was shown to be absent from the ER vesicles in PSACH, indicating a selective process by which specific matrix proteins are either retained or secreted. Large oligomeric glycoproteins such as thrombospondin I and thyroglobulin have been shown to specifically associate with rER proteins during their intracellular processing and secretion (Kuznetsov, Chen *et al.* 1997). These rER proteins most likely function as molecular chaperones in the protein folding process, and they may serve as retention anchors within the quality control system in the secretory pathway for immature or misfolded proteins (Hammond and Helenius 1995; Tatu, Hammond *et al.* 1995; Liu, Choudhury *et al.* 1997; Ellgaard, Molinari *et al.* 1999). Presumably such factors play a role in the cellular response to a mutant COMP molecule, as in the case of PSACH. Whether the abnormal rER accumulations in PSACH are due to selective interactions with mutant COMP molecules or are the result of a secretory pathway impaired by abnormal COMP is not known. In an effort to better understand not only the pathology of PSACH, but also the selective secretory processes for ECM molecules, we have identified distinct ECM and molecular chaperone proteins retained within the rER inclusions of chondrocytes from a patient with PSACH. These studies suggest a shared secretory pathway whereby intracellular accumulation of mutant COMP molecules results in the selective retention of certain ECM molecules, a process that is mediated by specific molecular chaperones.

RESULTS

Mutation analysis. Tissue from the PSACH patient was extracted and analyzed for mutations in the COMP gene. Direct sequencing of a PCR fragment from exon 14 in the patient showed a heterozygous A/G peak at position 1544 of the coding sequence (GenBank accession no. L32137). This base substitution is predicted to result in a change from an aspartic acid to a glycine residue at position 515 in the amino acid sequence (Figure 1). Blood samples from the unaffected parents were analyzed and the DNA was found to carry only the normal COMP sequence indicating that PSACH in this patient was caused by a spontaneous mutation in the COMP gene. The D515G point mutation is located near the carboxy terminal end of the TSP type 3 domain of COMP, a domain consisting of a contiguous set of calcium binding sites rich in aspartic acid residues.

Ultrastructural characterization. Patient and control cartilage was prepared to reveal optimal ultrastructure and examined in the electron microscope. Patient chondrocytes typically contained abnormally distended rER, although the appearance of these inclusions varied considerably and was heterogenous in terms of typical lamellar patterning. Frequently, rER inclusions were seen to occupy a majority of the intracellular space within individual chondrocytes, often with more than one large rER inclusion per cell, and thus affected the overall cell morphology (Figure 2A). Recurrently, the inclusions within chondrocytes residing in the resting and proliferative zones were non-lamellar and contained dense spherical bodies (Figure 2A). Inclusions within other chondrocytes in the same region appeared intermediate in lamellar morphology (Figure 2B). Within hypertrophic chondrocytes of the patient, most inclusions were large and

distinctly lamellar. No similar inclusions were observed in chondrocytes of control cartilage, although the rER was occasionally distended. The apparent metabolic condition of the non-hypertrophic chondrocytes also varied considerably. Often PSACH chondrocytes contained regions of normal rER and Golgi adjacent to the large lamellar bodies within the same cell (Figure 3A). Regions of narrow yet metabolically active rER appeared to connect to the larger rER inclusions (Figure 3B) suggesting the possibility that they are the intermediates in a stepwise progression to formation of the larger intracellular structures.

Intracellular retention and preferential localization to rER inclusions. In an effort to determine whether other molecules are retained abnormally in the rER inclusions in PSACH, we analyzed by immunofluorescence and immunoelectron microscopy the pattern of antibody localization first to ECM proteins in cartilage from the patient with PSACH and from normal, age-matched controls. In control cartilage COMP (Figure 4A) was localized to the interterritorial matrix surrounding the chondrocytes by immunofluorescence. In PSACH, COMP was found to be retained intracellularly both by immunofluorescence (Figure 4B) and by immunogold labeling (Figure 4C), and was shown to specifically localize within the enlarged rER inclusions by the latter method, as has been reported previously (Maddox, Keene *et al.* 1997; Delot, Brodie *et al.* 1998). COMP staining was also present at low but detectable levels in the extracellular matrix in PSACH cartilage, a finding different from what has been published previously (Maddox, Keene *et al.* 1997) and which may be due to the difference in mutation. Additionally we examined the antibody staining and localization for an ER-specific protein. In normal chondrocytes, antibody staining for HSP47, a collagen-specific ER-resident molecular

chaperone (Nagata 1996; Hosokawa and Nagata 2000) was moderate, and showed intracellular localization typical for an ER protein (Figure 4D). In PSACH chondrocytes staining for HSP47 was increased by immunofluorescence (Figure 4E) and was found to localize preferentially to the enlarged rER inclusions by immunoelectron microscopy (Figure 4F). It also appeared that the overall cell size was increased in the PSACH chondrocytes containing the rER inclusions (Figure 4E) as compared with the cells in normal tissue (Figure 4D).

Selective intracellular retention of matrix molecules. To further define the enlarged rER inclusions in PSACH we examined antibody staining and localization of the collagen molecules normally found in cartilage (reviewed in (Buckwalter and Mankin 1998)). In comparison with strong, intracellular staining for COMP within the rER vesicles of the patient tissue (Figure 5A), type II collagen stained only the extracellular matrix of PSACH cartilage (Figure 5B). Immunofluorescent staining for type VI collagen in patient tissue (Figure 5C) was similar to what has been shown previously in normal cartilage, namely that it is extracellular and immediately adjacent to chondrocytes forming part of the pericellular matrix (Keene, Ridgway *et al.* 1998). In agreement with these results, immunoelectron microscopy using the same antibody to type VI collagen only showed extracellular labeling in PSACH tissue (data not shown). In contrast, type IX collagen, a fibril-associated and cartilage-specific molecule that is normally found at low levels in the territorial region surrounding chondrocytes, was strongly localized within the patient cells by immunofluorescence (Figure 5D). Type IX collagen was shown to be preferentially located within the rER inclusions by immunoelectron microscopy, as has been demonstrated previously (Maddox, Keene *et al.* 1997; Delot,

Brodie *et al.* 1998). Type XI collagen, a fibrillar collagen and minor component of mature cartilage that normally associates with type II collagen, was localized at low levels within the rER inclusions of PSACH chondrocytes (Figure 5E). Some extracellular staining was also detectable by immunofluorescence. Type XII collagen is another fibril-associated collagen detected at very low levels surrounding chondrocytes in normal cartilage. Staining for type XII collagen in PSACH cartilage was similar to that for type XI collagen, present at low but detectable levels within the rER of chondrocytes (Figure 5F). Immunogold labeling using the same antibodies was in agreement with the immunofluorescent data showing low levels of labeling for types XI and XII collagen and high levels of type IX collagen within the rER inclusions of PSACH chondrocytes, whereas type VI collagen was secreted out into the cartilage matrix.

Since specific proteoglycans have been shown to interact with collagens in cartilage (Hedbom and Heinegard 1993; Hagg, Bruckner *et al.* 1998) further investigations were carried out to determine whether these proteoglycans were also present in the rER inclusions of PSACH chondrocytes. Aggrecan, the predominant proteoglycan in cartilage, was located throughout the extracellular matrix and in the matrix adjacent to chondrocytes in control cartilage as well as PSACH cartilage, by immunofluorescence (Figure 5G). In contrast to previously published results (Stanescu, Stanescu *et al.* 1993), the rER inclusions in the PSACH tissue did not appear to contain aggrecan, especially in comparison to the high levels of COMP staining (Figure 5A) within the same cells. Further analysis of aggrecan labeling by immunoelectron microscopy with the same antibody confirms these results by demonstrating its presence only in the ECM of PSACH cartilage (Figure 6). Moreover aggrecan is not present above

background labeling within the enlarged rER vesicles of PSACH cartilage. The small proteoglycans decorin and fibromodulin, integral components of articular cartilage that interact with fibrillar collagens, were localized at moderately high levels throughout the extracellular matrix in control cartilage. In PSACH cartilage both decorin (Figure 5H) and fibromodulin (Figure 5I) were not only present in the extracellular matrix, but also detected intracellularly in the rER inclusions at moderate levels by immunofluorescence. Immunoelectron microscopy using the same antibodies to decorin and fibromodulin was consistent with these findings and showed specific localization of the small proteoglycans to the rER inclusions (data not shown.) These results demonstrate that retention of mutant COMP molecules is associated with the selective retention of other matrix molecules, including type IX collagen, fibromodulin, decorin, and low level amounts of types XI and XII collagen, whereas aggrecan and types II and VI collagen appear to be secreted normally.

Intracellular concentration of molecular chaperones. To further characterize the apparent secretory defect of mutant COMP molecules in PSACH, we investigated whether typical ER resident proteins that are normally involved in the quality control mechanism of the secretory pathway were also present in the abnormally enlarged rER inclusions. Antibodies recognizing various ER-chaperones were used for immunofluorescence and immunoelectron microscopy. Included were antibodies to HSP47, the collagen-specific molecular chaperone (Nagata 1996), protein disulfide isomerase (PDI), an ER-resident protein with isomerase and chaperone activities (Wang 1998) and calnexin, an integral membrane protein of the ER that binds glycoproteins and functions as a chaperone (Bergeron, Brenner *et al.* 1994; Tatu and Helenius 1997). Also

included were antibodies to grp78 (BIP), a major chaperone of the ER lumen, and cyclophilin B, an ER-localized peptidyl prolyl cis-trans-isomerase (PPIase). All of the chaperone proteins analyzed were localized intracellularly to the ER at moderate levels in control cartilage by immunofluorescence, as is illustrated in Figure 7, with antibodies to calnexin (Figure 7A), HSP47 (Figure 7C), PDI (Figure 7E), cyclophilin B (Figure 7H), grp78 (Figure 7I) and ubiquitin (Figure 7K). In PSACH chondrocytes staining for calnexin (Figure 7B), HSP47 (Figure 7D), and PDI (Figure 7F) was significantly increased and present specifically within the enlarged rER inclusions by immunofluorescence. The levels of staining of these three chaperone proteins in the rER inclusions from PSACH cartilage, as determined by immunofluorescence, was comparable in intensity to that of COMP staining. In contrast, immunostaining for cyclophilin B (Figure 7H) and grp78 (Figure 7J) was detectable but appeared to be absent from the enlarged rER inclusions in PSACH chondrocytes. Immunoelectron microscopy with the antibodies recognizing calnexin and HSP47 further support the immunofluorescence data showing the presence of calnexin (Figure 8A) and HSP47 (Figure 8B) within the enlarged rER vesicles of PSACH chondrocytes. Calnexin appears to be present in the cytoplasm of the PSACH chondrocytes as well as in the enlarged rER vesicles (Figure 8A), whereas HSP47 appears to be localized predominantly to the enlarged rER vesicles (Figure 8B).

The fact that mutant COMP and other matrix molecules are abnormally retained within rER inclusions indicates that they are not be targeted for proteosomal degradation as has been shown to occur with other abnormally folded proteins (reviewed in (Bonifacino and Weissman 1998). Ubiquitination is one mechanism by which misfolded

proteins are targeted for proteolytic degradation via the 26S proteosome, a process requiring ER-retained proteins to be retrotranslocated to the cytosol (reviewed in (Sommer and Wolf 1997; Wilkinson 1999; Ciechanover, Orian *et al.* 2000). To determine whether ubiquitin was associated with abnormally retained matrix proteins within the rER inclusions in PSACH cartilage, we examined its intracellular localization by immunofluorescence. In control cartilage, ubiquitin was expressed at moderately low levels in the cytoplasm of chondrocytes (Figure 7K). In PSACH chondrocytes, ubiquitin appeared to be present at similar levels within the cells, but was absent from the rER inclusions (Figure 7I). This suggested that the mutant and presumably misfolded COMP molecules were escaping ubiquitination and thus the normal progression to ER-associated degradation by the proteosome.

DISCUSSION

In this study we report a novel D515G mutation in the calcium-binding TSP type 3 domain of COMP in a patient diagnosed with PSACH. It is in this region of the protein where the majority of reported PSACH mutations have been found to occur, demonstrating the importance of calcium binding to the structure and function of COMP. Although the function of COMP remains unclear, the mutations associated with PSACH presumably interfere with the protein's ability to bind calcium thereby affecting the overall structure, and function, of the protein. Distinct and enlarged rER accumulations within PSACH chondrocytes are a characteristic of the disease, and have previously been shown to contain COMP and type IX collagen c. In these studies we demonstrate selective retention of not only COMP and type IX collagen, but also other matrix molecules within the rER inclusions in PSACH chondrocytes suggesting that multiple

secretory pathways may exist for extracellular matrix proteins. Additionally, distinct molecular chaperones were found at elevated levels within the abnormal rER inclusions in PSACH chondrocytes indicating their involvement in the retention of mutant COMP molecules.

Cartilage from the patient was analyzed by electron microscopy and shown to contain chondrocytes with a heterogeneous population of rER inclusions. These intracellular structures varied in their size and appearance and not all contained the typical lamellar pattern of staining. Frequently the PSACH chondrocyte inclusions within the resting and proliferative zones were nonlamellar, whereas most inclusions in the hypertrophic chondrocytes were large and lamellar. Importantly, PSACH chondrocytes containing the enlarged rER vesicles were generally much larger in overall size with the rER vesicles often occupying a large proportion of the cell volume and thus altering the cell morphology. When antibodies to COMP and HSP47, the collagen-specific chaperone, were used for immunolocalization in the PSACH chondrocytes, the rER inclusions were found to contain elevated levels of both proteins as compared with age-matched control cartilage.

Further analysis of PSACH cartilage was performed by immunolocalization of various matrix proteins in order to determine whether they were retained along with mutant COMP or secreted normally. Type II and VI collagen were found to be secreted normally from the patient chondrocytes as determined by immunofluorescence. Data shown here are in agreement with previous results demonstrating that type IX collagen is retained at high levels within the rER inclusions (Maddox, Keene *et al.* 1997; Delot, Brodie *et al.* 1998). Type XI and XII collagen were found to reside within the rER

inclusions at much lower but detectable levels by both immunofluorescence and immunoelectron microscopy. In addition, various proteoglycans known to reside in the cartilage matrix were analyzed in PSACH tissue. In contrast to previously published results (Stanescu, Stanescu *et al.* 1993), we found that aggrecan was not retained within the rER inclusions in PSACH chondrocytes but was secreted out into the cartilage matrix. This was demonstrated clearly by both immunofluorescence and immunoelectron microscopy. In contrast, the small proteoglycans fibromodulin and decorin were found to be present within the rER inclusions at significant levels. These observations demonstrate that in PSACH cartilage there is selective intracellular retention of mutant COMP, types IX, XI, and XII collagen, decorin and fibromodulin. We hypothesize that different secretory pathways or different ER sorting mechanisms exist for various matrix molecules, which may help to explain why some matrix molecules are not abnormally retained within PSACH chondrocytes. This hypothesis is supported by earlier work from Vertel *et al* (Vertel, Velasco *et al.* 1989), who showed segregation of type II collagen and aggrecan into distinct ER subcompartments during their progression through the secretory pathway. The pathology of PSACH chondrocytes may thus be due to compartmentalization of mutant COMP and other matrix molecules during their secretion. Increased local concentrations of mutant COMP, as a result of ER retention, may facilitate interactions with other matrix molecules, especially with those usually found to occur in the cartilage matrix, resulting in the formation of organized protein aggregates within the rER inclusions.

Mechanisms of quality control within the ER are crucial for the proper folding, assembly, and processing of proteins destined for secretion. Misfolded and incompletely

assembled proteins are known to be retained within the ER and are subsequently targeted for degradation, a process referred to as ER-associated degradation (Bonifacino and Weissman 1998; Ellgaard, Molinari *et al.* 1999). This process has been shown to involve interactions with a variety of molecular chaperones and folding enzymes that normally reside in the ER. In order to determine whether specific ER proteins were involved in the selective retention of distinct matrix proteins and mutant COMP molecules, we stained control and PSACH cartilage with antibodies to calnexin, HSP47, PDI, cyclophilin B, and grp78 (BIP). Our results show that HSP47, PDI and calnexin were all detected at significantly elevated levels in the rER accumulations of PSACH chondrocytes as indicated by relative intensities of immunofluorescence staining. Immuno-electron microscopy with antibodies to calnexin and HSP47 confirmed their prescence specifically within the rER inclusions of PSACH chondrocytes. In contrast, cyclophilin B, an ER-resident chaperone with PPIase activity, and grp 78 were present within the PSACH chondrocytes but specifically excluded from the enlarged rER inclusions. These results imply that distinct molecular chaperones are associated with the retention of mutant COMP molecules. The induction or up-regulation of various genes encoding molecular chaperones has been shown to occur in response to the accumulation of unfolded proteins within the ER and is known as the unfolded protein response (UPR). It is thought to assist in the refolding of denatured proteins in the ER by increasing the concentrations of chaperones in this compartment and is suggested to be intimately coordinated with the ER-associated degradation pathway (Sidrauski, Chapman *et al.* 1998; Friedlander, Jarosch *et al.* 2000; Travers, Patil *et al.* 2000). The UPR may account for increased

levels of HSP47, PDI and calnexin detected within the rER inclusions of PSACH chondrocytes as reported here.

ER-associated degradation of misfolded proteins occurs via ubiquitination and ensuing degradation by the 26S proteasome, as reviewed in (Bonifacino and Weissman 1998). Proteins normally destined for this degradation pathway are targeted by ubiquitination after their retrotranslocation from the ER (Plemper and Wolf 1999). We were interested to know whether the mutant COMP molecules in PSACH chondrocytes were being targeted for this ER-associated degradation pathway. Immunolocalization with an antibody to ubiquitin demonstrated only low levels of intracellular staining and ubiquitin was not present within the enlarged rER inclusions. These results imply that the secretory defect in PSACH chondrocytes involves mutant, misfolded COMP molecules that are recognized by specific molecular chaperones but not by ubiquitin. We hypothesize that this is an ER-related event that involves the selective binding and recognition of mutant COMP molecules by specific molecular chaperones preventing its retrotranslocation into the cytosol for ubiqutination and degradation.

A large number of diseases are known in which quality control in the ER plays a role. Many of these involve mutant proteins that are targeted to and retained in the ER and often display a storage phenotype (reviewed in (Aridor and Balch 1999). For example, in α1-antitrypsin deficiency the mutant secretory glycoprotein is abnormally folded and retained in the ER with apparent hepatotoxic consequences (Perlmutter 1996). Although PSACH is associated with mutations in the COMP gene it is not known whether the resulting pathology of the disease is a consequence of the ER storage phenotype or a compromised ECM. In considering the overall structure of the COMP

molecule, a COMP monomer contains multiple domains that are predicted to fold independently of one another. It is reasonable to assume that the N-terminal α-helical coiled-coil domain of COMP assembles co-translationally as has been shown for TSP1 (Prabakaran, Kim *et al.* 1996). As the chain elongates the independently folding EGF domains also form their appropriate structure, and presumably upon completion of protein synthesis only a small portion of the COMP molecule (the type 3 domain) is adversely affected by the mutations. If the quality control system recognizes each domain as an independently folded entity, it faces the task of discriminating between six properly folded domains, which accounts for a majority of the protein, and a single misfolded domain which is most likely still associated with molecular chaperones. The resulting oligomeric complex, containing from one to five mutant COMP monomers and associated molecular chaperones, may either be unrecognizable by the ER-associated targeting and degradative machinery, or may be too large to be retrotranslocated through the translocon and out into the cytosol, resulting in prolonged ER retention.

In conclusion, our results suggest impaired secretion of mutant COMP and various ECM molecules causing abnormal accumulations within the rER of PSACH chondrocytes that likely alter normal cell function. This process is specifically mediated by distinct ER chaperones - acting individually or possibly as a complex of proteins - which presumably recognize the misfolded domain of the COMP molecules and play a role in ER retention of the mutant molecules. However, subsequent ER-associated degradation of the mutant molecules does not appear to occur in PSACH. We hypothesize that the selective ER retention of COMP, type IX collagen, low levels of types XI and XII collagen, fibromodulin, decorin and the ER-chaperones HSP47, PDI

and calnexin in PSACH leads to impaired cell function and reduced secretion of distinct matrix molecules, which may ultimately lead to a compromised ECM. Further investigations into the different secretory pathways or sorting mechanisms for matrix molecules, ER-associated pathways of folding and degradation, and the coordinated involvement of specific molecular chaperones throughout these processes, may provide further insight as to how impaired secretion of mutant COMP ultimately leads to PSACH.

MATERIALS AND METHODS

Patient Samples. The patient was an 11-year-old female diagnosed with PSACH. She exhibited the typical features of PSACH: short-limb short stature, joint laxity, and abnormal radiology. A sample of iliac crest cartilage was removed following informed consent protocol and used as the source of material for subsequent analyses. Tissue samples of iliac crest were also obtained from a 10-year old control for immunofluorescence. The chondrocytes were isolated following collagenase digestion: patient cartilage sample was incubated overnight at 37°C in Dulbecco's modified Eagle's medium (Life Technologies) containing 5% fetal calf serum and 3mg/ml collagenase D (Roche). The isolated cells were cultured for four passages.

Mutation Analysis. Total RNA was isolated from the cultured cells with Trizol (Life Technologies) using the protocol provided by the manufacturer. cDNA was generated by reverse transcription using random primers and the Superscript preamplification system (Life Technologies). Approximately 1μg cDNA was subjected to 40 cycles of PCR using Taq polymerase (Promega) with the forward primer 5'-AGTAGCTAGCTGGTCGCGAC ACTGACCTAGAC-3' and reverse primer 5'-ATAGTTTAGCGGCCGCTACGTGA CTTCAGCGTTCTCCGG -3' to amplify the type 3 domain as described previously

(Maddox, Mokashi *et al.* 2000). The 782bp PCR product was purified using the QIAquick agarose gel extraction kit (Qiagen) and both strands were directly sequenced by automated fluorescent dye terminator method (BigDye, Applied Biosystems, Inc.), using nested primers. The mutation identified on the cDNA was verified on the genomic DNA. Genomic DNA of the proband and her parents was extracted from whole EDTA-collected blood using the QIAmp Blood kit (Qiagen) according to the manufacturer's instructions. Approximately 500ng genomic DNA was subjected to 40 cycles of PCR using Taq polymerase with the forward primer HCOMPI12F1 (5'-CGGGTAGCCTTTGACAAAACG) reverse primer R0-1 (5'-TGGTTGAGCACCACCCAGTT-3'). The 539bp products were purified and directly sequenced as described above.

Antibodies. The rabbit polyclonal antibody to COMP (R3593) was prepared as described previously (Maddox, Keene *et al.* 1997). The rabbit antiserum against HSP47 was a gift from Dr. Nagata and is described elsewhere [(Takechi, Hirayoshi *et al.* 1992; Hosokawa and Nagata 2000). The mouse monoclonal antibody recognizing aggrecan was purchased from Chemicon (MAb2015). The rabbit polyclonal antibody recognizing decorin (LF 113) was a gift from Larry Fisher (Fischer, Tucker *et al.* 1997). The fibromodulin antibody was a gift from Dr. A. Plaas described in (Plaas and Wong-Palms 1993). The mouse monoclonal antibody that recognizes type II collagen (CIICI) was obtained from the Developmental Studies Hybridoma Bank. Antibody to type VI collagen (R279) was a rabbit polyclonal, used as described in (Keene, Ridgway *et al.* 1998). The type IX and type XI collagen antibodies are rabbit polyclonals (R9264, and R16123) prepared and used as described previously (Keene, Oxford *et al.* 1995; Davies,

Oxford *et al.* 1998; Morris, Oxford *et al.* 2000). Rabbit polyclonal antiserum raised against type XII collagen was used as described previously (Keene, Lunstrum *et al.* 1991). The antibodies recognizing cyclophilin B (PA1-027) and PDI (MA3-019) were purchased from Affinity Bioreagents, Inc. (ABR); the antibody recognizing calnexin was purchased from Stressgen (SPA-860) and the antibody recognizing grp78 was a goat polyclonal antibody (SC-1050) purchased from Santa Cruz Biotechnology. The antibody recognizing ubiquitin was a mouse monoclonal (MAB1510) and was purchased from Chemicon.

Immunofluorescence. Frozen sections of 8μm in thickness were made on tissue samples from the patient and from normal age-matched control cartilage. Sections were fixed with 4% paraformaldehyde in PBS for 1 hour at 25°C and then digested with chondroitinase ABC (Sigma) at a concentration of 0.2 units/ml in PBS at 25°C for 3 hours. Sections were washed in 100mM Tris, pH 7.6 and 0.1% Triton, followed by 100 mM Tris, pH 7.6, then blocked for 30 minutes at 25°C in 2% goat serum and incubated overnight at 4°C with primary antibodies. Sections were washed in 100 mM Tris, pH 7.6 and incubated in goat anti-mouse or goat anti-rabbit IgG antibodies conjugated with FITC (Sigma) at 25°C for 2 hours. They were stained with propidium iodide (Molecular Probes) to localize nuclei and then coverslipped in SlowFade Antifade medium (Molecular Probes). Immunofluorescent images were obtained either with a Zeiss Axiophot fluorescence microscope as color slides, which were then digitized, or with a Nikon E800 microscope equipped with a Sensys digital camera (Photometrics; Tucson, AZ) and utilizing Metamorph software (Universal Imaging; Wester, PA).

Electron Microscopy. Tissue samples included iliac crest from the 10-year old proband, navicular growth plate from a 5-year old control, and distal femur from a 15-year old control. Some cartilage was prepared for electron microscopy by fixation in 1.5 % paraformaldehyde and 1.5% glutaraldehyde in 0.1 M cacodylate buffer, pH 7.4 containing 0.05% tannic acid and 0.1% $CaCl_2$. The samples were rinsed in buffer then immersed in 1.0% buffered OsO_4, dehydrated in a graded series of ethanol to 100%, infiltrated in propylene oxide and embedded in Spurr's epoxy. Other tissue was fixed in cold 0.1% glutaraldehyde / 4% paraformaldehyde for 30 minutes, dehydrated at progressively lower temperatures (0 to -20°C), infiltrated in LR White media, and polymerized at 60°C. Sections were mounted on formvar coated nickel grids and floated sequentially on 0.15M Tris HCl, pH 7.5, blocking agent (0.05M glycine in Tris; 2% NFDM containing 0.5% ovalbumin and 0.5% fish gelatin in Tris), and incubated in primary antibody (diluted 1:10 in Tris) for 2 hours. Sections were rinsed several times in Tris, incubated with a combination of 5 and 10 nm gold conjugated secondary antibody (1:10 in Tris with 0.5% ovalbumin, 90 minutes), and given a final rinse in Tris. Labeled sections were examined either stained in uranyl acetate and lead citrate or unstained (Sakai and Keene 1994).

Acknowledgements

This work was supported by grants from Shriners Hospital for Children (to H. P. B., D. R. K., L.Y.S., and W.A.H.) and by Grant AR45582 from the Department of Health and Human Services (to H. P. B.).

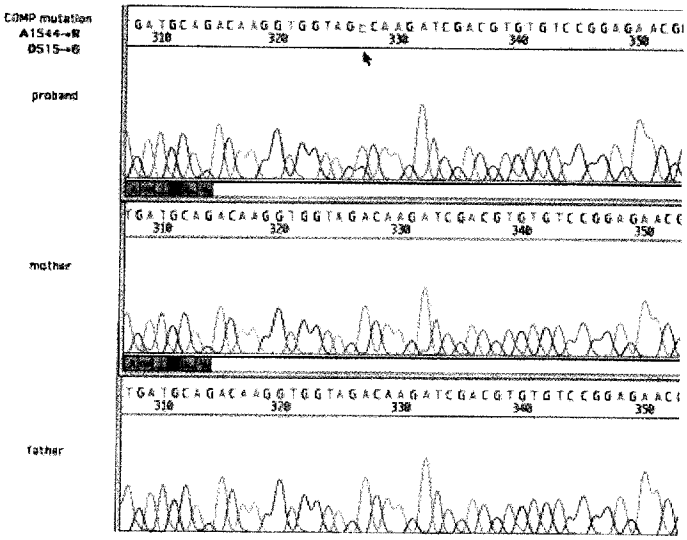


Fig. 1. Mutational analysis of DNA extracted from PSACH cartilage (proband) and unaffected parents. Direct sequencing of PCR fragment of the *COMP* gene showed a heterozygous A/G peak at position 1544 in the coding sequence (arrow). Both parents had a single A peak at the same position.

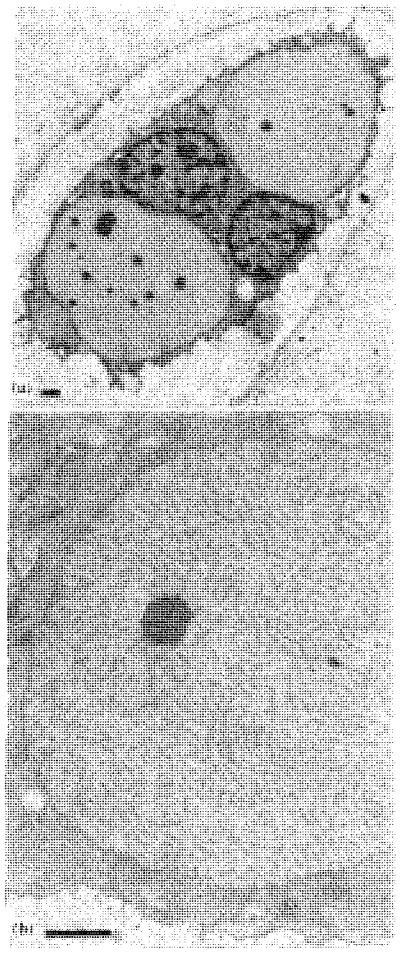


Fig. 2. Electron micrograph of a chondrocyte from PSACH patient cartilage. Enlarged rER inclusions occupy a large volume of the cell and affects overall cell morphology (a). Higher magnification of a PSACH chondrocyte (b) shows enlarged rER inclusion with intermediate lamellar appearance and dense, spherical bodies of unknown composition. (Scale BARS=1 μm.)

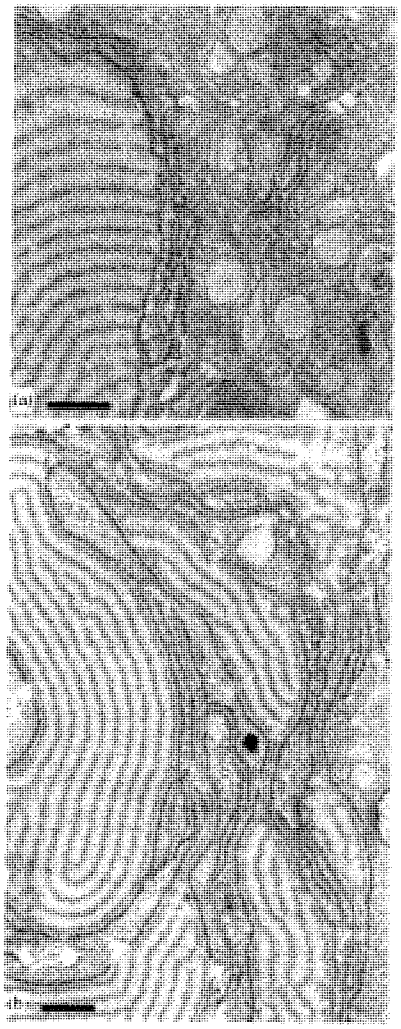


Fig. 3. Electron micrographs of rER inclusions in PSACH chondrocytes at higher magnification. Normal rER and golgi are within the same cell and adjacent to the abnormally enlarged rER inclusions (a). Lamellar appearance of rER inclusions is present throughout ER structures of varying thickness and some inclusions are interconnected (b). (Scale BARS=500 nm.)

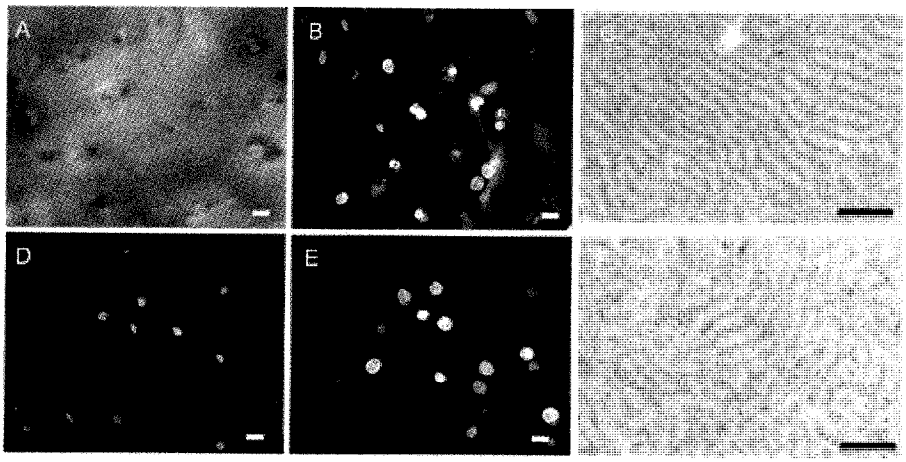


Fig. 4. Immunofluorescence (A, B, D, E) and immunoelectron microscopy (C, F) of cartilage from normal, age-matched control cartilage (A, D) and from the PSACH patient (B, E). Antibodies recognizing either COMP (A–C), HSP47 (D–F) were used for localization in tissues. COMP in control cartilage is extracellular (A) but intracellular in PSACH cartilage (B) and specifically localizes to the rER inclusions (C). HSP47 is intracellular in control and PSACH cartilage (D, E) and specifically localizes to the rER inclusions in PSACH chondrocytes (F). (Scale BARS=20 μm in panels A, B, D and E. Scale BARS=500 nm in panels C and F.)

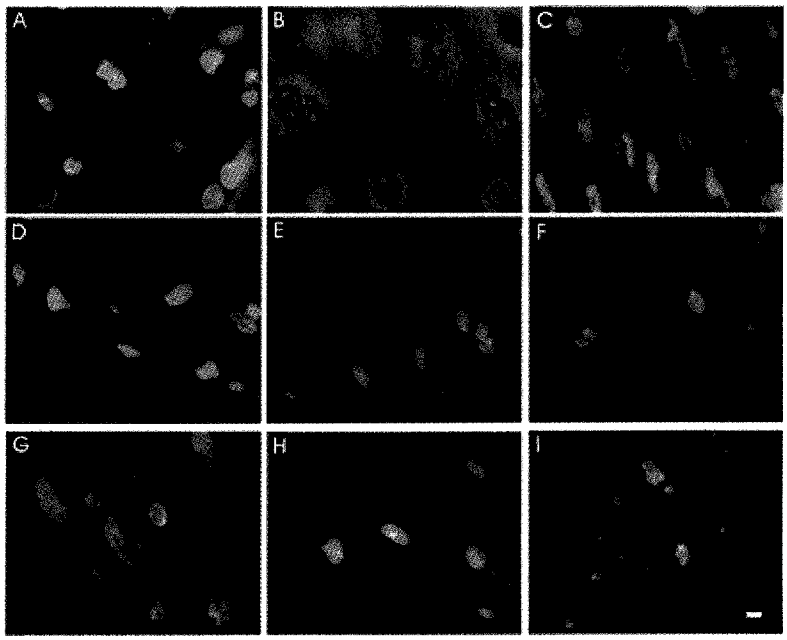


Fig. 5. Immunofluorescence of PSACH cartilage with FITC-labeled secondary antibodies (green stain) and propidium iodide in the nucleus (red stain). COMP (A) is localized within the chondrocytes and to the rER inclusions. In contrast, type II collagen (B) is localized to the cartilage matrix. Type VI collagen (C) is localized to the pericellular regions of the cartilage matrix and is not retained within the cells. Type IX collagen (D) is very strongly localized to the rER inclusions of PSACH chondrocytes. Type XI (E) and type XII (F) collagen are both localized within the PSACH chondrocytes at low levels. Immunostaining for aggrecan (G) was detected only in the cartilage matrix of PSACH tissue, whereas decorin (H) and fibromodulin (I) were detected at moderate levels within the rER inclusions of PSACH chondrocytes and at low levels in the cartilage matrix. (Scale BAR=10 μm and is the same for all panels in figure.)

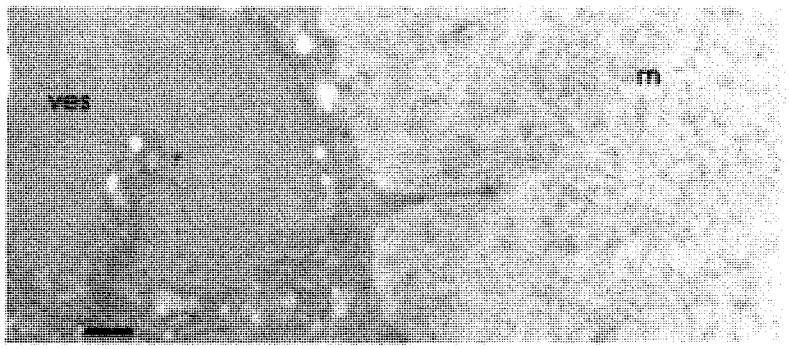


Fig. 6. Immunoelectron microscopy of PSACH chondrocyte with the antibody recognizing aggrecan. Aggrecan is localized at high levels in the extracellular matrix (m) surrounding the chondrocytes but is not present above background levels within the enlarged rER vesicles (ves) of the PSACH chondrocytes. (Scale BAR=500 nm.)

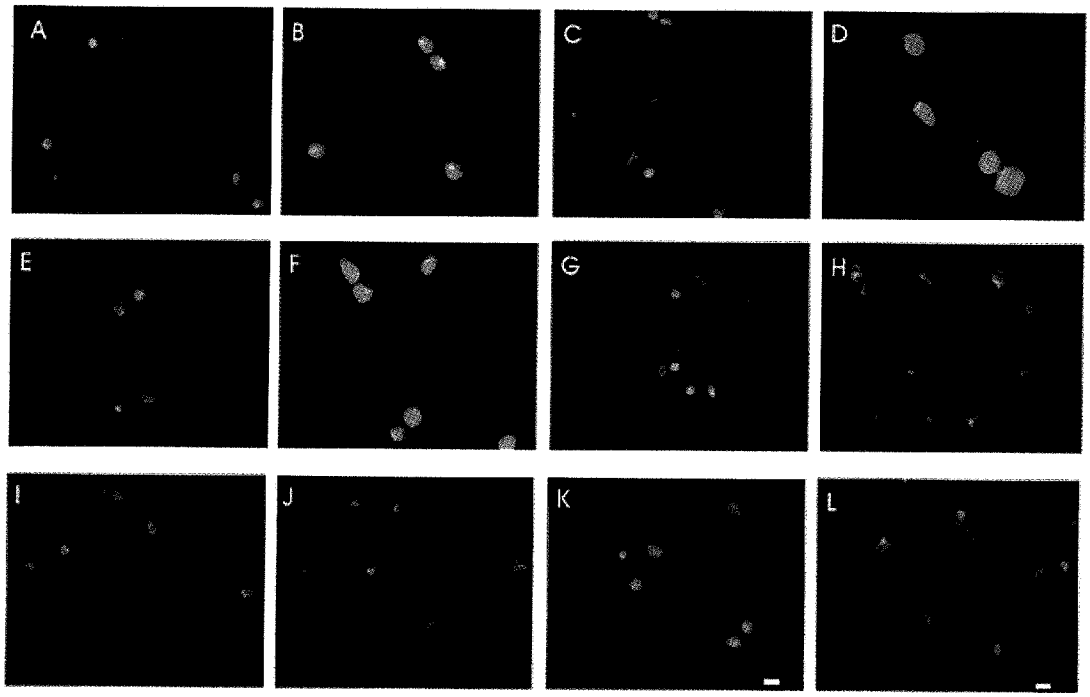


Fig. 7. Immunofluorescence on age-matched control cartilage (A, C, E, G, I and K) and PSACH cartilage (B, D, G, H, J and L) with FITC-labeled secondary antibodies (green stain) and propidium iodide in the nucleus (red stain). Staining for calnexin (A, B) is strongly detected and appears to stain much of the cell including the rER inclusions of PSACH chondrocytes (B). Immunostaining for HSP47 (C, D) is localized adjacent to the nucleus in the control cartilage (C), and is strongly detected within the enlarged rER inclusions of PSACH chondrocytes (D). Immunostaining for PDI is similar to that of HSP47 in that it is detected in the cytoplasm of control chondrocytes (E) and at increased levels within the enlarged rER inclusions of PSACH chondrocytes (F). Immunostaining for both cyclophilin B (G, H) and grp78 (BIP) (I, J) is detected intracellularly in control (G, I) as well as in PSACH chondrocytes (H, J). However the staining levels for both cyclophilin B and grp78 are relatively low in PSACH chondrocytes and both proteins appear to be absent from the enlarged rER inclusions, as determined by relative fluorescence intensity. Intracellular immunostaining for ubiquitin (K, L) is detectable in control chondrocytes (K) as well as in PSACH chondrocytes (L), but also appears to be not present at significant levels within the rER inclusions of PSACH chondrocytes. (Scale BARS=10 μm and is the same for all panels in figure.)

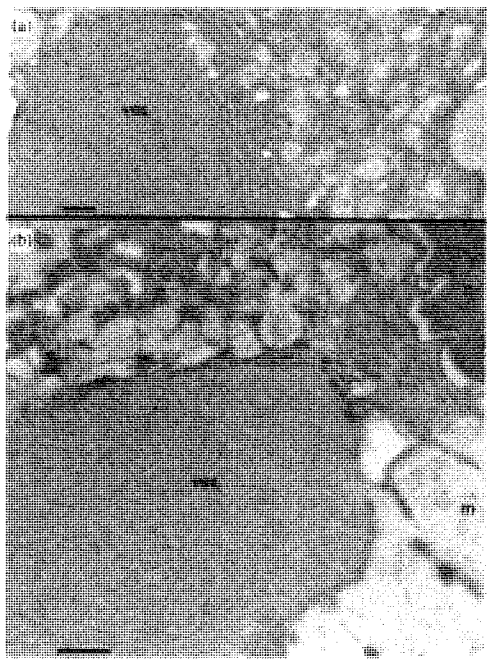


Fig. 8. High resolution immunoelectron microscopy with antibodies to calnexin (a) or HSP47 (b) on PSACH chondrocytes. Immunogold labeling demonstrates the presence of calnexin within the enlarged rER vesicle (ves) as well as in the cytoplasm of the patient chondrocytes (a). Immunogold labeling of HSP47 shows that it is also present within the enlarged rER vesicles of PSACH chondrocytes but does not appear to localize to the cytoplasm, nucleus (n) or extracellular matrix (m) of the cells. (Scale BARS=500 nm.)

CHAPTER 4

RATES OF SECRETION OF CARTILAGE OLIGOMERIC MATRIX PROTEIN VARY BETWEEN DIFFERENT CELL TYPES

Asawari Mokashi[1,2] and Hans Peter Bächinger[1,2]*

[1]Research Department, Shriners Hospital for Children, Portland, Oregon, 97201, USA
[2]Department of Biochemistry and Molecular Biology, Oregon Health Sciences University, Portland, Oregon, 97201, USA.

*Correspondence should be addressed to:
Hans Peter Bächinger, Ph.D.
Shriners Hospital for Children
3101 SW Sam Jackson Park Road
Portland, OR 97201
Phone: 503-221-3433
Fax: 503-221-3451
email: hpb@shcc.org

ABSTRACT

Pseudoachondrplasia (PSACH) is caused by mutations in an extracellular matrix (ECM) glycoprotein called cartilage oligomeric matrix protein (COMP). Hallmark fingerprint-like aggregates are seen in highly dilated rough endoplasmic reticulum of chondrocytes but not in the cells of other COMP secreting tissues like cultured chondrocytes, ligaments and tendons. Here the rate of COMP secretion was determined in bovine primary chondrocytes by metabolic labeling of COMP to compare the rates of COMP secretion in cartilage with other tissue types. Metabolic labeling of type II collagen was done as a control to ensure that the primary chondrocytes maintained their chondrocytic phenotype. These studies showed that COMP was secreted at much slower rates in the cartilage as compared to the other cell types. This suggests that COMP is exposed to different milieu in different cell types and that the longer retention of COMP in the chondrocytes may allow for abnormal associations of COMP with itself and other proteins.

Introduction:

Cartilage oligomeric matrix protein (COMP) is the fifth member of the thrombospondin family of extracellular matrix (ECM) glycoproteins (Newton, Weremowicz *et al.* 1994). Mutations in COMP are associated with pseudoachondroplasia (Briggs, Hoffman *et al.* 1995; Hecht, Nelson *et al.* 1995), a skeletal dysplasia that is characterized by disproportionately short stature, joint laxity, waddling gait and early onset osteoarthritis (Stanescu, Stanescu *et al.* 1982). This manifests at the onset of walking. Ultrastructural analyses of cartilage from patients with PSACH show distinctive and enlarged rough endoplasmic reticulum and highly organized insoluble structures composed of alternating bands of electron dense and electron lucent layers. These appear like fingerprints. These structures are the diagnostic markers of PSACH (Cooper R. 1973).

Mutations in the type III domain of COMP alters its structure (Maddox, Mokashi *et al.* 2000). This causes retention of COMP pentamers in the rER of the chondrocytes along with other proteins and lead to the PSACH pathology (Chen, Deere *et al.* 2000) (Hecht, Hayes *et al.* 2001; Vranka, Mokashi *et al.* 2001). COMP is a pentameric glycoprotein secreted by mainly by cartilage. Tendons, ligaments, vascular tissue and skin also secrete COMP to a lesser extent. However, in case of mutations that cause PSACH, cartilage is the only tissue that is so drastically affected by the mutation. Most of the research is focussed on determining the interactions of COMP with extracellular matrix (ECM) proteins (Rosenberg, Olsson *et al.* 1998; Holden, Meadows *et al.* 2001; Thur, Rosenberg *et al.* 2001). There is little information of COMP biosynthesis by cartilage.

In order to understand why COMP is retained in the chondrocytes, it is essential to look at the rates of secretion of COMP under normal conditions in the cartilage. COMP

secretion under normal and PSACH conditions have been studied in cultured human chondrocytes, ligaments and tendons (Hecht, Deere *et al.* 1998). These studies showed that mutant COMP was not retained in the any of these cells. These data supported the observation by Maddox *et al.* that the fate of cartilage oligomeric matrix protein is determined by the cell type in the case of a novel mutation in pseudoachondroplasia (Maddox, Keene *et al.* 1997). COMP processing was also studied in human recombinant COMP expressed in bovine primary chondrocytes using a gutless adenoviral vector (Dinser, Zaucke *et al.* 2002). The goal of this project is to study COMP secretion by primary chondrocytes.

Methods and Materials:

Isolations of primary bovine chondrocytes: Bovine fetal articular cartilage was collected in ice-cold D-PBS (GIBCO ™) containing 5μg/ml Amphotericin B (GIBCO ™) and 50μg/ml Gentimicin (GIBCO ™). The cartilage pieces were washed twice with 80% ethanol and placed in a beaker containing D-PBS. Each joint was sliced into 1mm pieces and incubated overnight in a digestion medium (Dulbecco's Modification Eagle's Medium (GIBCO™), 5% fetal bovine serum (FBS) form GIBCO™, 0.2% (w/v) collagenase (Sigma), 5 μg/ml amphotericin B (GIBCO ™) and 50 μg/ml Gentimicin (GIBCO ™)) at 37 °C. The collagenase-digested cartilage was spun at 3000 g to pellet out the chondrocytes. The pellet was resuspended in DMEM and layered on OptiPrep (Nycomed Pharma AS, Oslo, Norway) cushion diluted 1:1 in DMEM with 5% FBS. The tubes were spun at 1100 rpm for 15 minutes. The tissue debris formed a pellet at the bottom of the tube, while chondrocytes formed a layer in the OptiPrep-DMEM

interphase. The chondrocytes were washed 5 times in 5X volumes of D-PBS and incubated in whole medium (DMEM containing 5% FBS, 10 ng/ml TGF β and the antibiotics) at 37 °C for one hour on a shaker. The cell number (cells/ml) was counted with a hemocytometer and aliquots prepared for the subsequent experiments.

Analysis of COMP biosynthesis by bovine primary chondrocytes: Primary chondrocytes (10^8 cells/ aliquot at 10^7cells/ml) were suspended in cold, methionine-free medium prepared from MEM SelectAmine kit (GIBCO BRL) with 10 ng/ml TGF β and incubated at 37 °C for 1 hour. All the incubation steps were carried on the shaker. Radioactive [^{35}S] L-methionine (Amersham AG1094 Redivue cell labeling grade) at final concentration of 50 μCi/ml was added to each tube and the cells were labeled for 30 minutes. The radioactive medium was replaced with cold, complete medium containing excess unlabeled (10 mM) methionine. Chase times were from 0 to 18 hours. At each time point, cells were placed on ice for two minutes and pelleted by centrifugation at 4°C. The supernatant was transferred to another tube and the pellet and the supernatant were frozen in isopropanol-dry ice shelling bath and stored at -70°C until further use.

Protein Precipitations: Cell lysates and the medium representing equal cell number (estimated by DNA quantitation) were used for each step. Total proteins were precipitated from the medium with 10% trichloroacetic acid (TCA), washed twice in ice cold acetone and resuspended in sample buffer (Tris-glycine buffer). Full length COMP was immunoprecipitated from the cell lysate using the following method: 100 μl/aliquot protein G sepharose beads (GammaBind™ G Sepharose™, Amersham, Pharmacia Biotech AB) were washed in buffer A according to manufacturer's instructions. The

beads were incubated with 50 µl rabbit 3593 polyclonal anti-COMP antibody (described in (Vranka, Mokashi *et al.* 2001) at 4 °C overnight on an end-over-end nutator. Excess antibodies were washed off with buffer A.: The cell lysates were precleared of non-specific binding proteins by incubating them with 1/5 volume of gelatin-sepharose in a nutator for 1 hour. The pellets were spun down by centrifugation at 12000 g at 4 °C. The pre-cleared lysates were applied to the tubes containing Protein G-Sepharose–antibody beads. The tubes were nutated for 1 hour at 4°C and the GammaBind beads were pelleted. The beads were washed twice with buffer A, spun through a 10% sucrose cushion to remove free radioactivity followed by a final wash with buffer A. Total protein precipitation by TCA was done on the supernatant to ensure that all COMP had been immunoprecipitated. *Antigen elution:* Full length COMP was eluted from the beads by boiling the beads in 10% Laemmli's sample buffer for 10 minutes and spinning down the beads.

The cell lysate and the medium were analyzed by autoradiography following SDS-PAGE on a 5% tris-glycine gel. ^{14}C-labeled full length COMP was also run parallel to the samples. Storm image capture system (Molecular® Dynamics) was also used to visualize the gel and Image Quant software (Molecular® Dynamics) was used to quantitate the radioactivity with volume quantitation method. Volume quantitation calculates the volume under the surface created by a 3-D plot of pixels intensities and pixel location within a selected region.

DNA was estimated using Hoechst 33258 dye (Pharmacia Biotech) that binds to DNA and not RNA and measured by Dyna Quant 200 fluorometer (Hoeffer Pharmacia

Biotech) according to the manufacturer's instruction. DNA quantity represented the cell number in each aliquot.

Radiolabeling of full length COMP by reductive alkylation: Method of Dottavio-Martin *et al.* (Dottavio-Martin and Ravel 1978) was used to label full-length bovine COMP with [^{14}C] radiolabel. Full length COMP was isolated from the bovine cartilage by the method of Roseberg *et al.*(Rosenberg, **DNA estimation:** Cell pellets were lysed in buffer A (50 mM Tris/HCl buffer, pH 7.5 containing 0.5 M NaCl, 8 mM EDTA, 0.1% Tween 20, 1 mM PMSF) and cell debris pelleted out by spinning at 16000g at 4 °C. Olsson *et al.* 1998). 83 μCi [^{14}C] formaldehyde (NEN™ Life Science products, Inc, NEC-039H) was added to 25 μg full-length COMP in 0.04 M potassium phosphate buffer, pH 7. 10 μg freshly prepared $NaBH_3CN$ (6 mg/ml in the phosphate buffer) was added to the mixture and incubated at 25 °C for 1 hour. The solution was shaken gently at regular intervals. The volume was made to 250 μl with the phosphate buffer and the smaller components were dialyzed away for 16 hours at 4 °C. The radiolabeled full length COMP was visualized by SDS-PAGE on 5% Tris-glycine gel followed by autoradiography.

Pulse chase analysis of type II collagen secretion by bovine primary chondrocytes: The experimental protocol was based on the pulse chase described by Maddox *et al.* (Maddox, Garofalo *et al.* 1997). Primary chondrocytes were isolated as previously described. A wash with D-PBS containing 4mM EDTA was incorporated to inhibit any collagenase that was carried over form the digestion step. The cells (10^8cells/time point at 10^7 cells/ml were allowed to recuperate by incubating in complete medium containing 10% FBS, 50 μg/ml ascorbic acid and antibiotics for 1 hour at 37 °C. Ascorbic acid is

necessary to initiate collagen production. The cells were incubated in proline-free MEM SelectAmine medium with 10% FBS and 50 μg/ml ascorbic acid. *Radiolabeling :* [^{14}C] L-proline (Amersham CFB71) at the final concentration of 50 μCi/ml was added to each tube and incubated at 37 °C for 30 minutes. The hot medium was replaced by cold complete medium containing excess L-proline (10 mM) and 50 μg/ml ascorbic acid. Chase was carried at time points from 0 to 3 hours at 37 °C. All steps were carried on a shaker to keep the chondrocytes in suspension. After each time point, cells were placed on ice for two minutes and pelleted by centrifugation at 4 °C. The supernatant was transferred to another tube and the pellet and the supernatant were frozen in isopropanol-dry ice shelling bath and stored at –70 °C till further use. The cells were lysed with buffer A and the DNA was estimated as described above. The proteins present in the medium were precipitated at with 10% TCA as described previously. They were resuspended in 50 mM Tris/HCl buffer, pH 7.4 containing 200 mM NaCl. Non-collagenous proteins were degraded by pepsinization as follows: one volume of the medium was incubated with 1/10 volume of 0.3 M HCl and 1/10 volume of 1.5 mg/ml pepsin made in 10 mM acetic acid for 15 minutes at 37 °C. The reaction was stopped with 5 μl saturated Tris base. Cell associated proteins were extracted in 50 mM Tris-HCL, pH 7.4, containing 1 M NaCl and 0.1% Triton X-100. Volumes of the medium and cell lysate, representing equal number of cells (as estimated from DNA estimation using Hoechst 33258 dye) were analyzed by SDS-PAGE under reducing conditions and visualized and quantitated as described earlier.

Results

The secretion of COMP by chondrocytes was studied with metabolic labeling experiments. Primary chondrocytes dedifferentiate in serial monolayer with respect to their morphology and phenotype. They change from a round to a flattened fibroblast–like cells and secrete type I collagen instead of types II and IX collagens which are the markers of chondrocytic phenotype (Zaucke, Dinser *et al.* 2001). These cells, are therefore, not ideal for studying *in vivo* conditions. Biopsy material is rare, and available in insufficient quantities to follow protein biosynthesis. Hence, primary chondrocytes freshly isolated from fetal bovine articular cartilage is the best system to imitate the environment in cartilage. We carried out metabolic labeling of COMP in bovine primary chondrocytes isolated from fetal bovine joint cartilage. These cells were kept in suspension and prevented from growing in monolayers. Growth factors like IGF-I, TGF β are entrapped in the cartilage matrix and released under specific conditions affect chondrocyte behavior (van der Kraan, Buma *et al.* 2002). Among them, transforming growth factor-beta (TGF-β) plays a significant role in promoting chondrocyte anabolism *in vitro* (enhancing matrix production, cell proliferation, osteochondrogenic differentiation) and *in vivo* (short-term intra-articular injections lead to increased bone formation and subsequent cartilage formation, beneficial effects on osteochondrogenesis) (Grimaud, Heymann *et al.* 2002). Since the basal levels of COMP biosynthesis in the cells harvested from cartilage are low, TGF β was included in the metabolic labeling studies to increase COMP expression (Recklies, Baillargeon *et al.* 1998). COMP secretion by primary chondrocytes was observed in absence of TGF β (data not shown). TGF β increased the amounts of COMP synthesized by the cells without affecting the

rates of secretion. Metabolic labeling of type II collagen, a marker for chondrocytes, was done in parallel to ensure that the isolated cells have retained their phenotype. Collagen II biosynthesis has been studied previously (Maddox, Garofalo *et al.* 1997) and is a useful control for the experiments to ensure that the cells have recovered form isolation process.

Pulse chase experiments of [^{35}S] L-methionine (Fig. 1) labeled COMP determined that COMP is secreted out of the cells after an hour of chase. This indicates that COMP pentamer takes at least 90 minutes to become transport competent (30 minutes pulse and 60 minutes chase). Most of the secretion occurs between 2 hours (when only 15% total COMP has been secreted in the medium) to 4 hours. Even at 4 hours, the cell is not emptied of COMP, but retains 8.5% of total COMP. Secretion of type II collagen (Fig. 2) was not altered by the rigorous cell-isolation procedures and was comparable to previously published data (Maddox, Garofalo *et al.* 1997).

Discussion:

PSACH, a developmental disorder that leads to stunted growth, is caused mainly by mutations in COMP. COMP is retained in the rough endoplasmic reticulum (rER) of the chondrocytes along with other proteins that are either cartilage matrix proteins or rER chaperones. The retained proteins can be observed as fingerprint like structures in a highly dilated rER. In order to understand the pathology of PSACH, a close look needs to be taken at the biosynthesis of COMP in chondrocytes. While several cell types make and secrete COMP, none of them, except for cartilage are affected so drastically by the mutation. While there are some studies on COMP biosynthesis in ligaments, tendons, cultured chondrocytes and primary bovine chondrocytes expressing recombinant human

COMP, no thorough studies have been done to investigate COMP biosynthesis by primary chondrocytes. We studied COMP secretion by isolated primary bovine chondrocytes as these cells have morphological and biochemical characteristics similar to those embedded in the cartilage matrix. The primary chondrocytes secrete COMP molecules at rates that are different from those by other cell types.

Studies in cultured chondrocytes showed that COMP is present in the media after 1 hour and is depleted from the cells by two hours (Hecht, Deere *et al.* 1998). In cultured tendons and ligaments, COMP is present in the medium 30 minutes and completed completely depleted intracellularly by 2 hours. In contrast, in the primary chondrocytes, 85% of COMP secretion takes place after two hours. Dinser *et al.* introduced human normal and mutant full-length COMP molecules in bovine chondrocytes through a gutless adenovirus vector. They observed that the cells are depleted of 90% normal, exogenous COMP within 5 hours and all COMP is secreted out shortly after 20 hours. However, the time course of COMP production following the viral infection was not equivalent to the regulation of endogenous COMP synthesis. The difference in their system may lead to the slight differences in rates of COMP secretion seen between the recombinant system and our system.

Comparison of COMP processing to thrombospondin I, which is similar to COMP, with 60% identical residues (Morgelin, Engel *et al.* 1992; Oldberg, Antonsson *et al.* 1992; Bornstein and Sage 1994 and Adams and Lawler 1993)) shows that thrombospondin is processed at much faster rate than COMP ($t_{1/2} \sim 30$ mins). However, these studies have been done in thyroid epithelial cells. (Prabakaran, Kim *et al.* 1996) and cultured enodthelial cells (Vischer, Beeck *et al.* 1985).

Differences in the processing rates of COMP in different cell types indicate that COMP is exposed to different milieu in different cell types. COMP may encounter chondrocyte-specific proteins that may cause its longer retention within the cell. Immunohistological studies on PSACH cartilage biopsy material by Vranka *et al.* (Vranka, Mokashi *et al.* 2001) have shown that ECM proteins like types fibromodulin, decorin, types IX, XI, XII collagens immunolocalize within the fingerprint structures in the dilated rER of the chondrocytes. Longer retention times may allow enough time for COMP to prematurely form complexes within the rER, which should have been formed extracellularly. This could explain tissue specificity of the fingerprint structures that are the histological markers of PSACH (Maddox, Keene *et al.* 1997). Further studies need to be done to determine the causes of slower rates of COMP secretion by primary chondrocytes as compared to the other cell types.

Conclusion: This is the first study that elucidates processing of COMP in primary chondrocytes, a system that closely resembles COMP processing *in vivo*. Comparisons in rates of COMP secretion in different cell types indicates that tissue specificity of PSACH markers may partly be due to longer retention period of COMP within the chondrocytes complimented with chondrocyte specific environment in the cell rER.

Acknowledgements

This work was supported by grants from Shriners Hospital for Children (to H. P. B.) and by Grant AR45582 from the Department of Health and Human Services (to H. P. B.).

List of abbreviations: COMP, cartilage oligomeric matrix protein; PSACH, pseudoachondroplasia; ECM, extracellular matrix; TGF β, transforming growth factor β and TCA: trichloroacetic acid.

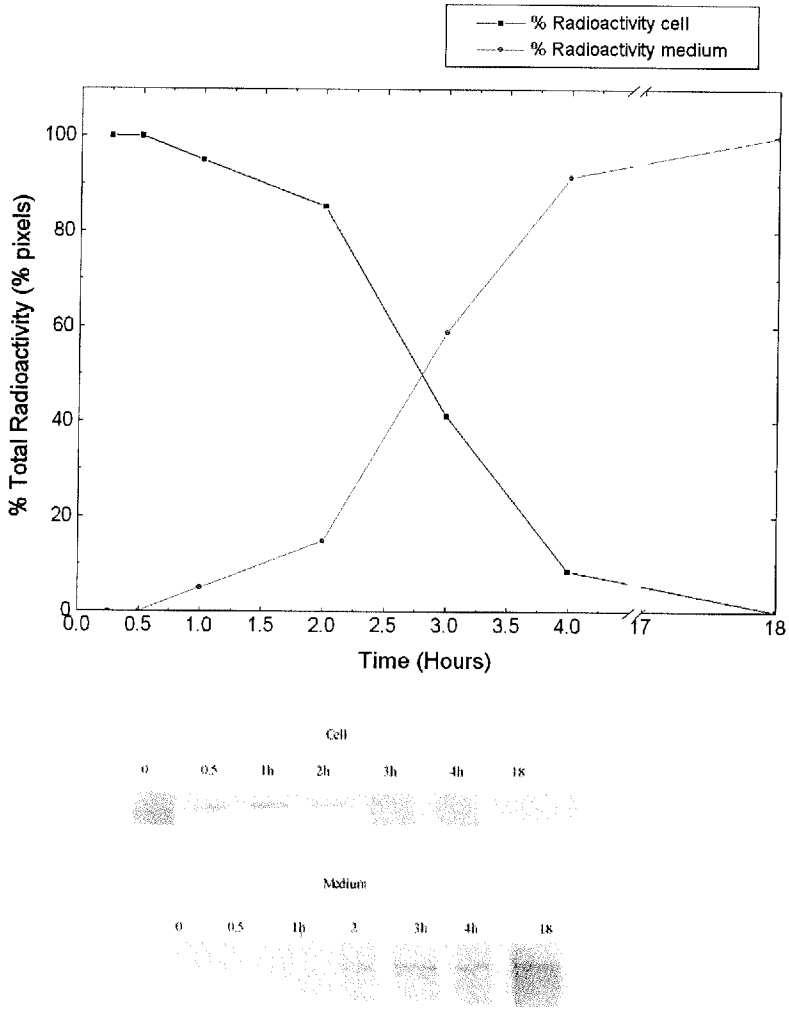


Figure 1. Secretion of COMP by bovine primary chondrocytes. Freshly isolated bovine chondrocytes were pulsed with [^{35}S] L-methionine for 30 minutes and chased for time points ranging from 0 to 18 hours. (Top) shows proportion of COMP found in chondrocytes and medium as compared to the total COMP. Red circles represent % total COMP present in the cell lysates and the black line and circle represent COMP in the medium. (Bottom) shows full length COMP in cell lysates and medium in the pulse chase experiment, visualized by autoradiography.

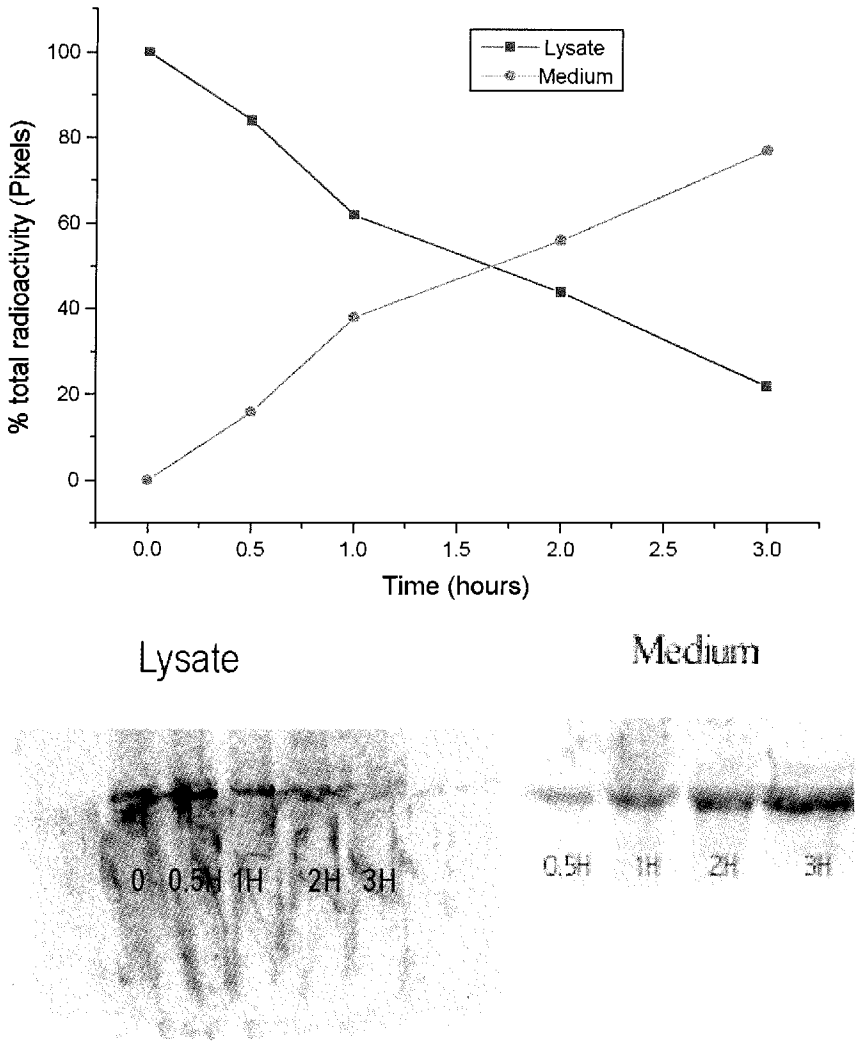


Figure 2.Secretion of type II collagen by bovine primary chondrocytes. Freshly isolated bovine chondrocytes were pulsed with [^{14}C] L-proline for 30 minutes and chased for time points ranging from 0 to 3 hours. Fig 2 (top) shows the proportion of type II collagen found in the chondrocytes and the medium as compared to the total type II collagen. Red squares and line represent % total type II collagen present in the cell lysates and the green squares and line represent type II collagen in the medium. Fig 2 (bottom) shows type II collagen in cell lysates and medium in the pulse chase experiment, visualized by autoradiography.

Chapter 5

PSACH Phenotype is caused by Retention of Mutant COMP in the rER and Through Premature Associations of COMP with ECM proteins

Asawari Mokashi[1,2] and Hans Peter Bächinger[1,2]*

[1]Research Department, Shriners Hospital for Children, Portland, Oregon, 97201, USA
[2]Department of Biochemistry and Molecular Biology, Oregon Health Sciences University, Portland, Oregon, 97201, USA.

*Correspondence should be addressed to:
Hans Peter Bächinger, Ph.D.
Shriners Hospital for Children
3101 SW Sam Jackson Park Road
Portland, OR 97201
Phone: 503-221-3433
Fax: 503-221-3451
email: hpb@shcc.org

ABSTRACT

Mutations in an extracellular matrix (ECM) glycoprotein - cartilage oligomeric matrix protein (COMP) cause a developmental disorder called pseudoachondroplasia. The mutation causes the retention of COMP along with other ECM proteins in the rough endoplasmic reticulum (rER) of the chondrocytes as highly organized fingerprint-like structures. A mutation D→G in the COMP-3 domain is known to alter the structure of this domain. Affinity binding studies with this domain revealed the early stages of COMP retention in the chondrocytic rER. The mutated domain selectively bound to GRP78, calreticulin and ERp72, but not to PDI and cyclophilin B. Thus the effect of COMP mutation alters the structure of the COMP-3 domain and leads to its retention by specific chaperones. We studied the effects of this mutation on the adjacent, C-terminal globular domain (CT). The region of COMP molecule containing the COMP-3 domain and the adjacent, CT domain (C3CT) was expressed in a mammalian culture system, with and without the mutation. Circular dichroism studies on the recombinant peptides showed that the mutation did not alter the structure of the C-terminal domain. This suggested that retention of mutant COMP in the rER, through its mutated COMP-3 domain was followed premature interactions between the CT-domain and nascent ECM proteins being processed by the rER.

Introduction

Pseudoachondroplasia (PSACH) is a developmental disorder of long bones that leads to stunted growth of the affected individual (reviewed in (Horton and Hecht 2000)). It belongs to a group of diseases called chondrodysplasias where the function of the cartilage to act as a template of bone growth is impaired. This results in varying degrees of dwarfism and bone and joint deformities. PSACH is diagnosed in the second or third year of life and is accompanied by a waddling gait, sudden slowing of bone growth and joint laxity. Weight bearing joints like the knees and hips are drastically affected, resulting in deformities and osteoarthritis. Electron microscopy of the joint cartilage reveals highly organized, fingerprint like, alternating electron dense and electron lucent bands in the extensively dilated rER of the chondrocytes (fig.1). These highly symmetrical structures are the hallmark of this disorder (Cooper, Ponseti *et al.* 1973).

PSACH is caused by mutations in the gene encoding cartilage oligomeric matrix protein (COMP), an extracellular matrix (ECM) glycoprotein that is secreted in abundance by the cartilage (Hecht, Nelson *et al.* 1995; Briggs, Mortier *et al.* 1998). These mutations lead to sequestration of COMP as well as other ECM proteins like types IX, XI and XII collagens, fibromodulin and decorin in the rough endoplasmic reticulum (rER) of the chondrocytes (Maddox, Keene *et al.* 1997) (Hecht, Hayes *et al.* 2001; Vranka, Mokashi *et al.* 2001). The retained ECM proteins are absent or present in very low levels in the extracellular matrix (ECM) (Maddox, Garofalo *et al.* 1997; Hecht, Montufar-Solis *et al.* 1998; Vranka, Mokashi *et al.* 2001).

The retention of proteins in a grossly dilated rER are characteristic of disorders known as endoplasmic reticulum storage diseases (ERSDs), hence it is apparent that PSACH is an ERSD. Other examples of ERSDs affecting the skeletal development are Osteogenesis imperfecta (OI) and lethal hypochondrogenesis (Kim and Arvan 1998). OI is caused by the retention and degradation of abnormal type I collagens, a major component of the bone matrix, in the cellular rER. Absence of this collagen in the ECM leads to abnormal development of the bone. Lethal hypochondrogenesis, caused by the retention of misfolded type II collagens in form of granular aggregates in the rER of the chondrocytes Unlike any other known ERSD, retention of COMP is tissue specific. The aggregates have been found only in chondrocytes, but not in other tissues that produce COMP in abundance, like ligaments and tendon cells. Tendons, cultured chondrocytes and ligaments secrete mutant COMP (Maddox, Keene *et al.* 1997and Hecht, 1998 #92; Delot, Brodie *et al.* 1998).

COMP is a 524 kD glycoprotein consisting of five identical chains. It belongs to the thrombospondin family of ECM proteins (Oldberg, Antonsson *et al.* 1992). It is a modular protein. Each monomer has one N-terminal coiled coil oligomerization domain, four EGF-like domains, a thrombospondin type III (COMP-3) domain that contains eight sequence repeats and a C-terminal globular domain (fig. 5). Of all the known PSACH mutations linked to COMP, most mutations are present in the COMP-3 domain (Briggs and Chapman 2002). This domain binds to calcium ions in a cooperative manner. The mutation disrupts the structure of the COMP-3 domain and reduces its calcium binding properties (Chen, Deere *et al.* 2000) (Maddox, Mokashi *et al.* 2000) (Thur, Rosenberg *et al.* 2001). This leads to abnormal interactions between the COMP molecules and other

proteins (Maddox, Mokashi *et al.* 2000; Hecht, Hayes *et al.* 2001; Vranka, Mokashi *et al.* 2001). COMP-3 is not known to interact with any proteins under normal conditions.

Vranka *et al.* and Hecht *et al.* localized selective rER chaperones like GRP78, HSP47, PDI, calnexin, calreticulin, GRP94 and ERp72 to the dilated rER in patient chondrocytes (Hecht, Hayes *et al.* 2001; Vranka, Mokashi *et al.* 2001). These proteins are a part of quality control mechanisms in place within the rER, which clear the organelle of misfolded proteins by promoting their folding or by degrading the irreversibly misfolded proteins (Kim, Kwon *et al.* 1996). However, since there are several proteins retained in the dilated rER, these studies do not give information on the specific substrates of these chaperones. Neither do they provide clues about the early interactions involved in the aggregate formation. We do not know why COMP is retained specifically within the chondrocyte rER or why rER and cytosolic proteases do not degrade it. It is of interest to know how the quality control mechanisms within the chondrocytic rER prevent the exit of COMP from the organelle. We do not know if chondrocyte-specific proteins are responsible for the retention mutated COMP in the rER. In this study we determined that specific molecular chaperones that associate with mutant COMP and play a significant part in the early stages of aggregate formations through affinity chromatography. Similar techniques have been previously used to identify proteins binding to denatured collagen (gelatin) (Zeng, MacDonald *et al.* 1998).

There is growing evidence that the CT domain plays an important role in PSACH pathogenesis. Binding studies show that COMP interacts with some ECM proteins through its C-terminal domain. It binds to with types I, II, and IX collagens in a Zn^{2+} dependent manner (Rosenberg, Olsson *et al.* 1998; Holden, Meadows *et al.* 2001; Thur,

Rosenberg *et al.* 2001) and with fibronectin in a divalent cation dependent manner (Di Cesare, Chen *et al.* 2002). These interactions might be essential for maintaining the structure of the ECM matrix, but it is unclear if the ECM molecules found in the lamellar inclusions aggregate non-specifically or form premature associations with the retained COMP in the PSACH rER.

PSACH mutations in COMP-3 domain cause minor alterations (increased Zn^{2+} dependency) on binding efficiency of CT to types I, II and IX collagens(Thur, Rosenberg *et al.* 2001). It is suggested that the mutations in the COMP-3 domains alter the conformation of the CT domain leading to abnormal interactions with the newly synthesized ECM molecules present in the chondrocytes, (Hou, Putkey *et al.* 2000; Thur, Rosenberg *et al.* 2001). However the highly symmetrical structures of the rER aggregates in the PSACH chondrocytes as well as the retention of selective ECM proteins in the rER suggest that the ECM molecules are not retained within the rER by a random aggregation process, but through very specific interactions with the C-terminal domain of COMP (Vranka, Mokashi *et al.* 2001).

Since the effects of the PSACH mutations in the COMP-3 domain on the CT domain are not known, structural analyses of recombinant peptides consisting of both the domains was done here. Here we show that a previously characterized D446→N mutation in the COMP-3 domain (Maddox, Mokashi *et al.* 2000) does not alter the structure of the CT domain and conclusively resolve nature of interactions between CT domain and the ECM molecules in the PSACH rER.

Results

Affinity chromatography studies: Some mutations in the COMP-3 domain of COMP cause PSACH. Electron microscopy of the affected cartilage reveals fingerprint like alternating electron dense and electron lucent patterns of retained material in highly extended rER. Biophysical characterization of the COMP-3 peptides has shown that these mutations lead to changes in the secondary and the tertiary structure of this domain. The mutant domain has less structure, hence it is of interest of determine if these structural anomalies are detected by the rER quality control systems and lead to retention of COMP in the chondrocyte rER. Immunofluorescence studies were previously done to identify the chaperones and the ECM proteins that were retained in the PSACH rER (Hecht, Hayes *et al.* 2001; Vranka, Mokashi *et al.* 2001). We were unable to distinguish between chaperones that formed complex with COMP and those that were associated with other ECM proteins trapped in the fingerprint structures. One of the goals of this study was to identify the rER proteins that interacted directly with mutant COMP molecules.

The main obstacle to studying this ERSD is the inability to obtain a cell culture system that represents the events occurring in the cartilage. The primary chondrocytes isolated from the patient tissues dedifferentiate with respect to their morphology and phenotype on being cultured in a serial monolayer. They change from a round to a flattened fibroblast–like cells and secrete type I collagen instead of types II and IX collagens that are markers of chondrocytic phenotype (Zaucke, Dinser *et al.* 2001). These cells, are therefore, not ideal for studying in vivo conditions. Biopsy material is rare, and available in insufficient quantities to perform thorough investigations of PSACH. Alginate beads have been successfully used to duplicate the diagnostic fingerprint phenotype in PSACH

chondrocytes (Hecht, Montufar-Solis *et al.* 1998; Hecht, Hayes *et al.* 2001; Dinser, Zaucke *et al.* 2002). However the success of this system is determined by the formation of the aggregates in the cells grown in alginate beads. Thus this system provides us with the end-stages of the disorder, while the early stages of PSACH pathogenesis remain unknown. In this study we circumvented the problems of cell dedifferentiation by developing a system that closely resembles the environment within the chondrocytes to which the mutant COMP molecules are exposed. We were the first to observe the early steps in the processing of mutant COMP molecules in the chondrocytes.

Recombinant WT and mutant peptides were used to make affinity chromatography columns. Lysates from freshly isolated primary bovine chondrocytes kept in suspension, which resemble the chondrocytic milieu seen in vivo, were applied to the column to identify the differences in chondrocyte chaperones that bound directly to the WT and the mutant peptide.

Selective binding of chaperones to the mutant affinity column: A shoulder was consistently observed during the washing step (Fig 3). The proteins eluted in this peak represent weak but specific associations with the COMP peptides. This fits with the definition of chaperones, which are defined as a wide variety of factors that facilitate formation of native protein structures without themselves forming a part of the protein or specifying the protein structure (Laskey, Honda *et al.* 1978). The tightly bound complexes were eluted by changing the pH, ionic strength, removing the chelating ions or with ATP/Mg buffers, as discussed in the Methods and Materials section. The protein fractions corresponding to the shoulder observed in the wash step and the eluates were pooled. Proteins that bound to the column were identified by Western blotting and the

intensities of the protein bands that bound to mutant and WT. columns were compared (fig. 4). Reproducible results were obtained for the blots probed with antibodies against PDI, cyclophilin B, calreticulin, GRP78 and ERp72.

Cyclophilin B (CyP) is a peptidyl prolyl *cis-trans* isomerase that accelerates the folding of some proteins both in *vivo* and *in vitro* by catalyzing the initial steps in the folding and rearrangement of proline containing proteins. It is one of the target genes for unfolded protein response (UPR) and the expression of CyP gene is upregulated under conditions that perturb normal protein processing in the rER (Derkx and Madrid 2001). UPR is a protective mechanism in the ER, which tries to prevent accumulation of misfolded proteins by upregulating the expression of chaperones that enhance folding. As seen in fig.4 there is no significant difference in the amounts of CyP. B bound to the mutant and the WT. columns. This is in agreement with our immunofluorescence data on patient cartilage, where comparable levels of CyP B were observed in patient cartilage and age matched cartilage.

Calreticulin retains the misfolded soluble glycoproteins in the ER through series of bind-and-release cycles and eventually targets the misfolded protein for degradation or storage in the ER (Chevet, Jakob *et al.* 1999). It can also discriminate between native and non-native conformations of non-glycosylated proteins and prevents their aggregation. Thus it is a true chaperone that prevents the aggregation of both glycosylated and non-glycosylated substrates and keeps them in conformations competent for refolding. (Saito, Ihara *et al.* 1999). It also recruits folding catalysts like ERp57 to the misfolded substrates. Higher levels of calreticulin were seen in eluates from the mutant peptide column as compared to WT. peptide column.

GRP78, also known as BiP has been extensively studied. It binds to hydrophobic patches that are usually buried inside a correctly folded protein and a) are exposed on nascent polypeptides during translation or b) become available after perturbations in their structure or c) exposed on unassembled sub units of proteins. By binding to the nascent polypeptides, GRP78 prevents such improper associations (Corsi and Schekman 1996). There is strong evidence that GRP78 targets misfolded proteins for degradation by cytosolic protease complex called the proteosomes. GRP 78 also showed preferential binding to the mutant column compared to the WT. column.

ERp72 belongs to PDI family of thioreductases. It catalyzes correct disulfide bond formation in nascent proteins in the ER. It is found to be associated with misfolded mutant proteins and has yet to be seen in complex with normal proteins. It is induced during unfolded protein response by a cell under stress conditions. It has also been implicated in ER-associated degradation of misfolded proteins (Kim and Arvan 1998). ERp72 was found to bind preferentially to the mutant column.

Table 1 shows the comparison between protein interaction data obtained from immunohistochemistry and electron microscopy of patient and control biopsy cartilage (Vranka, Mokashi *et al.* 2001) and the present study. Cyclophilin B is involved in the processing of COMP-3 peptides, however, data from both studies do not show relation of this chaperone with PSACH phenotype. Though PDI is upregulated in the patient chondrocytes and is included in the enlarged rER, affinity studies show that PDI has preference for COMP3-Wt as compared to the mutant. This suggests that higher levels of

PDI seen in the biopsy material may be due to its association with other structural proteoglycans and collagens trapped in the lamellar aggregates. Our data, that shows preferential binding of GRP78 to the mutant peptide, contradicts the immunohistochemistry data of Vranka *et al.*, which shows that there is no inclusion of GRP 78 in the enlarged rERvesicle. However, it is in agreement with that of Hecht *et al.* [(Hecht, Hayes *et al.* 2001), who show that GRP78 is associated with PSACH aggregates.

Structural studies on the C3CT peptides: We expressed recombinant C3CT peptide comprising the COMP-3 domain and the globular domain in human embryonic kidney cells. C3CT peptides with the D 446→N mutation were expressed along with the wild-type (WT) peptides. This is the first study aimed at deducing the consequences of mutations in the COMP-3 domain on the structure of the adjacent CT domain.

Recombinant C3CT peptides expressed in human embryonic kidney cells (293-EBNA cells) were secreted in the medium and the cell medium was collected and the peptides were purified by anion exchange and size exclusion chromatography (fig. 6a and 6b). The peptides migrated at ~54 kD under non-reducing conditions and at ~70 kD in reducing conditions on a 4-20% tris-glycine gel (Novex). As this peptide is extensively disulfide-bonded, reducing these bonds lead to the formation of an extended structure which traveled at higher molecular weights. Under non-reducing conditions, the C3CT band appears smeared, but under reducing conditions, the bands are sharp. This indicates that higher molecular weight species of C3CT, consisting of small degradation products were disulfide-bonded to C3CT peptides. These degradation products are seen the lower molecular weight bands seen in the reduced gel and are absent under non-reducing

conditions. These degradation products were not expected to interfere with the secondary structure analysis.

Secondary structure analysis of C3CT peptides: The CD spectra of the WT and the Mut peptides were recorded and in the presence and the absence of 5mM calcium chloride. The secondary structure of the C3CT peptides was determined by using the variable selection method. Since just the CT domain was not expressed, we obtained the spectra of this domain in isolation by subtracting the C3CT spectra from previously recorded COMP-3 peptide spectra (Maddox, Mokashi *et al.* 2000). The spectra were weighted to reflect the difference in the sizes of the COMP-3 peptides and the C3CT peptides using the formula:

$$CT = C3CT - [(M_{COMP3}/M_{C3CT}) * COMP3]$$

where *CT, C3CT* and *COMP3* are the spectra of the CT, C3CT and COMP-3 peptides normalized to 1M amino acids. M_{COMP3} and M_{C3CT} are the molecular weights of the COMP-3 and C3CT peptides.

Since the COMP-3 and the CT domains are adjacent but independently folding domains, we hypothesize that the D446→N mutation in the COMP-3 domain would not affect the structure of the CT domain. The changes in the secondary structure of the C3CT peptides would be mainly due to the changes in the COMP-3 domain and would be similar to the effects on the secondary structure seen in the COMP-3 peptides observed previously (Maddox, Mokashi *et al.* 2000).

Unlike the COMP-3 peptides, where the addition of Ca^{2+} led to structural differences between the Wt and the Mutant peptides (Maddox, Mokashi *et al.* 2000), the WT and the mutant C3CT and CT peptides have different structures both in presence and absence of calcium (Fig. 7A). However, only the spectra measured in presence of calcium are physiologically relevant and are discussed in detail. The structural differences between the WT and mutant C3CT peptides in presence of Ca^{2+} are comparable to those in the COMP-3 peptides, in that the mutation causes a loss of the helical structure and a gain in the random coil structure of the C3CT peptides. Unlike the COMP-3 peptides, which showed a gain of anti-parallel-β sheets in the mutant, there is a loss β-sheets in the mutant C3CT peptides (Table 2).

Structural analyses of the CT spectra obtained by subtracting the C3CT spectra from the weighted COMP-3 spectra revealed that there were significant differences between the mutant and the WT CT in absence of Ca^{2+}. These differences were abolished in the presence of Ca^{2+} and only subtle differences were seen between the mutant and the WT C-terminal domain (Fig. 7B and Table 3). This implies that Ca^{2+} ions are required to stabilize the structure of the CT domain and that PSACH mutation in the COMP-3 domain does not significantly alter the structure of the adjacent CT domain.

The present data shows that the mutations causes very slight changes in the CT domain, which include slight loss of the α helical content and minor increase in the unordered structures. Thus, the effect of the PSACH mutations is mainly on the structure of the COMP-3 domain, which includes a loss of α-helical content and a gain in anti-parallel β sheets and unordered structures.

Discussion:

Mutations in proteins often lead to their retention in the ER followed by upregulation of various chaperones through a process called the unfolded protein response (UPR) (Patil and Walter 2001). If correct folding is not possible and degradation is not initiated rapidly, the proteins can interact with other unfolded proteins and form aggregates (Wickner, Maurizi *et al.* 1999) giving rise to conformational disorders. PSACH can be categorized as such a disorder as mutations in COMP molecule lead to its retention along with other proteins and to the formation of fingerprint aggregates. It is not known why COMP is not degraded intracellularly or why COMP retention is tissue-specific. In the present study we identified the proteins that directly interact with misfolded COMP and eventually cause the formation of lamellar aggregates.

Although it was likely that COMP-3 mutations altered the structure of the CT domain, secondary structure analysis of C3CT peptides showed that this was not so. Thus a mutant COMP monomer contains a single unstructured COMP-3 domain (Maddox, Mokashi *et al.* 2000). The non-native domain is recognized as such by rER chaperones like GRP78, Erp72, calreticulin and HSP47 and the COMP pentamers containing the mutant molecules are retained in the rER. Since we were not able identify any COMP-3 binding proteins by amino-acid sequence, we cannot say whether any other chondrocyte specific proteins, besides type IX collagen, were involved in preventing the misfolded COMP from being exported out of the ER in the early stages of aggregate formation.

Another theory for COMP retention in the rER proposes that the misfolded COMP molecules are not escorted out of the rER because they are not recognized as misfolded by their specific molecular escorts. However the direct association of chaperones, known

to retain misfolded proteins in the rER with the mutant COMP suggests that anchoring of COMP in the rER, rather than the absence of forward transport leads to COMP retention within the rER. This conclusion is further strengthened by the finding by Dinser *et al.* (Dinser, Zaucke *et al.* 2002) that primary chondrocytes are able to secrete at least a small proportion of mutant COMP molecules.

At the early stages of cartilage development, the retained pentamers may not be sufficiently unstructured to be recognized as the targets for degradation though they may have sufficient non-native structure to prevent their exit out of the rER (Vranka, Mokashi *et al.* 2001) (Horwich 2002). Even if they were targeted for degradation, they would be unlikely substrates for proteosomal degradation. The COMP pentamers oligomerize as they are translated; hence the mutation is unlikely to effect the structure of the N-terminal domain (Vranka, Mokashi *et al.* 2001). In order to be degraded by the ubiquitin-proteosome system located in the cytosol, COMP has to be translocated out of the rER through a 20Å Sec61 pore (Wickner, Maurizi *et al.* 1999). It is highly improbable that the cylindrical oligomerization domain, with a diameter of 33 Å or the CT domains that have a diameter of 48 Å (Morgelin, Engel *et al.* 1992) could be retrotranslocated out of the 20Å pore. That the COMP molecule has an extended bouquet structure makes such an event more unlikely.

Our data shows that the structure of the CT domain is not significantly altered by a mutation in the COMP-3 domain. Since the mutation alters the Zn^{2+} dependence of interactions between the CT domain and the ECM molecules (Thur, Rosenberg *et al.* 2001), it could be argued that the COMP-3 mutation does not alter of the structure of CT domain in a Ca^{2+}-dependent manner but in a Zn^{2+} dependent manner. However, Hou *et*

al. and Thur *et al.* show that Zn^{2+} does not alter the conformation of the COMP-3 domain under normal or PSACH conditions (Hou, Putkey *et al.* 2000; Thur, Rosenberg *et al.* 2001). Hence the conformation of the CT domain would not be greatly influenced by the mutation in a Zn^{2+} dependent manner and would be amenable to interacting normally with its ECM ligands.

Our data allows us to speculate on the nature of the aggregates retained in the chondrocyte rER. The structural studies show that the effect of the PSACH mutations is mainly on the conformation of the COMP-3 domain. This includes a loss of α-helical content and a gain in anti-parallel β sheets and unordered structures. Studies on aggregated states of various proteinaceous amyloid plaques that are the diagnostic markers of neurodegenerative diseases have shown that proteins in aggregated states are enriched in anti-parallel β-strands as compared to the native, functional proteins (Goloubinoff, Mogk *et al.* 1999). The enrichment of β-strands in the mutant domain of COMP may have a physiological relevance and promote aggregate formation *in vivo*. Proteins retained in the chondrocytic rER are observed as highly organized structures that are unlike the protofibrils and the fibrils associated with amyloid plaques (Bucciantini, Giannoni *et al.* 2002). This indicates that although the mutant molecules may self-associate, they do not form amyloid deposits in the chondrocytes.

It is to be noted that *in vitro* refolding studies as well as studies on protein deposition diseases have shown that misfolded proteins *self-associate* and this association is specific (Horwich, 2002). The contacts between these molecules are not influenced by presence of other denatured proteins. The ECM proteins which were localized in the lamellar aggregates of PSACH chondrocytes (Vranka, Mokashi *et al.* 2001) hence, could not be a

part of a general aggregation process in the chondrocyte ER. Further studies need to be done to determine the nature of the inclusions.

The present studies show that the PSACH mutation has very minor effects on the structure of the CT domain. While these changes may alter it potential for binding with other ECM molecules slightly, the high local concentrations of COMP in the rER, as a result of retention, would compensate for these minor changes and allow these interactions to occur. Thus newly synthesized ECM proteins that are being processed through the rER might form the premature interactions with COMP molecules, typically initiated in the ECM.

Studies on COMP processing in the rER show that bovine primary chondrocytes retain COMP within the cell at least twice as long as other COMP secreting tissues like cultured chondrocytes, ligaments and tendons do. Thus, 1) the slow rates of COMP processing in the chondrocytes 2) in unison with the retention by chaperones that are the part of ER QC and self-association of unstructured COMP-3 domains, 3) as well as the ability of the CT domains to form premature interactions with other ECM proteins being processed by the rER may lead to the fingerprint structures seen specifically in the chondrocytes.

These structures insoluble, hence they are protease resistant. The extensive lamellar network is an impossible substrate for retrotranslocation through the 20Å Sec61 pore. The aggregates cause extensive dilation of the rER and also impair cellular functioning, leading to PSACH pathology. The absence/low levels of retained proteoglycans and collagens in the ECM impede cartilage matrix formation.

This is the first study that looks at the early steps in the processing of mutant COMP molecules. The insights provided by the affinity studies cannot be obtained by other systems being presently used. In conclusion, we not only have identified the proteins that are involved in normal processing of COMP, but also those that actively participate in the retention of misfolded COMP molecules. We also demonstrated that an PSACH-associated COMP mutation in the COMP-3 domain does not alter the conformation of the adjacent C-terminal domain significantly. The C-terminal domain may play a role in the PSACH pathology by prematurely interacting with nascent ECM proteins in the rER.

Methods and Materials:

Generating C3CT peptides: Primary chondrocytes of PSACH patient as well as a normal control were isolated as previously described (Maddox, Garofalo *et al.* 1997) and cultured for up to eight passages. The region containing type 3 domain and CT domain (C3CT) was amplified following reverse transcription of total RNA (Life technologies, Inc., Superscript preamplification kit) using a forward primer contained in a NheI cleavage site (5'agtagctagctagcacccagctccccg3'), whereas the reverse primer contained an NotI site (3'atagtttagcggccgctaggcttgccgcagcccg5'). The resulting polymerase chain reaction bands were gel –purified and ligated into the pCRII vector (Invitrogen). One wild-type clone and one carrying the PSACH mutation were selected following cDNA sequence analysis of the plasmids. The C3CT DNA fragments were ligated into the cloning site in vector pCEPSP-rF17H (Reinhardt D.P. *et al*, 1997) engineered to contain signal sequence of BM40 for the secretion of the recombinant peptide into the medium.

The pCEP4-C3CT vectors were cultured in DH52 cells and the DNA was isolated under sterile condition. Human embryonic kidney cells were transfected with the construct in presence of $CaCl_2$. Selection antibiotics, Geneticin (Life Technologies, Inc., G418) and hygromycin-B (Calbiochem) were added to the cell culture medium 48 hours after transfection at the concentration of 500μg/ml. The concentration of the antibiotics was reduced to 250μg/ml after additional 48 hours. The serum free medium was collected and tested for C3CT peptides using SDS-polyacrylamide gel electrophoresis.

Purification of the C3CT peptides: The medium from the cultured cells was collected and treated with protease inhibitors (0.25 mg/ml AEBSF and 0.5 μg/ml leupeptin). It was extensively dialyzed against 20 mM Tris, pH 7.4. and applied to a Q-Sepharose anion-exchange column (Amersham Pharmacia Biotech). A gradient from 0.225M-1M NaCl in the same buffer was applied and C3CT peptide was eluted at 400mM NaCl. The eluate was concentrated on a Mono-Q HR 5/5 ion-exchange column (Amersham Pharmacia Biotech) and further purified on Superose 12 molecular sieve column (Amersham Pharmacia Biotech). The fractions were analyzed by SDS PAGE gel electrophoresis as well as N-terminal sequencing to test for their homogeneity.

Generating COMP-3 peptides: Recombinant peptides corresponding to normal and mutant (Asp 446→ Asn) type III domain (described in (Maddox, Keene *et al.* 1997)) were generated using human embryonic kidney cells (293-EBNA cells, Invitrogen) as previously described (Maddox, Mokashi *et al.* 2000).

Secondary structure analysis: Circular dichroism (CD) spectra of C3CT peptides in 10 mM MOPSO, pH 6.8 was measured at 5°C in the presence and absence of 5 mM Calcium chloride. The spectra were recorded between 180 nm to 260 nm on a Jasco J-500 spectropolarimeter using thermostatted cells for temperature. Secondary structure analysis was determined using a variable selection method (Compton and Johnson 1986; Compton, Mathews *et al.* 1987).

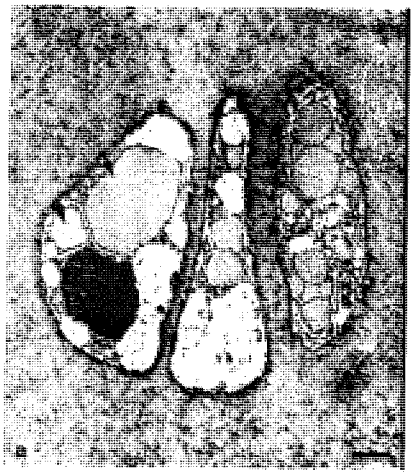
Affinity chromatography: COMP-3 peptides were coupled to NHS-activated Sepharose 4 fast flow (Amersham Pharmacia Biotech) to make 1 ml columns according to manufacturer's directions. The ligands are coupled to the matrix through their primary amine groups. The mutant and wild type COMP-3 peptides were dissolved in coupling buffer (0.1 M $NaHCO_3$ pH 7.4 containing 0.5 M NaCl) at the concentration of 1 mg/ml. The resin was washed with ice-cold 1mM HCl. The ligand was applied to the column and coupling was allowed to proceed overnight at 4°C in an end-over-end mixer. The unbound ligand was washed off with coupling buffer and any remaining active groups were blocked with 1M ethanolamine, pH 7.4. Excess ligand and remaining blocking buffer was removed by alternating washes with coupling buffer and acidic buffer (0.1 M acetate buffer, pH 4, 0.5 M NaCl. The starting material and the flow through that was collected after binding to the column were run on 10% SDS-PAGE gels. All the material had bound to the column.

Fetal bovine chondrocytes were isolated as described above. Briefly, the cell pellet was suspended in lysis buffer (0.1M Tris/HCl, pH 7.2 containing 150mM NaCl, 0.1% Triton X-100 and 0.25 mg/ml AEBSF and 0.5μg/ml leupeptin) to get 10^7 cells/ml. The cell

suspension was kept on a shaker at 5°C for 30 minutes followed by ultrasonication for 30 seconds. The resulting suspension was further lysed by mechanical douncing and centrifuged at 12000g for 30 minutes. The pellet contained the cell debris and nuclei. $CaCl_2$ was added to supernatant at the final concentration of 5mM and was applied to the column. The column was equilibrated in binding buffer (0.1mM Tris/HCl pH 7.2 containing 150 mM NaCl, 0.1% Triton X-100 and 50μg/ml Gentimicin). The supernatant was applied to the column and the unbound material was removed with 5 column volumes of the binding buffer. The bound proteins were eluted by changing the pH, ionic strength and chelating the Ca^{+2} ions with EDTA (0.1M Tris/HCl pH 4 containing 1M NaCl, 0.1% Triton X-100, 4mM EDTA and 0.1M acetic acid). The eluate fractions were sequenced or separated by SDS-PAGE on 10-20% tricine gels (Novex), transferred to sequencing grade PVDF membrane and probed with various antibodies. In some cases, ATP elution step was carried out with binding buffer containing 1mM ATP, 2 mM $MgCl_2$, 2 mM KCl. The column was washed with 5 column volumes of binding buffer containing 20% ethanol and stored in binding buffer.

The proteins that bound to the column were run on a 4-20% SDS-PAGE and transferred to PVDF membrane. They were probed with antibodies against various chaperones. Polyclonal anti-Erp72 (SPA-720StressGen), rabbit anti-grp78 antibodies (SPA-826 StressGen), rabbit anti-PDI antibody (SPA-890, StressGen) rabbit anti-cyclophilin B antibody (PA1-027 ABR) and rabbit anti-calreticulin antibody (PA3-900) were used. N-terminal protein sequencing was also attempted to identify the nature of the bound materials.

Acknowledgements: We acknowledge the excellent technical assistance of Bruce A. Boswell. In addition, we thank Dr. Kerry Maddox for providing us with human embryonic kidney cells were transfected with the C3CT constructs. We also thank Drs. Kerry Maddox and Janice A. Vranka for their comments on this manuscript. This work was supported by grants from Shriners Hospital for Children (to H. P. B.)



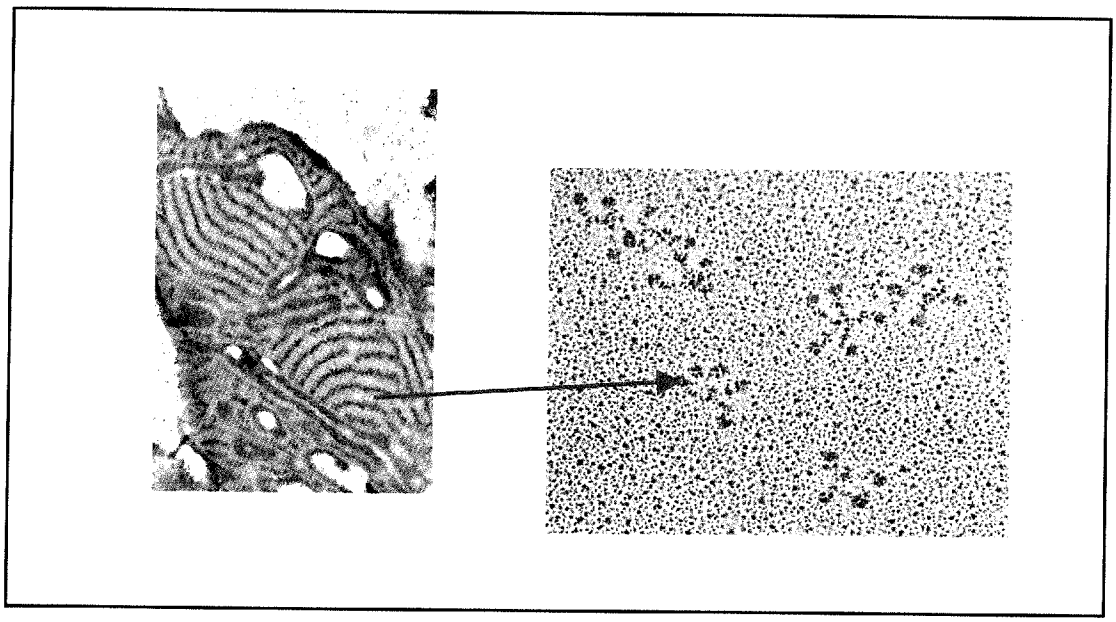


Figure1: Enlarged chondrocytes in cartilage of PSACH patient (top). The electron microscopy of chondrocytes show that rER is enlarged and contains alternating electron-dense and electron lucent fingerprint like aggregated structures (Bottom left). Bouquet-shaped COMP molecules are localized to the dark bands. Reproduced from (Maddox, Keene et al. 1997).

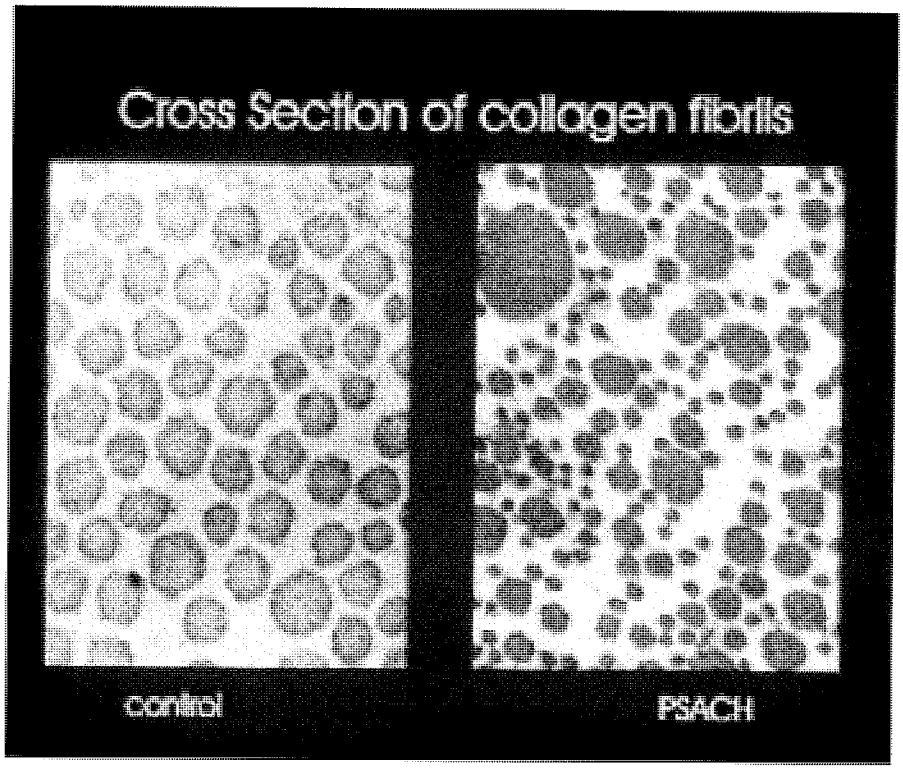


Figure 2. PSACH mutations disrupt fibril formation: Cross section of collagen bundles show that PSACH patients have irregular fibril formation as compared to age-matched non-PSACH control.

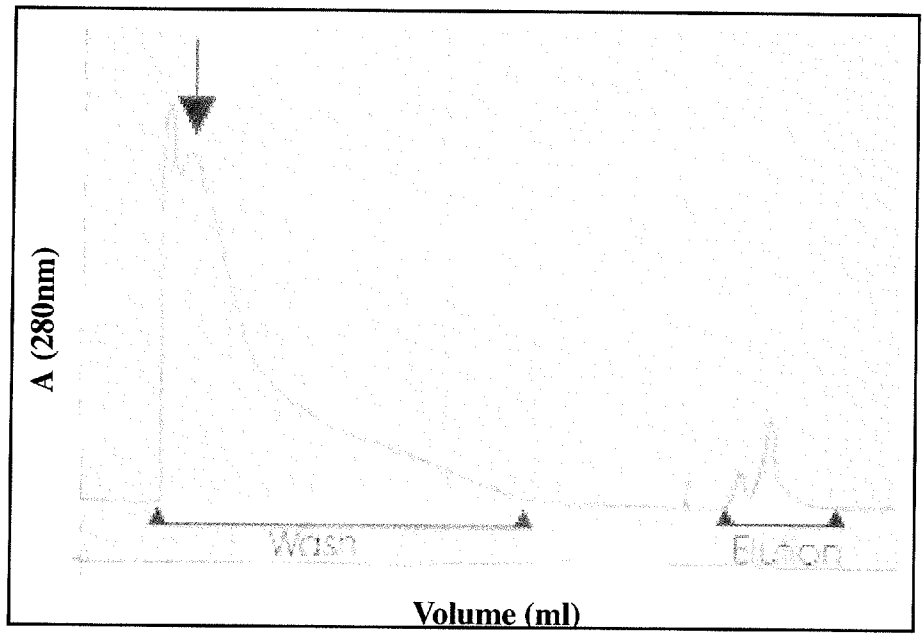


Figure 3:Specific chondrocyte proteins interact with the COMP-3 domain: NHS-sepharose affinity columns were prepared with the recombinant WT and mutant COMP-3 peptides. Lysates of freshly isolated bovine chondrocytes were applied on the affinity columns. Proteins that bound to the COMP-3 peptides were eluted under conditions described in the text. The arrow indicates the proteins that interact in a weak, but specific manner with the COMP-3 peptides.

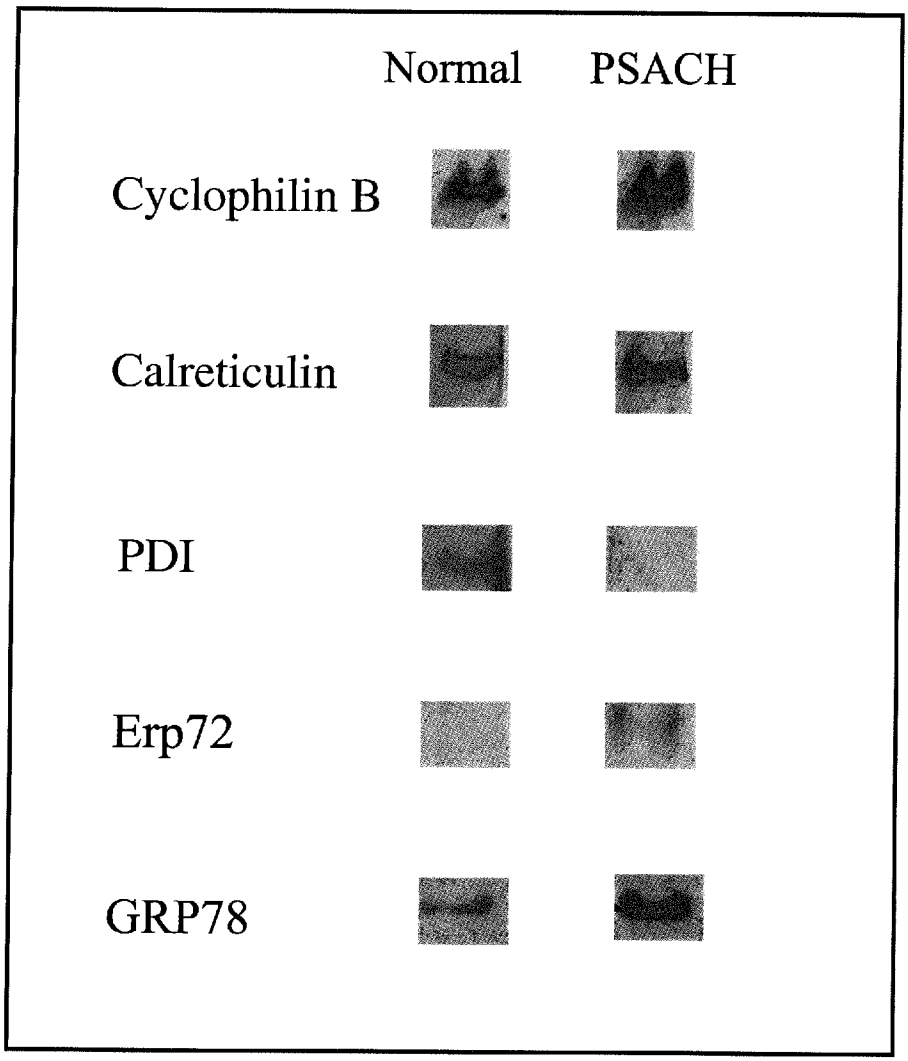


Figure 4: Selective binding of rER chaperones to the COMP3-Mutant peptides: Lysates from fetal bovine primary chondrocytes were applied to COMP3-NHS-sepharose affinity columns. The eluates were separated by SDS-PAGE, transferred to PVDF membrane. Fig.4 shows the Western blots of proteins eluted from COMP3-WT and -mutant peptide columns.

Chaperones	WT. Cartilage (Vranka, Mokashi et al. 2001)	PSACH cartilage (Vranka, Mokashi et al. 2001)	Wt. Column	Mutant Column
Cyp. B	+++	+++ (excluded from enlarged vesicles)	+++	+++
Ubiquitin	+++	+++ (excluded from enlarged vesicles)	N/D	N/D
HSP47	+++	+++++ in vesicles	N/D	N/D
PDI	+++	+++++ in vesicles	++++	+
Calnexin	+++	+++++ in vesicles	N/D	N/D
ERp72	N/D	N/D	+	+++
Calreticulin	N/D	N/D	+++	++++++
GRP78	+++	+++ excluded from enlarged vesicles	+++	+++++

Table1: Comparison between protein interaction data obtained from immunohistochemistry and electron microscopy of PSACH biopsy cartilage (Vranka, Mokashi, et al. 2001) and the present study.

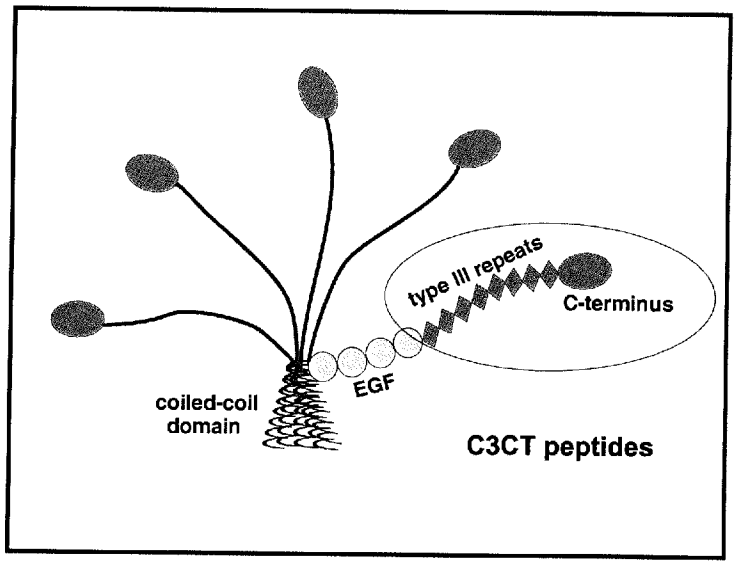


Figure 5: Expression of C3CT recombinant peptides: The region of COMP molecule representing the type III repeats and the C-terminus was recombinantly expressed as C3CT peptides in transfected 293-EBNA cells. The cell culture medium was purified as shown in figure 2. Adapted from (Maddox, Mokashi et al. 2000).

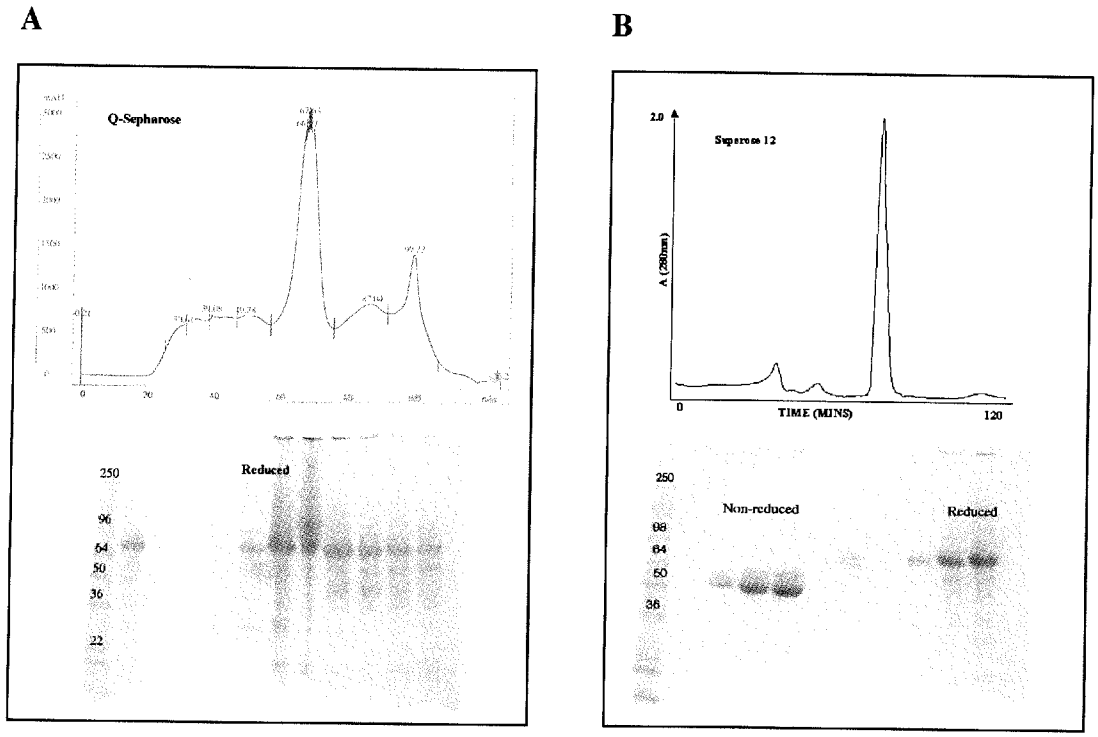


Figure 6. Purification of C3CT peptides: Medium containing recombinant wild type and mutant peptides was separated on Q-sepharose column *(A)* followed by Mono-Q column. The final step in purification was on Superose-12 column *(B)*. The fractions were analyzed by SDS-PAGE on 4-20% Tris-glycine gel.

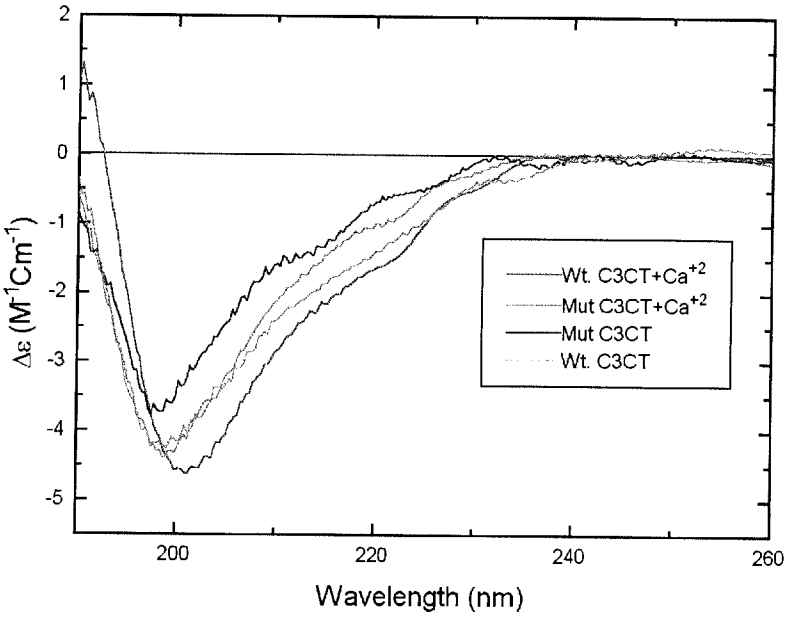


Figure 7A: Secondary structure analysis of the C3CT peptides: Circular dichroism spectra of wild type and mutant C3CT peptides were obtained in presence and absence of 5 mM Ca^{2+}.

Structure	%H	%A	%P	%T	%O	% Total	RMSE
C3CT-wt	7.2	19.3	2.7	28.2	43.1	100	<0.20
C3CT-wt+Ca	9.7	24.1	3.1	28	35	100	<0.34
C3CT-Mut	6.1	23.9	2.8	25.6	41.6	100	<0.20
C3CT-Mut+Ca	6.4	21.2	2.6	27.5	42	100	<0.20

Table 2:CD analysis of C3CT peptides at 5 ℃ in MOPSO buffer: The spectra were analyzed by variable selection method. H, helix; A, Anti-parallel β-sheet; P, parallel β-sheet; T, turn; O, other aperiodic structures; RMSE, root mean square error.

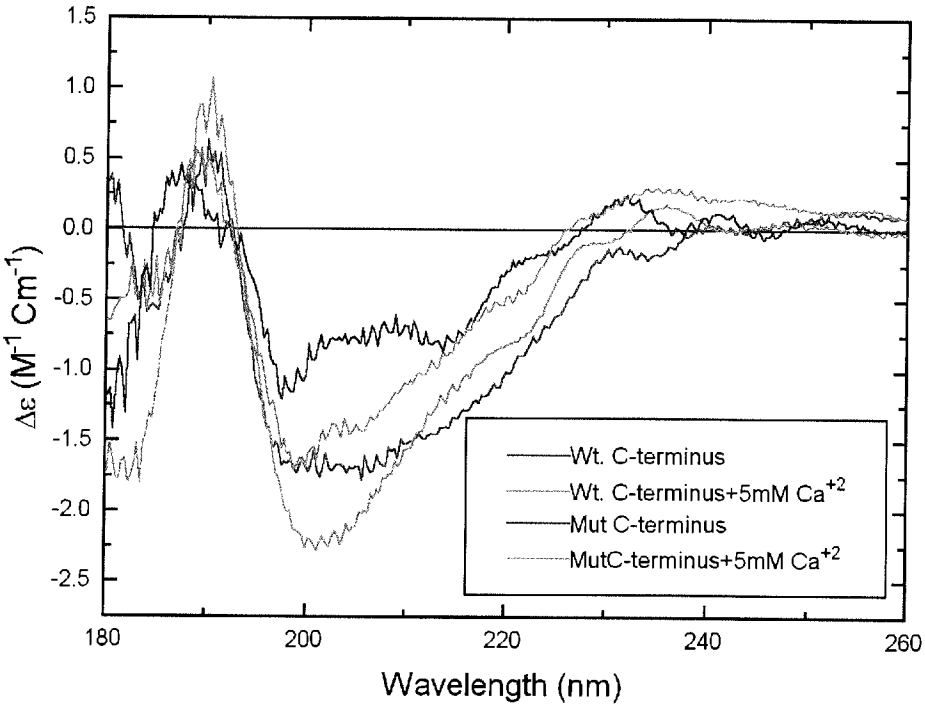


Figure 7B: Secondary structure analysis of the C3CT peptides: Circular dichroism spectra of wild-type and mutant CT peptides were obtained by subtracting weighted COMP-3 spectra from C3CT spectra as described in the text. The spectra were measured in presence and absence of 5mM Ca^{+2} at 5°C.

Structure	%H	%A	%P	%T	%O	% Total	RMSE
CT-wt	7.8	35.4	3.5	20.7	32.7	100	0.20
CT-wt+Ca	7.7	37.4	3.4	21.3	30.7	100	<0.34
CT-Mut	6.4	41.4	3.4	18.2	30.4	100	<0.20
CT-Mut+Ca	6.9	37	3.3	20.5	32.2	100	<0.20

Table 3:CD analysis of CT domain at 5 ℃ in MOPSO: The spectra were analyzed by variable selection method. H, helix; A, Anti-parallel β-sheet; P, parallel β-sheet; T, turn; O, other aperiodic structures; RMSE, root mean square error.

CHAPTER 6

Summary and Conclusions

Mutations in COMP, a protein abundantly secreted by cartilage cells, lead to retention of COMP and several other proteins as lamellar fingerprint structures in the rER of the chondrocytes. This retention eventually causes stunted growth and osteoarthritis associated with pseudoachondroplasia. The primary goal of this thesis was to understand the molecular and cellular mechanisms that caused pseudoachondroplasia. The difficulty in obtaining culture conditions that reproduced the cartilage environment was circumvented by using recombinantly expressed peptides and primary chondrocytes that maintained the chondrocytic phenotype.

Biophysical characterization of recombinant type III domain and C3CT peptides showed the first measurable consequences of a mutation in COMP and provided a molecular mechanism involved in generating a phenotype. The D446→ N mutation in the type III domain resulted in profound alterations of this domain and reduced its calcium binding properties. This effect of the mutation was confined to the type III domain, and it had very little impact on the secondary structure of the adjacent globular C-terminal domain.

Mutational analyses on another PSACH patient identified a novel D515→G substitution in the eighth sequence repeats of the type III domain. Immunohistochemistry studies on the patient cartilage and an age-matched control showed the effects of the structural aberrations on the processing of COMP molecules in the chondrocytic rER. Collectively, the data obtained from affinity chromatography studies, electron microscopy and immunofluorescence helped identify many of the chaperones and ECM proteins involved in PSACH pathology. Chaperones involved in the normal COMP processing along with

those involved in the processing of mutant COMP were identified. I was able to distinguish between the chaperones that interacted directly with COMP and those that interacted with other proteins retained along with COMP. More quantitative binding studies like ELISA will be performed to confirm the specificity of the interactions between the chaperones that preferentially bound to the mutant COMP3 column like GRP78 and calreticulin and the mutant peptide.

The studies also determined that the retained lamellar structures resulted in highly enlarged rER, which occupied a large portion of the cell volume. This resulted in altered cell morphology. Electron microscopy studies and biochemical studies on PSACH chondrocytes grown in alginate bead cultures by Hecht *et al.* and Dinser *et al.* suggest that the affected chondrocytes undergo necrosis (Hecht, Montufar-Solis *et al.* 1998; Dinser, Zaucke *et al.* 2002). Our data (Vranka, Mokashi *et al.* 2001) also supports the studies by Vertel *et al.* and Bonfanti *et al.* (Vertel, Velasco *et al.* 1989; Bonfanti, Mironov *et al.* 1998) indicating that multiple secretory pathways might exist for extracellular matrix proteins. However, attempts to identify novel chondrocytic proteins that bound selectively to normal or mutant COMP molecules were unsuccessful.

The conundrum of tissue specificity of the ER aggregates was partly solved by obtaining the rates of COMP secretion by primary chondrocytes and comparing them with the previously published rates of COMP secretion by other cell-types known to secrete mutant COMP molecules, namely cultured chondrocytes, ligaments and tendons. The results indicated that primary chondrocytes retained COMP molecules for longer times than other tissues. This suggested that longer processing time in the chondrocytes may allow the COMP molecules to concentrate in the rER and provide sufficient time for

abnormal interactions of COMP molecules with other nascent ECM proteins to occur. Since the fingerprint structures are seen only in tissues that synthesize type IX collagen (Maddox, Keene et al. 1997), associations between mutant COMP and type IX collagen may be a major factor in the retaining COMP in the rER. Studies that measure the effects of COMP concentration on the interactions between mutant COMP (D→446N) and type IX collagen and other ECM molecules associated with the inclusions, as well as the time dependence of these associations need to be done. Studies need to be done to determine the cause of longer COMP retention time in the primary chondrocytes. First, we need to determine whether the longer COMP processing time was caused by slower export out of the rER or the Golgi by analyzing the glycosylation pattern of COMP. Sensitivity to either endoglycosidase H digestion or *N*-glycosidase F digestion would indicate whether COMP was localized in the rER or the Golgi respectively. Second, co-immunoprecipitation of COMP with chaperones like PDI, calreticulin, GRP 78 or HSP47 would determine whether these chaperones play a role in longer retention times of COMP in chondrocytes as compared to the other cell types. This could also explain why the mutant COMP molecules escape retention in ligament cells, tenocytes and cultured chondrocytes.

The data presented leads to the to the conclusion that PSACH pathology is caused by the following factors:

1) **Mutant COMP is retained in the chondrocyte rER by ER chaperones:** Heterozygous mutations in the COMP result in 97% COMP pentamers with non-native type III domain. ER quality control mechanisms recognize the abnormal structures in mutant COMP molecules and retain COMP pentamers in the rER. It is possible that

COMP specific molecular escorts are unable to export the mutant COMP molecules out of the rER. However, the association of chaperones known to retain misfolded proteins in the rER with the mutant COMP suggests that anchoring of COMP in the rER, rather than the absence of forward transport leads to COMP retention within the rER. This conclusion is further strengthened by the finding by Dinser *et al.* (Dinser, Zaucke et al. 2002) that primary chondrocytes are able to secrete at least a small proportion of mutant COMP molecules.

2) **Mutated COMP molecules cannot be degraded by cellular proteases:** That the ER proteases cannot efficiently degrade the mutant molecules was shown by the studies in our lab (unpublished data). There are two possibilities as to why the retained COMP molecules are not degraded by the ubiquitin-proteosome pathway. First, the bouquet shaped COMP pentamers may be too large to be retrotranslocated out of the Sec 61p pore for proteosomal degradation. Second, although the mutated COMP molecules may be recognized by the ER retention systems, they may not be unstructured enough to be recognized as targets for proteosomal degradation. Alternatively, they may be restructured into conformations that are not the substrates for the ERAD targeting machinery. We have not been able to rule out any of the possibilities.

3) **Retained COMP molecules need to reach threshold concentrations to associate with themselves or with other ECM proteins:** Tissues that secrete low levels of COMP or have shorter intracellular processing times do not retain COMP. The disorder manifests only when COMP biosynthesis is rapidly increased enhanced. Both these findings imply that high levels of COMP in the cell are critical for forming the fingerprint aggregates. Though it may be difficult to measure the concentration dependence of

fingerprint inclusion formation in cell systems, dynamic light scattering and ultracentrifugation studies would give indication whether mutant COMP molecules undergo self-associations at high concentrations.

4) **Premature interactions between the COMP and type IX collagen cause the formation of the fingerprint aggregates:** The mutation had very little, if any effect on the integrity of the C-terminal domain. As a result it has the structure which allows it to bind normally with its ligands. The very regular nature of the retained structures, along with the presence of at least one known protein (Type IX collagen) which binds to COMP in the ECM suggests that these are specific interactions occurring prematurely in the rER. The cells that do not synthesize type IX collagens secrete mutant COMP. This emphasizes necessity of COMP-type IX collagen interactions in PSACH phenotype.

5) **Abnormal retention of proteins affects cellular metabolism:** The rER of the affected chondrocytes is extensively dilated, occupies a very large proportion of the cell volume and causes the cells to be highly enlarged. Many chondrocytes contain two nuclei which suggests that cell division is affected. Though some normal protein transport does occur in the affected rER, the abnormal ER is almost completely occupied by the insoluble fingerprint structures. This suggests that not only is the protein processing by rER impaired, but normal cell functioning is probably hindered. The accumulation of proteins results in cell death and an overall reduction in the number of viable cells in these tissues (Hecht, Montufar-Solis et al. 1998).

6) **Absence of ECM proteins in the cartilage matrix may disrupt its architecture:** Absence/low levels of COMP, some ECM proteoglycans and non-fibrillar collagens in the ECM lead to disruption of cartilage ECM. This suggests COMP mutations cause

impaired chondrocyte functions as well as abnormal formation of the cartilage ECM. This would result in abnormal cartilage structure and development, causing the PSACH phenotype.

These studies have given us an indication on the pathology of PSACH and further studies are needed to confirm some of the hypothesis suggested. Further *in vitro* studies need to be done to look at the process of formation of the inclusions.

APPENDIX

J Biol Chem, Vol. 276, Issue 44, 40537-40544, November 2, 2001

The Basic Helix-Loop-Helix Domain of the Aryl Hydrocarbon Receptor Nuclear Transporter (ARNT) Can Oligomerize and Bind E-box DNA Specifically*

Joy L. Huffman[§¶], Asawari Mokashi[§], Hans Peter Bächinger[**], and Richard G. Brennan[*]

From the Department of Biochemistry and Molecular Biology, Oregon Health Sciences University and Portland Shriners Hospital, Portland, Oregon 97201
Received for publication, June20, 2001, and in revised form, August 9, 2001

All the experimentation in this paper has been performed by Asawari Mokashi except anisotropy measurements at lower salt concentrations and dynamic light scattering which were performed by Joy Huffman.

Abstract

The aryl hydrocarbon receptor nuclear transporter (ARNT) is a basic helix-loop-helix (bHLH) protein that contains a Per-Arnt-Sim (PAS) domain. ARNT heterodimerizes *in vivo* with other bHLH PAS proteins to regulate a number of cellular activities, but a physiological role for ARNT homodimers has not yet been established. Moreover, no rigorous studies have been done to characterize the biochemical properties of the bHLH domain of ARNT that would address this issue. To begin this characterization, we chemically synthesized a 56-residue peptide encompassing the bHLH domain of ARNT (residues 90-145). In the absence of DNA, the ARNT-bHLH peptide can form homodimers in lower ionic strength, as evidenced by dynamic light scattering analysis, and can bind E-box DNA (CACGTG) with high specificity and affinity, as determined by fluorescence anisotropy. Dimers and tetramers of ARNT-bHLH are observed bound to DNA in equilibrium sedimentation and dynamic light scattering experiments. The homodimeric peptide also undergoes a coil-to-helix transition upon E-box DNA binding. Peptide oligomerization and DNA affinity are strongly influenced by ionic strength. These biochemical and biophysical studies on the ARNT-bHLH reveal its inherent ability to form homodimers at concentrations supporting a physiological function and underscore the significant biochemical differences among the bHLH superfamily.

INTRODUCTION

The aryl hydrocarbon receptor nuclear transporter (ARNT)(Zhulin, Taylor *et al.* 1997) protein belongs to the basic-helix-loop-helix Per-Arnt-Sim (bHLH PAS) family of transcriptional regulator proteins. These functionally oligomeric proteins are important for cell cycle and developmental regulation and for sensing and responding to environmental conditions (Zhulin, Taylor *et al.* 1997; Crews 1998) ARNT shows high sequence homology to other bHLH motifs of this family ((Antonsson, Arulampalam *et al.* 1995; Bacsi and Hankinson 1996)), particularly at residues known to contact DNA (Ferre-D'Amare, Prendergast *et al.* 1993; Ellenberger, Fass *et al.* 1994; Ferre-D'Amare, Pognonec *et al.* 1994; Ma, Rould *et al.* 1994; Shimizu, Toumoto *et al.* 1997; Parraga, Bellsolell *et al.* 1998). In general, bHLH domains bind a consensus DNA element, the so-named E-box (CANNTG), and are required for oligomerization. PAS domains, which are found in all kingdoms, are involved in protein-protein interactions and ligand/inducer binding, acting as environmental sensors (Zhulin, Taylor *et al.* 1997) PAS-containing proteins typically have two such conserved, repeated domains that are separated by a spacer region. PAS domains are not always contiguous with the bHLH DNA binding domain.

ARNT heterodimerizes *in vivo* with other bHLH PAS proteins, including the aryl hydrocarbon receptor (AHR) and hypoxia-inducible factor 1 (HIF1), to form activated DNA binding complexes (Gradin, McGuire *et al.* 1996). Formation of the AHR/ARNT heterodimer requires the binding of polycyclic and halogenated aromatic hydrocarbons such as 2,3,7,8-tetrachlorodibenzo-*p*-dioxin, which are known exogenous ligands for AHR and mediate carcinogenesis via AHR·ARNT complex activation (reviewed in

Ref.(Hankinson 1995). The resulting AHR·ARNT complex binds an atypical E-box DNA sequence, TNGCGTG, thereby activating the transcription of a number of target genes by direct interaction with general transcription factors, such as transcription factor IIB (Swanson and Yang 1998). Among the activated genes are CYP1A1 and CYP1A2, cytochrome oxidases that metabolize polycyclic and aromatic compounds to electrophilic derivatives, which are the ultimate chemical agents that attack DNA. The AHR/ARNT heterodimer also activates transcription of the *mdr1* multidrug transporter, albeit indirectly (Mathieu, Lapierre *et al.* 2001). The HIF1/ARNT heterodimer (HIF1) senses the oxygen tension in cells and, under hypoxic conditions, activates the transcription of a number of genes, the promoters of which contain the E-box sequence, TACGTGCT. Transcription is activated by formation of a HIF1-(CREB (cAMP-response element-binding protein)/ATF1)-(p300/CBP (CREB-binding protein)) complex on the cognate DNA (Ebert and Bunn 1998). Among the genes regulated are erythropoietin, vascular endothelial growth factor, glycolytic enzymes, tyrosine hydroxylase, inducible nitric-oxide synthase, and heme oxygenase-1, all of which allow the cell to cope with lower oxygen levels (reviewed in Ref.(Guillemin and Krasnow 1997)). HIF1 also plays a role in iron homeostasis by its activation of the ceruloplasm gene (Mukhopadhyay, Mazumder *et al.* 2000). This HIF1-mediated response has been found to be crucial for angiogenesis and solid tumor formation ((Maxwell, Dachs *et al.* 1997; Carmeliet, Dor *et al.* 1998; Ryan, Lo *et al.* 1998)). Neither AHR or HIF1 homodimers nor AHR/HIF1 heterodimers have been observed.

In contrast to the *in vivo* importance of AHR/ARNT and HIF1/ARNT heterodimers, the biological relevance of ARNT homodimers is unclear. No physiological role for ARNT

homodimers has yet been defined, and *in vitro* coimmunoprecipitation studies have been unable to detect the homodimeric ARNT complex (Reisz-Porszasz, Probst *et al.* 1994). However, *in vivo* reporter gene assays have demonstrated that putative ARNT homodimers can activate transcription via E-box binding (Antonsson, Arulampalam *et al.* 1995),(Sogawa, Nakano *et al.* 1995), and a preferred DNA binding site has been identified that contains the E-box sequence (Swanson and Yang 1999). It is expected that ARNT homodimers would bind to the CACGTG E-box as do other canonical bHLH proteins, whereas ARNT in one of its heterodimeric complexes would bind one half-site of an asymmetric consensus binding site, with the heterodimeric partner binding the other, nonconsensus half-site. As demonstrated for the AHR/ARNT heterodimer, ARNT is located on the GTG E-box half-site, and AHR is situated on the less restrictive (A/C)(G/C/T)(A/T) non-E-box half-site (Bacsi, Reisz-Porszasz *et al.* 1995).

The structures of several bHLH domain-containing peptides bound to DNA have been determined by x-ray crystallography, including MAX, USF, MyoD, Pho4, E47, and SREBP1 (Ferre-D'Amare, Prendergast *et al.* 1993; Ellenberger, Fass *et al.* 1994; Ferre-D'Amare, Pognonec *et al.* 1994; Ma, Rould *et al.* 1994; Shimizu, Toumoto *et al.* 1997; Parraga, Bellsolell *et al.* 1998). From these structures, it has been determined that the basic region is helical, and its residues make the primary contacts to the DNA, whereas the helix-loop-helix region is largely responsible for dimerization. Additionally, synthetic peptides of several bHLH domains have been characterized biochemically and biophysically, of which the Deadpan bHLH was perhaps the most rigorously investigated (Anthony-Cahill, Benfield *et al.* 1992; Bishop, Jones *et al.* 1995; Muhle-Goll, Nilges *et al.* 1995; Kunne and Allemann 1997; Wendt, Thomas *et al.* 1998; Winston, Millar *et al.*

1999). However, these latter studies do not address the biochemical and biophysical properties of the bHLH-PAS family members, in particular ARNT.

To delineate the structural mechanism of transcription regulation by ARNT either as a homodimer or an AHR/ARNT or HIF1/ARNT heterodimer, we have undertaken a series of biophysical and biochemical studies on a chemically synthesized, 56-residue peptide that encompasses the ARNT-bHLH domain. Specifically, we have used circular dichroism (CD) to determine the extent to which the peptide is folded in both the presence and absence of DNA and equilibrium sedimentation to determine its oligomerization state. Furthermore, using fluorescence anisotropy, we have determined the binding affinity of this peptide for E-box DNA under a variety of experimental conditions. As a complement to our sedimentation equilibrium experiments, we conducted dynamic light scattering studies to evaluate the oligomerization state and monodispersity of the peptide at higher concentrations. Unexpectedly, significant biochemical differences from other bHLH proteins such as Deadpan and Tal were found. Moreover, the data provide evidence that the bHLH of ARNT can form homodimers, which might have biological relevance.

EXPERIMENTAL PROCEEDURES

Solid Phase Peptide Synthesis

A peptide encompassing residues 90-145 of ARNT (ARNT-bHLH) was synthesized using Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry on a Milligen/Biosearch peptide synthesizer. Cleavage and deprotection reactions were carried out in trimethylsilane bromide, ethanediol, *m*-cresol, thioanisol, and trifluoroacetic acid for 1 h at 0 °C under a

nitrogen blanket to ensure that cysteine residues remained reduced after removal of the trityl protection groups. The cleaved peptide was filtered through a medium-sintered glass filter to separate it from resin. The peptide was then washed with trifluoroacetic acid. The filtrate and washings were combined, and all liquid was evaporated using a rotary evaporator. The peptide was precipitated with diethyl ether and filtered through a medium-sintered glass funnel. The peptide was dried under a stream of nitrogen, dissolved in 10% acetic acid, and lyophilized. The peptide was then dissolved in 0.1% trifluoroacetic acid and purified on a Vydac C-18 reverse-phase HPLC column with a mobile phase of 0.1% trifluoroacetic acid and a linear 60-min gradient of 0-100% acetonitrile. Chromatography runs were recorded with a diode array detector and analyzed using Millennium 2000 software (Waters). The peptide eluted at 73% acetonitrile. Fractions from 70 to 74% were pooled and rechromatographed. The purity of the peptide was ascertained by mass spectrometry (data not shown).

Circular Dichroism

Circular dichroism experiments were performed with a Jasco-J500 instrument. An E-box-containing oligonucleotide, ARNDNA (5'-GGCTCAGTCACGTGACTGAGC-3'), was purchased from Oligos, Etc. This sequence was chosen to contain the consensus RTCACGTGAY sequence determined to be recognized by ARNT using a site affinity and amplification assay (Swanson and Yang 1999). The oligonucleotide, which has an unpaired 5' guanosine upon duplex formation, was resuspended from the lyophilized pellet in 10 mM sodium cacodylate, pH 6.5, such that the final concentration of single-stranded oligonucleotide was 2.0 mM. The concentration was calculated using data that

were provided by Oligos, Etc. for each strand. The palindromic strands were annealed by heating to 80 °C, followed by slow cooling to 25 °C. Complete annealing was confirmed by high pressure liquid chromatography and gel electrophoresis (data not shown). The final concentration of duplex DNA was determined by measuring the absorbance of the solution at 260 nm and using the equation $A = cl$, where A is the absorbance at 260 nm, is the known extinction coefficient for double-stranded DNA, 0.02 μg^{1} cm^{1} (Sambrook 1989), c is the concentration of DNA in μg/ml, and l is the path length through the cuvette, 1 cm. The molecular weight of the duplex oligonucleotide (13.0 kDa) was subsequently used to calculate the molar concentration of ARNDNA. The CD spectra of both 100 μM ARNT-bHLH peptide, calculated for a monomer, and 50 μM duplex ARNDNA in Buffer A (50 mM NaCl or 50 mM NaF, 20 mM Tris, pH 7.4) were measured from 260 to 180 nm at 20 °C in a 0.01-cm cell. The substitution of NaF for NaCl had no impact on the spectra (data for NaCl not shown). The spectrum of 100 μM ARNT containing 50% trifluoroethanol (TFE) or 50 μM duplex ARNDNA oligonucleotide was also measured. The final reported spectra are averages of 10 runs. To measure any changes in peptide secondary structure upon binding DNA, the CD difference spectrum was calculated by subtracting the spectrum of ARNDNA alone from that of the ARNDNA/ARNT mixture. The concentration of the peptide solution was verified by amino acid analysis. The secondary structures were analyzed using the variable selection method (Compton, Mathews *et al.* 1987).

DNA Binding at Lower Salt Concentrations; Fluorescence Polarization

Fluorescence polarization experiments were done with a PanVera Beacon fluorescence polarization system (PanVera Corp.). 5'-Fluoresceinated oligonucleotides corresponding to ARNDNA and the *Escherichia coli purF* operator (Oligos, Etc.) (ARNDNA, 5'-F-GCTCAGTCACGTGACTGAGCCCCTTGCTCAGTCACGTGACTGAGC-3', *purF* 5'-F-AAAGAAAACGTTTGCGTACCCCCTACGCAAACGTTTTCTTT-3') were self-annealed in 10 mM sodium cacodylate, pH 6.5, by heating to 80 °C followed by flash cooling to form a stem-loop structure with the E-box motif (underlined) or *purF* operator motif, respectively, at the center of the stem. Oligonucleotide concentrations were calculated as described for ARNDNA used in CD experiments. Binding was assayed in a 1-ml volume at 25 °C. Unless otherwise noted in the text, the components of each binding experiment were 2 nM fluoresceinated DNA and 1.0 μg/ml poly(d[I·C]) in Buffer A (50 mM NaCl or NaF, 20 mM Tris, pH 7.4). Poly(d[I·C]) (Sigma) was included as a control for nonspecific DNA binding. It is expressed in μg/ml rather than molar to reflect the fact that the exact length of the poly(d[I·C]) molecules is not discrete but averages between 1200 and 3000 base pairs. In one set of experiments, the amount of poly(d[I·C]) included was varied to 0.0, 0.1, or 1.0 μg/ml. After each addition of peptide, samples were incubated in the Beacon instrument at 25 °C for 30 s before a measurement was taken. The 30-s incubation allowed equilibrium to be reached. The millipolarization ($P \times 10^3$ where P is polarization) at each titration point represents the average of eight measurements integrated over 6 s. Samples were excited at 490 nm, and emission was measured at 530 nm.

The data of each binding isotherm were analyzed by curve fitting using SigmaPlot software (Jandel Corp.). Because the calculated dissociation constants were all greater

than 20 nM and the experimental DNA concentration was 10-fold less than this value, it was assumed that the concentration of protein bound to DNA was negligible in comparison with the total protein concentration. Therefore, the following equation could be applied (Lundblad, Laurance *et al.* 1996),

$$P = ((P_{bound} - P_{free})[ARNT]/K_d + [ARNT])) + P_{free} \quad \text{(Eq. 1)}$$

where P is the polarization measured at a given total concentration of peptide ([ARNT]), P_{free} is the initial polarization of the free DNA, and P_{bound} is the maximum polarization of specifically bound DNA. Nonlinear least squares analysis was used to determine P_{free}, P_{bound}, and K_d.

DNA Binding at Higher Salt Concentrations; Fluorescence Anisotropy

Anisotropy studies at higher salt concentrations (Buffer B: 150 mM NaCl, 100 mM Tris, pH 7.4) were done with an SLM 8000 spectrofluorometer with T optics at 25 °C. The sample was excited at 480 nm, and the parallel and perpendicular polarization components of fluorescein emission were measured at 520 nm. Except for the buffer, the components in the assay were the same as those used in measuring anisotropy in lower salt buffer. After each addition of peptide, the solution was incubated for 1 min to attain equilibrium. Each titration point is an average of 12 measurements, each integrated over 30 s. Data were analyzed using Scientist MicroMath software. Anisotropy was measured as the function of concentration of ARNT-bHLH (μM) added to the binding reaction, and the data were fit using a two-step binding model that follows. Anisotropy, A, is related to polarization, P, by $A = 2P/(3\ P)$ ((Lundblad, Laurance *et al.* 1996)).

Step 1; Cooperative Binding of ARNT-bHLH to ARNDNA-- The equilibrium model that best describes the initial phase of the higher salt binding isotherm is two ARNT-bHLH monomers binding to the ARNDNA duplex, *i.e.* 2 ARNT-bHLH monomers + ARNDNA ARNT-bHLH·ARNDNA + ARNT-bHLH ARNT-bHLH dimer·ARNDNA. The equation to describe this cooperative binding is as follows,

$$\nu_1 = C^P/C_{HALF}{}^P/C_{HALF}{}^P + C^P \qquad \text{(Eq.2)}$$

where $_1$ is the average number of bound ARNT-bHLH molecules per molecule of ARNDNA, C is the peptide concentration (μM), and C_{HALF} is the concentration at half-saturation of the cooperative binding event. P, the Hill coefficient, is the average number of interacting sites on ARNDNA, which is two. P is also the measure of cooperativity in a binding event.

Step 2; Noncooperative Oligomerization of ARNT-bHLH Dimer on ARNT-bHLH Dimer-ARNDNA-- The following equilibrium model describes the second phase of binding: 1 ARNT-bHLH dimer·ARNDNA + 1 ARNT-bHLH dimer 2 ARNT-bHLH dimers·ARNDNA. In noncooperative binding, all binding sites are equivalent and independent of each other. The binding curve is a rectangular hyperbola (van Holde 1985) and is defined by the equation,

$$\nu_2 = K[ARNT]/(1+K[ARNT] \qquad \text{(Eqn.3)}$$

where ν_2 is the average number of ARNT-bHLH dimers bound to the ARNT-bHLH dimer·ARNDNA complex formed in the first step. K is the association constant of the noncooperative event, and [ARNT] is the concentration of free ARNT-bHLH monomers.

At high concentrations (>23 μM), ARNT-bHLH is dimeric. In this step of binding, only the dimer-bound DNA is available for further binding of ARNT.

Dynamic Light Scattering

Dynamic light scattering studies were done using a DynaPro-801 instrument (Protein Solutions, Inc.). All solutions were filtered through 0.1-μm Anotop filters (Whatman) to remove aggregated peptide and other particulates. Scattering of the ARNT-bHLH peptide was analyzed at concentrations of 5.0 and 10.0 mg/ml (0.78 and 1.56 mM ARNT-bHLH monomer, respectively) in its storage buffer (40 mM KCl, 2 mM dithiothreitol, 0.4 mM EDTA, 5% glycerol, and 20 mM Tris, pH 7.4). Peptide-DNA experiments were performed by mixing 1.56 mM ARNT-bHLH peptide with 1.0 mM oligonucleotide (in 10 mM sodium cacodylate, pH 6.5), resulting in final concentrations of 0.78 and 0.50 mM each molecule, respectively. Under these experimental conditions, there is stoichiometric binding (data not shown). Reported scattering values are the averages of at least 25 scans of 30 s each. All data were analyzed using AutoPro 4.0 PC software (Protein Solutions, Inc.).

Dynamic light scattering uses the Brownian motion of molecules in solution, which causes scattered light intensity to fluctuate (Schurr 1977). These fluctuations are measured by the DynaPro instrument at 20 different times between 3 and 3000 μs. An exponential decay function is generated from the scattering data. The rate of decay of this function is used to determine the translational diffusion coefficient, D_T. The radius of hydration, R_H, is then calculated using the following $D_T = kT/6\ R_H$, where k is the Boltzmann constant, T is temperature in Kelvin, and is the solvent viscosity. R_H is defined as the radius of a hypothetical hard sphere that diffuses with the same speed as the particle

under examination. However, macromolecules are non-spherical and solvated. Therefore, the molecular weight (M_r) of a macromolecule is estimated using M_r *versus* R_H calibration curves developed from standards of known molecular weight and size. Thus, the M_r estimate of a given particle is subject to error if it deviates from the shape and solvation of the molecules used as standards. The molecular weight for protein macromolecules is estimated from a curve that fits the equation $M_r = [1.6800 \cdot R_H]^{2.3398}$, as implemented in the AutoPro software. A similar calibration curve for oligonucleotides has not yet been developed and, hence, precludes an analysis of the ARNDNA alone.

Analysis of Peptide-DNA Complexes by Sedimentation Equilibrium Studies

Sedimentation equilibrium measurements were done with a Model E analytical ultracentrifuge (Beckman Corp.) at 25 °C. A solution of 5 μM ARNDNA and 10 μM ARNT-bHLH was diluted 1:1 and 1:2 in Buffer B (150 mM NaCl, 100 mM Tris, pH 7.4). Buffer B was placed in the reference compartment, and sample solutions were loaded into sample compartments of double sector cells. The initial experiment was run at 20,000 rpm for 24 h because the heaviest species expected was a 38-kDa complex of an ARNT-bHLH tetramer bound to double-stranded DNA. The distribution of the components at equilibrium was determined by measuring the UV absorption of each cell at 280 nm. However, at this wavelength we were unable to differentiate between peptide or DNA alone and peptide-DNA complexes. Given the inability to distinguish DNA from poly(d(I·C]) peptide at 280 nm, we used 5'-fluorescein-labeled DNA in some studies. Samples with 10 μM fluoresceinated ARNDNA and 40 μM ARNT-bHLH peptide were allowed to equilibrate for 20 h at 20,000 rpm. The samples were monitored at 494 nm, the

absorption maximum of fluorescein. Only fluoresceinated DNA and peptide-DNA complexes are observed at this wavelength, without interference from free peptide.

Data acquisition and analyses were done with Ultrascan software (Borries Demeler), and Scientist (MicroMath) was used for experimental data fitting. Weight average molecular weight for a single species model was calculated using the formula. ;

$$C = C_0 + (C_M - C_0) \cdot e^{MA(1-\nu\rho)\,\omega 2(r2-rM2)/2RT} \qquad \text{(Eq.4)}$$

The value *MA* represents the average molecular weight of all sedimenting species. r_M is the distance from the meniscus to the axis of the rotor, r is the distance between each point along the concentration gradient and the rotor axis, and C_0 and C_M are the concentrations at points r and r_M, respectively. ω, ν, ρ, $(1-\nu\rho)$, R, and T represent angular velocity, partial specific volume, density, buoyancy, gas constant, and temperature, respectively.

RESULTS

Polypeptide Synthesis and Circular Dichroism-- We synthesized a 56-aminoacyl residue polypeptide that corresponds to the bHLH DNA binding domain of the ARNT protein as determined by amino acid sequence homology with other bHLH proteins (Ferre-D'Amare, Prendergast *et al.* 1993; Ferre-D'Amare, Pognonec *et al.* 1994; Ma, Rould *et al.* 1994) Shimizu, Toumoto *et al.* 1997) (Ellenberger, Fass *et al.* 1994) (Parraga, Bellsolell *et al.* 1998) and Fig.1. CD spectra were measured to determine the secondary structure content of the ARNT-bHLH peptide and to observe any changes in secondary structure upon binding to DNA. CD spectra were also taken in the presence of TFE. TFE increases helicity of polypeptides by selectively destabilizing solvent-amide group interactions.

Compact conformations such as helices, which maximize intramolecular polypeptide backbone hydrogen bonding and lessen solvent exposure are favored. Medium-sized peptides with an intrinsic tendency to assume a helical conformation in water show an increase in helicity upon the addition of TFE. Hence, the CD spectra of the peptide with and without 50% TFE were measured to affirm that the synthetic peptide had the ability to adopt a helical structure. In both cases, spectra were obtained with strong maxima at 190 nm and double minima at 200-210 and 222 nm, characteristic of helices. The amplitude of the spectrum for ARNT-bHLH with TFE was approximately three times that of peptide in buffer alone (Fig.2A). Secondary structure analysis using the variable selection method showed that the helicity increased from 12.8 to 68.0% upon the addition of TFE (Table I). The predicted maximum helicity attainable by the peptide is 82.1%. This maximum is calculated by assuming that the basic region, helix 1, and helix 2 (Fig.1) would be completely -helical, as observed in the crystal structures of other bHLH peptides bound to DNA.

The bHLH family of transcription factors binds to specific DNA sequences primarily as dimers and tetramers (Phillips 1994). bHLH transcription factors are known to recognize cognate DNA by means of an intrinsically flexible basic region that forms an helix upon binding to cognate E-box DNA. The helix-loop-helix region is involved in dimerization and might undergo structural transitions as well, albeit smaller. Therefore, the CD spectrum of ARNT-bHLH peptide and duplex ARNDNA in 2:1 stoichiometry (monomer peptide-DNA) was measured, and the CD spectrum of ARNDNA alone was subtracted from this to obtain a difference spectrum representing peptide structure in the presence of

DNA (Fig. 2*B* and TableI). The secondary structure analysis showed a 5-fold increase in the helical content of the peptide (from 12.8 to 65.0%) and a 3-fold decrease (40.1 to 12.9%) in the random coil structure in the presence of ARNDNA. The percentage turn also decreased 3-fold, from 39.2 to 12.5%. The random coil-to- helix transition correlates with the increased helical content and decreased random coil content observed on association of ARNT-bHLH with the ARNDNA oligonucleotide. Crystal structures of bHLH peptides bound to cognate DNA have shown that the DNA is not appreciably bent as a result of peptide binding. Oligonucleotide base stacking is a major contributor to optical activity in 260-300-nm range. No change in the optical activity was seen in this range on binding ARNT-bHLH, indicating that there is no gross change induced in the structure of the oligonucleotide (data not shown).

Fluorescence Anisotropy-- We used a fluorescence anisotropy binding assay to measure the equilibrium binding affinity of the ARNT-bHLH peptide for oligonucleotides with and without the E-box sequence. Fluorescence anisotropy is a straightforward technique for directly measuring macromolecular interaction in solution (reviewed in (Lundblad, Laurance *et al.* 1996)), and hence, the effects of a number of variables can be tested readily. The technique is based on the observation that the rotational motion of a fluoresceinated oligonucleotide is slowed by peptide binding, thus increasing the measured anisotropy (or polarization) of the DNA. It is assumed that peptide binding to DNA is directly proportional to the increase in anisotropy if the temperature and viscosity of the solution are constant. This is confirmed by the dynamic light scattering and equilibrium sedimentation studies reported later in which no large, nonspecific aggregate particles were observed.

Specificity of DNA Binding-- The specificity of the ARNT peptide for E-box DNA was determined by comparing the binding isotherms of peptide and E-box DNA (ARNDNA) to peptide and noncognate DNA (*E. coli purF* operator) in 50 mM NaCl, 20 mM Tris, pH 7.4. The resulting binding curves are shown in Figs.3, *A* and *B*. The calculated equilibrium dissociation constant, K_d, is 56.2 ± 11.7 nM for the ARNT-bHLH·ARNDNA complex. The K_d for *purF* operator DNA was calculated to be 1.91 ± 0.73 μM.

Effects of Lower Concentrations of Monovalent and Divalent Salts on DNA Binding-- To characterize DNA binding by the ARNT-bHLH peptide further, the buffer and salt concentrations were varied for a series of ARNT-bHLH peptide titrations into ARNDNA. Representative curves are shown in Figs. 3 and 4. The best binding was observed in the solution containing 50 mM NaCl or 50 mM NaF and 20 mM Tris HCl, pH 7.4 (Fig. 3*A* and 4*A*). In some binding experiments, NaF was used in place of NaCl because it is a more suitable salt for spectroscopic studies that utilize far UV wavelengths, such as circular dichroism. The affinity of the ARNT-bHLH for ARNDNA was not changed by this substitution, with K_d values equal to 56.2 ± 11.7 nM in 50 mM NaCl and 57.2 ± 4.8 nM in 50 mM NaF. We found that the addition of 10 mM $MgCl_2$ to the solution containing 50 mM NaF and 20 mM Tris HCl, pH 7.4, also had no effect on binding (K_d = 56.4 ± 13.9 nM, data not shown).

Effects of Higher Concentrations of Salt on DNA Binding-- Anisotropy measurements of ARNT-bHLH binding to F-ARNDNA in higher salt conditions (150 mM NaCl, 100 mM Tris, pH 7.4) resulted in a complex biphasic curve (Fig. 3*C*). The first phase of the curve

has sigmoidal shape, implying cooperative binding. Such cooperative binding of ARNT-bHLH to ARNDNA is evident to ~23 μM. C_{half}, the peptide concentration at half-saturation, was determined to be 11.7 ± 0.1 μM. The value of the Hill coefficient, *P*, is 1.6 ± 0.1 and indicates that ARNT-bHLH monomers dimerize cooperatively on ARNDNA. The second phase of the binding curve is able to be fit by a rectangular hyperbola, implying noncooperativity. This binding mode has a dissociation constant, K_D, of about 20 μM. However, we cannot elucidate the molecular complexes formed during this second binding event, *i.e.* peptide tetramerization or a second bHLH dimer binding to DNA cannot be discerned from this analysis.

Effects of Poly(d[I·C]) on Equilibrium Binding-- As a control for DNA binding specificity, we included 1.0 μg/ml poly(d[I·C]) in all fluorescence polarization experiments. The use of poly(d[I·C]) ensures that any DNA binding observed at low concentrations of peptide is specific because the high concentration of poly(d[I·C]) offers a huge excess of nonspecific binding sites (on the order of 10 μM) compared with the ARNDNA (2 nM). To determine the effects of poly(d[I·C]) on the measured equilibrium binding constants, additional experiments were conducted in which the poly(d[I·C]) was either removed (0.0 μg/ml) or at a concentration diminished 10-fold (0.1 μg/ml). The lower salt binding buffer, containing 50 mM NaF and 20 mM Tris HCl, pH 7.4, was used in each binding experiment. The resulting binding isotherms are shown in Fig. 4. The equilibrium dissociation constants were $K_d(0.0)$ = 30.0 ± 4.3, $K_d(0.1)$ = 48.1 ± 12.1, and $K_d(1.0)$ = 57.2 ± 4.8 nM. Thus, the ARNT-bHLH binding affinity decreases less than 2-fold as poly(d[I·C]) concentration is increased. The very weak dependence of the K_d of the ARNT-bHLH for ARNDNA on the concentration of poly(d[I·C]) contrasts sharply

with the DNA binding characteristics of the bHLH peptide of Deadpan (Winston, Millar *et al.* 1999). Whereas we observe less than a 2-fold increase in K_d (30-57 nM for 0.0 to 1.0 μg poly(d[I·C]), respectively), Winston *et al.* (Winston, Millar *et al.* 1999) report a 15-fold increase (2.5-37 nM (14.8-fold) for 0.0-0.6 μg of poly(d[I·C]), respectively). The fold differences for the Deadpan peptide would be expected to be even greater at 1.0 μg poly(d[I·C]).

Dynamic Light Scattering-- To determine the oligomeric state of the complex observed in the lower salt fluorescence polarization experiments, we carried out dynamic light scattering studies on high concentrations of free peptide and peptide-DNA mixtures. Dynamic light scattering allows the assessment of the oligomeric state as well as the "dispersity" of a macromolecule or its complex solution (Schurr 1977). Dispersity is the degree to which the particles in a solution are the same size. The results of the dynamic light scattering experiments on the ARNT-bHLH in the presence and absence of ARNDNA are given in Table II. At 5.0 mg/ml (0.78 mM), the ARNT-bHLH peptide (6.5 kDa/monomer) forms a dimer with an apparent molecular weight of 11.1 ± 1.5. At 10.0 mg/ml (1.56 mM), the peptide is also a dimer with apparent molecular weight 13.5 ± 1.0, which is nearly identical to the calculated molecular weight. When solutions of 1 mM duplex ARNDNA oligonucleotide (13.0 kDa) and 1.5 mM ARNT-bHLH are mixed, the average molecular weight of the components in solution is 20.3 ± 4.5. This is consistent with a population containing free peptide dimers (13.0 kDa), free duplex DNA (13.0 kDa), and dimeric peptide-DNA complexes (26.0 kDa). The lower molecular weight average is expected under our experimental conditions because of the molar excess of DNA (0.5 mM) to peptide dimers (~0.4 mM). It is not possible to accurately resolve two

species that differ only 2-fold in molecular weight using this method. However, the change we see upon addition of DNA to the peptide indicates that one peptide dimer binds one duplex DNA molecule at these concentrations. Moreover, the aggregation of components is not significant as judged by the standard deviations of the measured diffusion coefficients (Table II).

Equilibrium Sedimentation Studies-- To investigate the effects of lower peptide concentrations and the higher salt environment on the peptide-DNA oligomerization state, a series of sedimentation studies were done. Using 2.5 μM peptide monomer and 1.25 μM duplex DNA, the analysis yielded complexes with an apparent molecular weight of 19.9 ± 0.2 (Table III). This value corresponds to either one ARNT monomer (6.5 kDa) bound to ARNDNA (13.0 kDa) or to the average molecular weight of free duplex DNA or peptide dimer and one ARNT-bHLH dimer bound to DNA. Under the conditions of 2:1 ARNT-bHLH monomer:ARNDNA stoichiometry at 2-fold higher concentrations (5 μM peptide and 2.5 μM DNA), a 23.0 ± 0.1-kDa species was observed. This is likely the average molecular mass of two species, a 26-kDa complex of one 13-kDa ARNT-bHLH dimer bound to one 13-kDa ARNDNA duplex and the peptide dimer and ARNDNA alone, where the fraction of the 26-kDa species is higher. At higher concentrations (10 μM peptide and 5 μM DNA), a 39.0 ± 0.2-kDa species was observed. Because the absorption wavelength was 280 nm, it was not possible to determine whether this was a tetramer bound to DNA (39 kDa) or a mixture of aggregated states. To differentiate peptide-DNA complexes from free peptide or DNA, fluorescein-labeled ARNDNA was employed, and the absorbance of the sample was monitored at 494 nm, a wavelength at which only the fluoresceinated deoxyoligonucleotide absorbs. At a 4:1 ARNT-bHLH

monomer:ARNDNA stoichiometry (40 μM:10 μM, respectively), a complex with a molecular mass of 43.3 ± 2.1 kDa was observed. This is consistent with an ARNT-bHLH tetramer bound to ARNDNA.

DISCUSSIONS

ARNT is the common heterodimeric partner of a number of bHLH-PAS family members and, thus, plays an essential role in many important pathways (reviewed in (Crews 1998)). Among these, the AHR-mediated pathway is essential for the xenobiotic response (Sogawa, Nakano *et al.* 1995), and the HIF-mediated pathway controls the hypoxic response and, subsequently, hypoxia-mediated apoptosis (Guillemin and Krasnow 1997) and (Bruick 2000)Both pathways are involved also in the development and progression of cancer, where the AHR pathway is the target of action of carcinogenic compounds found in cigarette smoke and Agent Orange(Schecter, Dai *et al.* 1995), and the HIF-mediated hypoxic response is critical for the formation and growth of solid tumors (Ryan, Lo *et al.* 1998) ; (Maxwell, Dachs *et al.* 1997) and (Carmeliet, Dor *et al.* 1998). A chromosomal translocation resulting in the production of a TEL-ARNT fusion protein, which contains the N-terminal domain of TEL and almost all of ARNT including the bHLH domain, has been described (Salomon-Nguyen, Della-Valle *et al.* 2000). This protein contributes to leukemogenesis, likely via dysregulation of ARNT-mediated pathways (Salomon-Nguyen, Della-Valle *et al.* 2000).

To understand the biochemical and hence, physiological, function of these members of the bHLH-PAS family, we synthesized a 56-residue peptide corresponding to the bHLH domain of ARNT. We have characterized this domain biophysically and biochemically to

ascertain its ability to form a homodimer and bind E-box DNA and compared it to other bHLH domains. Although there are similarities between these bHLH domains in their CD spectra and oligomerization properties, there are also distinct differences, particularly in their DNA binding properties. This is notable, especially, given the high degree of amino acid sequence similarity across the bHLH family of proteins and the near structural identities of those members, the structures of which are known. However, none of the bHLH peptides studied to date has been a member of the bHLH PAS family. In light of these studies, the biochemical and biophysical differences between ARNT and bHLH proteins such as Deadpan (Winston, Millar *et al.* 1999) indicate that subtle, but significant, differences exist within the bHLH super family.

As anticipated, the helical content of the ARNT-bHLH peptide increases dramatically in the presence of DNA or TFE. Our CD studies have shown that the peptide obtains its maximum helicity, 68.0%, in TFE. A nearly identical helicity, 65.0%, is seen upon its binding to specific DNA. This can be explained largely by the coil-to-helix transition of the flexible basic region and helix 1, which together account for 50% of the length of the peptide. The -helical content in the presence of DNA is significantly lower than the predicted maximum helicity, 82.1%, and that observed for other bHLH proteins, *e.g.* MyoD-E47, 80%, and Mash-1, 85% (Wendt, Thomas *et al.* 1998 and Kunne, 1997 #176)This difference can be ascribed to either our overestimation of the predicted helical content of the ARNT-bHLH protein or helical fraying. Our experimental result, however, is similar to that reported for Dpn, *i.e.* 63% -helical content in the presence of DNA. Although the helicity of the ARNT-bHLH peptide alone is significantly lower (12.8%) than that observed with other bHLH peptides, *e.g.* MyoD-E47 and Mash-1 (both 50%),

these differences may be explained by different methods used in CD data analysis. Alternatively, the homodimeric bHLH domain of ARNT may be intrinsically more flexible, a property that might be required for its biological function. Further studies on the bHLH domains of ARNT/AHR and ARNT/HIF1 heterodimers will be required to address these questions.

Our DNA binding studies indicate that the ARNT-bHLH homodimer is able to bind the E-box (CACGTG) specifically and with high affinity (56 nM) but not other unrelated DNA sequences. Using fluorescence anisotropy, we have determined the K_d for peptide-DNA binding under a number of experimental conditions. In our lower salt binding buffer, the effect of nonspecific DNA on the K_d is minimal. This is in sharp contrast to results of similar experiments with the Dpn-bHLH peptide. Whereas we observe less than a 2-fold decrease in K_d for ARNT-DNA binding upon decreasing the poly(d[I·C]) concentration from 1.0 to 0.0 μg/ml, the K_d of Dpn for its cognate DNA decreases 15-fold when the poly(d[I·C]) is reduced from 0.6 to 0.0 μg/ml (Winston, Millar *et al.* 1999). The very weak dependence of E-box binding by the ARNT-bHLH on the presence or absence of nonspecific DNA indicates that the ARNT-bHLH peptide binds nonspecific DNA far less readily than the Dpn-bHLH. In further contrast to Dpn, ARNT-bHLH DNA binding is diminished in the presence of 150 mM NaCl and 100 mM Tris HCl but not in moderate amounts (10 mM) of divalent cation. The Dpn peptide binds cognate DNA more specifically (in the presence of poly(d[I·C])) when salt concentrations are increased. Perhaps the different DNA binding properties of these bHLH proteins result in part from their net charge differences. From residue 90 to the N terminus of helix 2 (residue 128), the region expected to interact with DNA, the net charge of ARNT is +3, whereas the

corresponding region of Dpn (residues 41-83) is +11. The less basic nature of the ARNT bHLH might make this polypeptide more sensitive to competition by Na^+ and K^+ ions for its DNA binding site, which suggests an important role for ARNT-phosphodiester backbone contacts for specific binding.

The dynamic light scattering data, obtained under conditions of lower ionic strength, reveal that at high concentrations, ARNT-bHLH is predominantly a dimer and that one dimer binds one duplex DNA molecule. A model for this simple binding mode is presented in Fig. 5*A*. In contrast, the data obtained from the binding isotherms and equilibrium sedimentation studies under conditions of higher salt is fit better by a two-step binding model (Fig.5, *B-C*). In this environment and at peptide concentrations 23 μM, we propose that ARNT-bHLH monomers bind to ARNDNA in a cooperative manner to form dimers on the DNA (Fig. 5*B*). Similar binding modes are proposed for other transcription regulators, including LexA (Kim and Little 1992). At concentrations <15 μM, LexA is a monomer that dimerizes cooperatively on cognate DNA to which it binds with nanomolar affinity. ARNT-bHLH binding differs from LexA binding to DNA, however, as the ARNT-bHLH dimer·ARNDNA complex provides an additional site for the binding of a second ARNT-bHLH dimer. Specifically, at peptide concentrations >23 μM we propose that significant concentrations of dimeric ARNT-bHLH are found and that the dimer may bind to the (2 ARNT-bHLH monomers)·ARNDNA complex to form the 39-kDa (4 ARNT-bHLH)·ARNDNA complex, which is observed in our equilibrium sedimentation experiments. Because we did not see a 32-kDa species ((3 ARNT-bHLH monomers)·ARNDNA) in the ultracentrifugation studies, the second ARNT-bHLH dimer

must bind in a one-step process to the (2 ARNT-bHLH)·ARNDNA complex (Fig. 5*C*).

The higher salt binding mode differs from the high affinity DNA binding of ARNT-bHLH in lower salt concentrations and with the DNA binding of other bHLH proteins (Phillips 1994). However, two simple mechanisms by which two dimers of the ARNT-bHLH bind ARNDNA can be proposed. In one, the increase in peptide structure caused by binding to DNA might allow the formation of a four-helix bundle-like tetramer. This is unlikely, because similar oligomerization phenomena should be expected at lower salt concentrations, and such binding is not observed under that condition (Table II). In the second, a dimer of dimers on the DNA could be stabilized by interactions between the loop regions of each homodimer. Protein-protein cross-linking and peptide sequencing should clarify the issue of tetramerization. It is possible that the sensitivity we see to ionic strength is a result of salt competition with dimerization surfaces that are more hydrophilic than observed for other bHLH domains. Indeed, a theoretical model of the HIF1/ARNT bHLH heterodimer (Michel, Minet *et al.* 2000) places several hydrophilic residues at or near the interface between the domains, unlike the MyoD bHLH interface, which is largely hydrophobic (Ma, Rould *et al.* 1994).

The sedimentation data presented here differ from those studies done on the bHLH of MyoD, which was shown to be dimeric and tetrameric in solution but bound DNA only as a dimer (Muhle-Goll, Nilges *et al.* 1995). However, it is important to note that not every homodimeric bHLH-containing protein is able to bind DNA. A peptide corresponding to the bHLH domain of TalI was shown to form homodimers in solution, but no DNA

binding was (Bishop, Jones *et al.* 1995). However, TalI was able to bind DNA as a heterodimer with the bHLH of E47.

In conclusion, we have carried out biophysical and biochemical studies on the oligomerization and DNA binding properties of the bHLH domain of ARNT that suggest homodimeric ARNT could be a viable transcription regulator *in vivo*. The ARNT-bHLH does homodimerize, as evidenced by our dynamic light scattering studies. We observe high affinity binding of ARNT-bHLH dimers to a specific DNA site at lower ionic strength and the induction of a coil to helix transition in the presence of DNA. At higher ionic strength, the ARNT-bHLH cooperatively dimerizes on ARNDNA, and at even higher concentrations of peptide, an additional ARNT-bHLH dimer is able to bind to the dimeric peptide-duplex DNA complex. This cooperativity may have a biological significance as it was seen in salt concentrations thought to be present in the cell (Scopes 1993). However, the biological significance of ARNT-bHLH tetramerization is unclear since the cellular concentration of ARNT has not been established. Furthermore, the PAS-A domain, which is critical for the heterodimerization of AHR and ARNT and is not present in the bHLH peptide, could play a significant negative or positive role in ARNT homodimerization (Kronenberg, Esser *et al.* 2000). Comparison of the sensitivity to ionic strength or other solution parameters of ARNT homodimers with ARNT heterodimers (*e.g.* ARNT·AHR or ARNT·HIF1) may yield information regarding the transcription regulation activity of each of these complexes. Further biochemical and structural studies to characterize the ARNT·AHR and ARNT·HIF1 interactions will be required to understand the mechanisms of ARNT heterodimerization and its attendant gene regulation.

FOOTNOTES*

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ These authors contributed equally to this work.

¶ Supported by a National Institutes of Health predoctoral training grant and a Tartar Trust Foundation grant.

** Supported by a Shriners Hospital grant.

Supported by National Institutes of Health Grant GM-49244. To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, Oregon Health Sciences University, 3181 S. W. Sam Jackson Park Rd., Portland, OR 97201. Tel.: 503-494-4427; Fax: 503-494-8393; E-mail: brennanr@ohsu.edu.

Published, JBC Papers in Press, August 13, 2001, DOI 10.1074/jbc.M105675200

ABBREVIATIONS

The abbreviations used are: ARNT, aryl hydrocarbon receptor nuclear transporter; bHLH, basic helix-loop-helix; PAS, Per-Arnt-Sim; AHR, aryl hydrocarbon receptor; HIF, hypoxia-inducible factor; Dpn, Deadpan; CD, circular dichroism; TFE, trifluoroethanol.

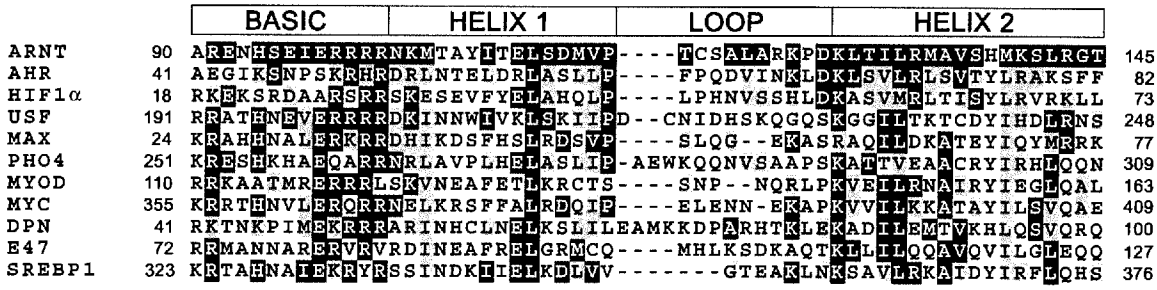


Fig. 1. Primary sequence alignment of selected bHLH domains. GenBank™ accession numbers for each, as well as the percent identity to the ARNT-bHLH, in parentheses, are: ARNT P53762 (100), AHR P41738 (21), HIF1 NP_001521 (21), USF Q07956 (25), MAX P28574 (27), PHO4 P07270 (26), MYOD P15172 (27), MYC P01106 (29), DPN Q26263 (28), E47 B31492 (25), SREBP1 P36956 (33). ARNT, AHR, and HIF1a are members of the bHLH-PAS family. Identical residues are *boxed* in *black*, and homologous residues are *boxed* in *gray*. The first and last residue number of the bHLH domain is given at the beginning and end of each sequence. Crystal structures have been solved for USF (6), MAX (5), PHO4 (8), MYOD (7), E47 (9), and SREBP1 (10). These structures reveal similar three-dimensional folds of bHLH domains. Alignments were performed using ClustalW, and this figure was generated using Boxshade.

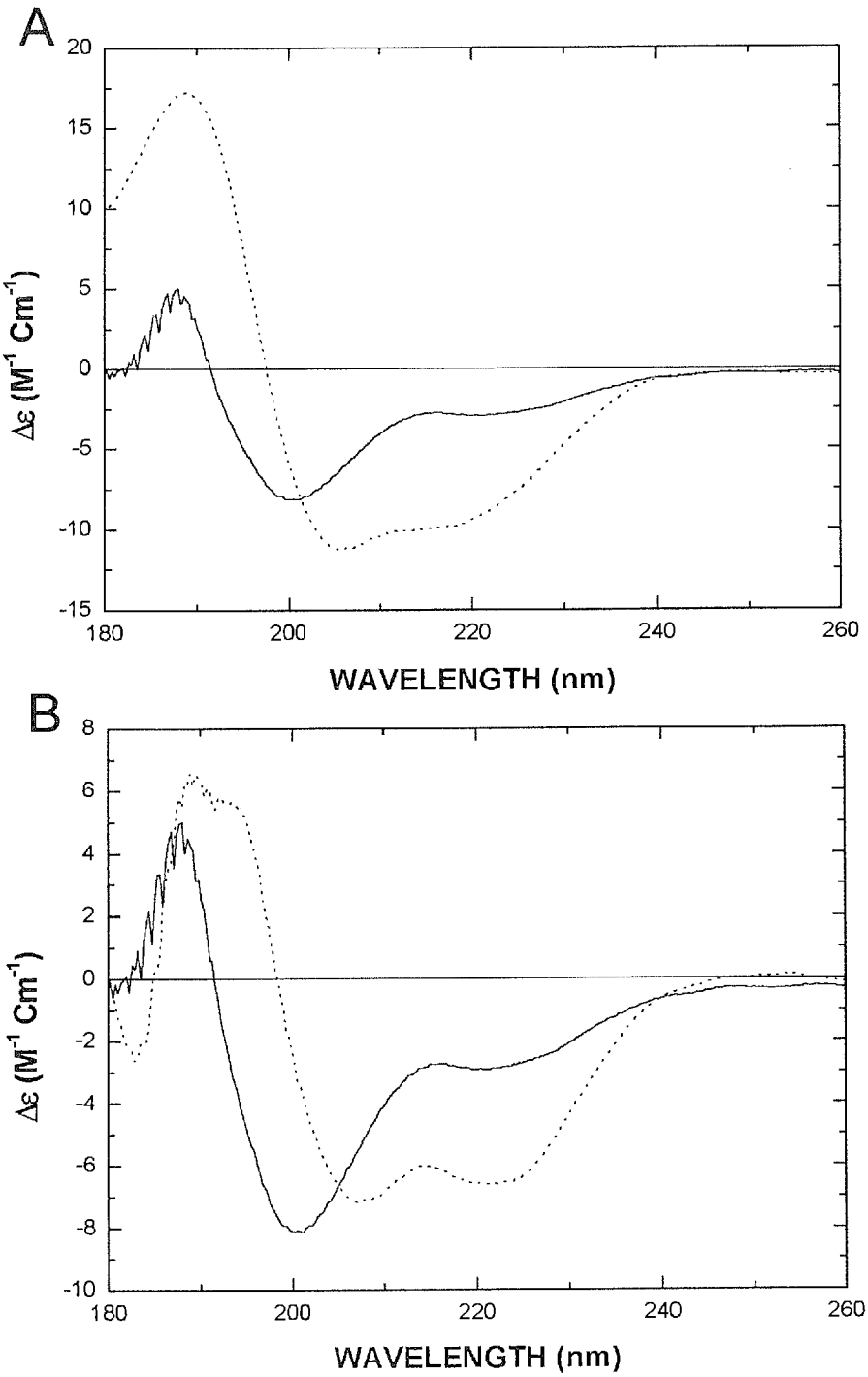


Fig. 2. Circular dichroism spectra of ARNT and ARNT-DNA. *A*, ARNT peptide alone (*solid line*) and in 50% TFE (*dotted line*). *B*, ARNT peptide (*solid line*) and the difference spectrum (*dotted line*) obtained by subtracting ARNDNA spectrum from that of ARNDNA·ARNT complex (1:2 DNA-peptide monomer ratio). The ARNT-ARNDNA spectrum was carried out in 50 mM NaF, 20 mM Tris, pH 7.4.

	%H	%A	% P	% T	% O
ARNT	12.8	11.4	3.4	39.2	40.1
ARNT+50%TFE	68.0	0.5	2.5	18.7	14.2
ARNT+DNA (2:1)	65.0	3.2	3.1	12.5	12.9

Table I
Secondary structure analysis of the CD spectra
H, helix; A, anti-parallel sheet; P, parallel sheet; T, turn; O, other aperiodic structures. Spectra were taken in 50 mM NaF, 20 mM Tris, pH 7.4.

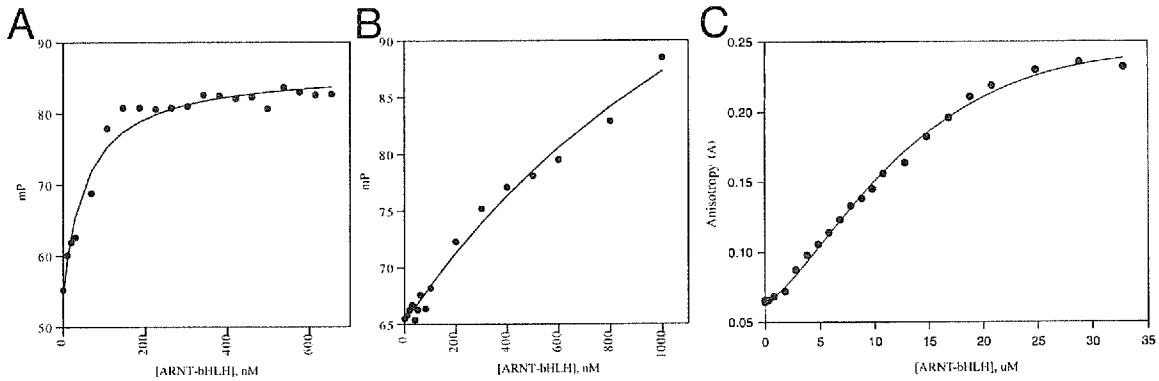


Fig. 3. Binding isotherms of the ARNT-bHLH to fluoresceinated oligodeoxynucleotides. Millipolarization (*mP*) is plotted against the concentration of ARNT-bHLH. *A*, ARNT-bHLH binding to F-ARNDNA in Buffer A. *B*, ARNT-bHLH binding to *purF* operator in Buffer A. *C*, ARNT-bHLH binding to F-ARNDNA in Buffer B.

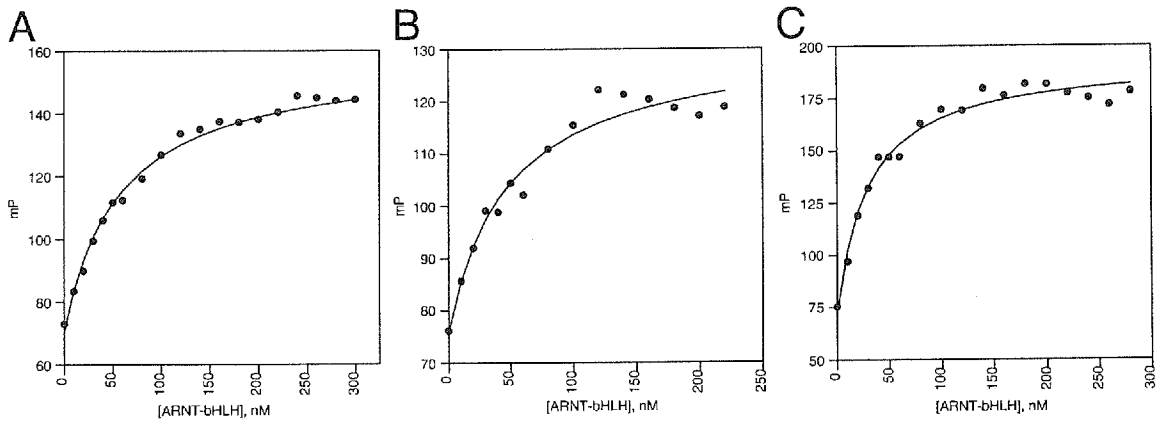


Fig. 4. Effects of poly(d[I·C]) on the binding of ARNT-bHLH to F-ARNDNA. Binding isotherms in the presence of 50 mM NaF, 20 mM Tris, pH 7.4, and 1.0 μg/ml poly(d[I·C]) (*A*), 0.1 μg/ml poly(d[I·C]) (*B*), and no poly(d[I·C]) (*C*). Millipolarization (*mP*).

Table II
Dynamic light scattering analysis of ARNT-bHLH with and without DNA

Polypeptide+/-DNA	Concentration *mM*	**Diffusion coefficient** 10^9 cm^2/s ***kDa***	**Average molecular weight kDa**
ARNT-bHLH	**0.78**	**1265 ± 75**	**11.1 ± 1.5**
ARNT-bHLH	**1.56**	**1167 ± 87**	**13.5 ± 1.0**
ARNT-bHLH + ARNDNA	**0.78 ± 0.50**	**1062 ± 100**	**20.3 ± 4.5**

Table III

Sedimentation equilibrium studies of peptide-DNA complexes

The buffer was 100 mM Tris, pH 7.4. The NaCl concentration was 150 mM in all experiments

Protein-DNA[a]	Average molecular weight *kDa*	Wavelength *nM*	Sedimenting species
2.5 μM ARNT-bHLH + 1.25 μM ARNDNA	19.9 ± 0.2	280	ARNT-DNA or DNA + 2 ARNT-DNA
5 μM ARNT-bHLH + 2.5 μM ARNDNA	23.0 ± 0.1	280	2 ARNT-DNA + DNA + ARNT-DNA + 2 ARNT
10 μM ARNT-bHLH + 5 μM ARNDNA	39.0 ± 0.2	280	4 ARNT-DNA
40 μM ARNT-bHLH + 10 μM ARNDNA[b]	43.3 ± 2.1	494	4 ARNT-DNA

[a] All references to ARNT-bHLH and ARNT are based on monomeric peptide, and all references to DNA correspond to double-stranded ARNDNA oligonucleotide. [b] ARNDNA used in this experiment was fluorescently labeled.

.

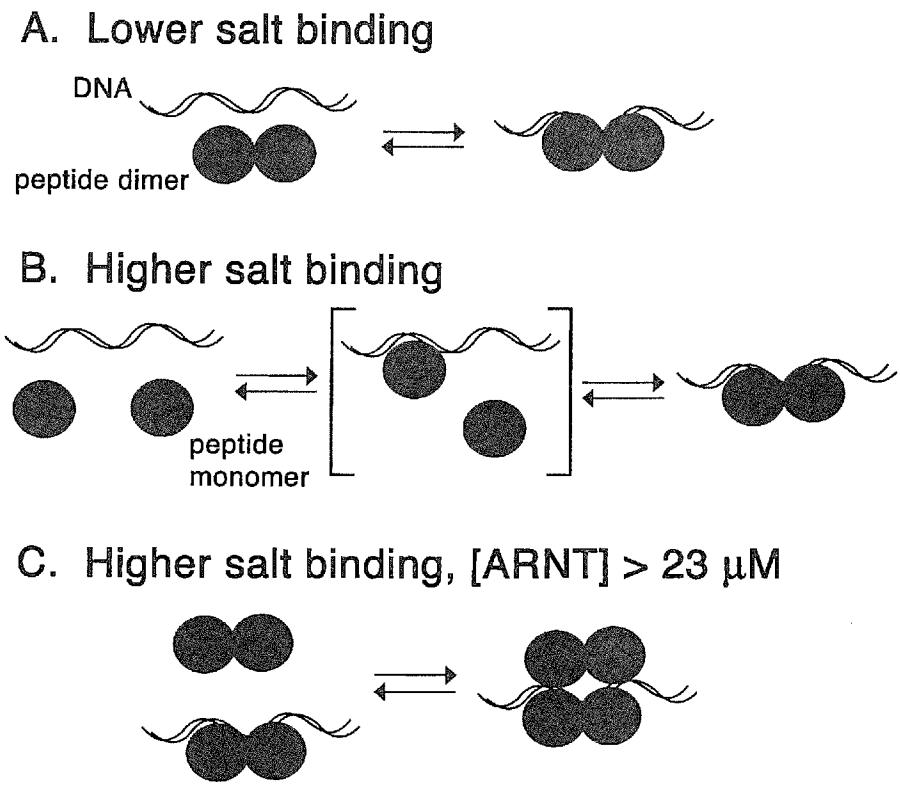


Fig. 5. Proposed model for ARNT-bHLH binding to cognate DNA under lower and higher salt conditions. *A*, In *lower* salt solution (50 mM NaCl, 20 mM Tris, pH 7.4), ARNT-bHLH is predominantly found as a dimer, which is able to bind cognate DNA with high affinity. In higher salt conditions (150 mM NaCl, 100 mM Tris, pH 7.4), ARNT-bHLH binds ARNDNA in a two-step process: *B*, *step 1*: below 23 μM ARNT-bHLH, one peptide monomer binds first to the DNA followed by cooperative recruitment of a second monomer to the DNA. *C*, *step 2*: above 23 μM, ARNT-bHLH is present largely as dimers, and one dimer binds to the preformed 2 ARNT-bHLH monomer-ARNDNA complex.

Literature Cited

Adams, J. C. and J. Lawler (1993). "Diverse mechanisms for cell attachment to platelet thrombospondin." J Cell Sci 104(Pt 4): 1061-71.

Annaert, W. G., B. Becker, et al. (1997). "Export of cellubrevin from the endoplasmic reticulum is controlled by BAP31." J Cell Biol 139(6): 1397-410.

Anthony-Cahill, S. J., P. A. Benfield, et al. (1992). "Molecular characterization of helix-loop-helix peptides." Science 255(5047): 979-83.

Antonsson, C., V. Arulampalam, et al. (1995). "Constitutive function of the basic helix-loop-helix/PAS factor Arnt. Regulation of target promoters via the E box motif." J Biol Chem 270(23): 13968-72.

Argon, Y. and B. B. Simen (1999). "GRP94, an ER chaperone with protein and peptide binding properties." Semin Cell Dev Biol 10(5): 495-505.

Aridor, M. and W. E. Balch (1999). "Integration of endoplasmic reticulum signaling in health and disease." Nat Med 5(7): 745-51.

Bacsi, S. G. and O. Hankinson (1996). "Functional characterization of DNA-binding domains of the subunits of the heterodimeric aryl hydrocarbon receptor complex imputing novel and canonical basic helix-loop-helix protein-DNA interactions." J Biol Chem 271(15): 8843-50.

Bacsi, S. G., S. Reisz-Porszasz, et al. (1995). "Orientation of the heterodimeric aryl hydrocarbon (dioxin) receptor complex on its asymmetric DNA recognition sequence." Mol Pharmacol 47(3): 432-8.

Ballo, R., M. D. Briggs, et al. (1997). "Multiple epiphyseal dysplasia, ribbing type: a novel point mutation in the COMP gene in a South African family." Am J Med Genet 68(4): 396-400.

Bannykh, S. I. a. B., W.E. (1997). "Membrane Dynamics at the Endoplasmic Reticulum-Golgi Interface." JCB vol 138(1): 1-4.

Baron, M., D. G. Norman, et al. (1992). "The three-dimensional structure of the first EGF-like module of human factor IX: comparison with EGF and TGF-alpha." Protein Sci 1(1): 81-90.

Bergeron, J. J., M. B. Brenner, et al. (1994). "Calnexin: a membrane-bound chaperone of the endoplasmic reticulum." Trends Biochem Sci 19(3): 124-8.

Bishop, P., C. Jones, et al. (1995). "Synthesis of the basic-helix-loop-helix region of the immunoglobulin enhancer binding protein E47 and evaluation of its structural and DNA binding properties." Int J Pept Protein Res 46(2): 149-54.

Bogaert, R., G. E. Tiller, et al. (1992). "An amino acid substitution (Gly853-->Glu) in the collagen alpha 1(II) chain produces hypochondrogenesis." J Biol Chem 267(31): 22522-6.

Bonfanti, L., A. A. Mironov, Jr., et al. (1998). "Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation." Cell 95(7): 993-1003.

Bonifacino, J. S. (1996). "Reversal of fortune for nascent proteins." Nature 384(6608): 405-6.

Bonifacino, J. S. and A. M. Weissman (1998). "Ubiquitin and the control of protein fate in the secretory and endocytic pathways." Annu Rev Cell Dev Biol 14: 19-57.

Bornstein, P. and E. H. Sage (1994). "Thrombospondins." Methods Enzymol 245: 62-85.

Briggs, M. D. and K. L. Chapman (2002). "Pseudoachondroplasia and multiple epiphyseal dysplasia: mutation review, molecular interactions, and genotype to phenotype correlations." Hum Mutat 19(5): 465-78.

Briggs, M. D., S. M. Hoffman, et al. (1995). "Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene." Nat Genet 10(3): 330-6.

Briggs, M. D., G. R. Mortier, et al. (1998). "Diverse mutations in the gene for cartilage oligomeric matrix protein in the pseudoachondroplasia-multiple epiphyseal dysplasia disease spectrum." Am J Hum Genet 62(2): 311-9.

Briggs, M. D., I. M. Rasmussen, et al. (1993). "Genetic linkage of mild pseudoachondroplasia (PSACH) to markers in the pericentromeric region of chromosome 19." Genomics 18(3): 656-60.

Brooks, D. A. (1997). "Protein processing: a role in the pathophysiology of genetic disease." FEBS Lett 409(2): 115-20.

Bruick, R. K. (2000). "Expression of the gene encoding the proapoptotic Nip3 protein is induced by hypoxia." Proc Natl Acad Sci U S A 97(16): 9082-7.

Bucciantini, M., E. Giannoni, et al. (2002). "Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases." Nature 416(6880): 507-11.

Buckwalter, J. A. and H. J. Mankin (1998). "Articular cartilage: tissue design and chondrocyte-matrix interactions." Instr Course Lect 47: 477-86.

Bukau, B. and A. L. Horwich (1998). "The Hsp70 and Hsp60 chaperone machines." Cell 92(3): 351-66.

Carmeliet, P., Y. Dor, et al. (1998). "Role of HIF-1alpha in hypoxia-mediated apoptosis, cell proliferation and tumour angiogenesis." Nature 394(6692): 485-90.

Chen, H., M. Deere, et al. (2000). "Cartilage oligomeric matrix protein is a calcium-binding protein, and a mutation in its type 3 repeats causes conformational changes." J Biol Chem 275(34): 26538-44.

Chevet, E., C. A. Jakob, et al. (1999). "Calnexin family members as modulators of genetic diseases." Semin Cell Dev Biol 10(5): 473-80.

Ciechanover, A., A. Orian, et al. (2000). "The ubiquitin-mediated proteolytic pathway: mode of action and clinical implications." J Cell Biochem Suppl 34: 40-51.

Cohn, D. H., M. D. Briggs, et al. (1996). "Mutations in the cartilage oligomeric matrix protein (COMP) gene in pseudoachondroplasia and multiple epiphyseal dysplasia." Ann N Y Acad Sci 785: 188-94.

Cole, W. G. (1997). "The Nicholas Andry Award-1996. The molecular pathology of osteogenesis imperfecta." Clin Orthop(343): 235-48.

Cole, W. G. (2002). "Advances in osteogenesis imperfecta." Clin Orthop(401): 6-16.

Compton, L. A. and W. C. Johnson, Jr. (1986). "Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication." Anal Biochem 155(1): 155-67.

Compton, L. A., C. K. Mathews, et al. (1987). "The conformation of T4 bacteriophage dihydrofolate reductase from circular dichroism." J Biol Chem 262(27): 13039-43.

Cooper, R. R., I. V. Ponseti, et al. (1973). "Pseudoachondroplastic dwarfism. A rough-surfaced endoplasmic reticulum storage disorder." J Bone Joint Surg Am 55(3): 475-84.

Corsi, A. K. and R. Schekman (1996). "Mechanism of polypeptide translocation into the endoplasmic reticulum." J Biol Chem 271(48): 30299-302.

Coux, O., K. Tanaka, et al. (1996). "Structure and functions of the 20S and 26S proteasomes." Annu Rev Biochem **65**: 801-47.

Crews, S. T. (1998). "Control of cell lineage-specific development and transcription by bHLH- PAS proteins." Genes Dev 12(5): 607-20.

Davies, G. B., J. T. Oxford, et al. (1998). "Temporal and spatial expression of alternative splice-forms of the alpha1(XI) collagen gene in fetal rat cartilage." Dev Dyn 213(1): 12-26.

Davis, E. C. and R. P. Mecham (1996). "Selective degradation of accumulated secretory proteins in the endoplasmic reticulum. A possible clearance pathway for abnormal tropoelastin." J Biol Chem 271(7): 3787-94.

Deere, M., T. Sanford, et al. (1998). "Identification of twelve mutations in cartilage oligomeric matrix protein (COMP) in patients with pseudoachondroplasia." Am J Med Genet 80(5): 510-3.

Deere, M., T. Sanford, et al. (1999). "Identification of nine novel mutations in cartilage oligomeric matrix protein in patients with pseudoachondroplasia and multiple epiphyseal dysplasia." Am J Med Genet 85(5): 486-90.

Delot, E., S. G. Brodie, et al. (1998). "Physiological and pathological secretion of cartilage oligomeric matrix protein by cells in culture." J Biol Chem 273(41): 26692-7.

Derkx, P. M. and S. M. Madrid (2001). "The foldase CYPB is a component of the secretory pathway of Aspergillus niger and contains the endoplasmic reticulum retention signal HEEL." Mol Genet Genomics 266(4): 537-45.

Di Cesare, P. E., F. S. Chen, et al. (2002). "Matrix-matrix interaction of cartilage oligomeric matrix protein and fibronectin." Matrix Biol 21(5): 461-70.

Di Cesare, P. E., C. Fang, et al. (1999). "Localization and expression of cartilage oligomeric matrix protein by human rheumatoid and osteoarthritic synovium and cartilage." J Orthop Res 17(3): 437-45.

Di Cesare, P. E., C. Fang, et al. (2000). "Expression of cartilage oligomeric matrix protein (COMP) by embryonic and adult osteoblasts." J Orthop Res 18(5): 713-20.

DiCesare, P. E., M. Morgelin, et al. (1994). "Cartilage oligomeric matrix protein and thrombospondin 1. Purification from articular cartilage, electron microscopic structure, and chondrocyte binding." Eur J Biochem 223(3): 927-37.

Dietz, H. C., G. R. Cutting, et al. (1991). "Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene." Nature 352(6333): 337-9.

Dinser, R., F. Zaucke, et al. (2002). "Pseudoachondroplasia is caused through both intra- and extracellular pathogenic pathways." J Clin Invest 110(4): 505-13.

Downing, A. K., V. Knott, et al. (1996). "Solution structure of a pair of calcium-binding epidermal growth factor- like domains: implications for the Marfan syndrome and other genetic disorders." Cell 85(4): 597-605.

Ebert, B. L. and H. F. Bunn (1998). "Regulation of transcription by hypoxia requires a multiprotein complex that includes hypoxia-inducible factor 1, an adjacent transcription factor, and p300/CREB binding protein." Mol Cell Biol 18(7): 4089-96.

Efimov, V. P., A. Lustig, et al. (1994). "The thrombospondin-like chains of cartilage oligomeric matrix protein are assembled by a five-stranded alpha-helical bundle between residues 20 and 83." FEBS Lett 341(1): 54-8.

Ellenberger, T., D. Fass, et al. (1994). "Crystal structure of transcription factor E47: E-box recognition by a basic region helix-loop-helix dimer." Genes Dev 8(8): 970-80.

Ellgaard, L., M. Molinari, et al. (1999). "Setting the standards: quality control in the secretory pathway." Science 286(5446): 1882-8.

Fang, C., C. S. Carlson, et al. (2000). "Molecular cloning, sequencing, and tissue and developmental expression of mouse cartilage oligomeric matrix protein (COMP)." J Orthop Res 18(4): 593-603.

Ferre-D'Amare, A. R., P. Pognonec, et al. (1994). "Structure and function of the b/HLH/Z domain of USF." Embo J 13(1): 180-9.

Ferre-D'Amare, A. R., G. C. Prendergast, et al. (1993). "Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain." Nature 363(6424): 38-45.

Fessler, L. I. and J. H. Fessler (1974). "Protein assembly of procollagen and effects of hydroxylation." J Biol Chem 249(23): 7637-46.

Fewell, S. W., K. J. Travers, et al. (2001). "The action of molecular chaperones in the early secretory pathway." Annu Rev Genet 35: 149-91.

Fife, R. S. and K. D. Brandt (1984). "Identification of a high-molecular-weight (greater than 400 000) protein in hyaline cartilage." Biochim Biophys Acta 802(3): 506-14.

Fife, R. S., G. L. Hook, et al. (1985). "Topographic localization of a 116,000-dalton protein in cartilage." J Histochem Cytochem 33(2): 127-33.

Fink, A. L. (1999). "Chaperone-mediated protein folding." Physiol Rev 79(2): 425-49.

Fischer, D., R. P. Tucker, et al. (1997). "Cell-adhesive responses to tenascin-C splice variants involve formation of fascin microspikes." Mol Biol Cell 8(10): 2055-75.

Franzen, A., D. Heinegard, et al. (1987). "Evidence for sequential appearance of cartilage matrix proteins in developing mouse limbs and in cultures of mouse mesenchymal cells." Differentiation 36(3): 199-210.

Friedlander, R., E. Jarosch, et al. (2000). "A regulatory link between ER-associated protein degradation and the unfolded-protein response." Nat Cell Biol 2(7): 379-84.

Gardner, A. M., S. Aviel, et al. (1993). "Rapid degradation of an unassembled immunoglobulin light chain is mediated by a serine protease and occurs in a pre-Golgi compartment." J Biol Chem 268(34): 25940-7.

Gething, M. J. (1999). "Role and regulation of the ER chaperone BiP." Semin Cell Dev Biol 10(5): 465-72.

Giannoni, P., M. Siegrist, et al. (2003). "The mechanosensitivity of cartilage oligomeric matrix protein (COMP)." Biorheology 40(1-3): 101-9.

Glover, J. R. and J. M. Tkach (2001). "Crowbars and ratchets: hsp100 chaperones as tools in reversing protein aggregation." Biochem Cell Biol 79(5): 557-68.

Goloubinoff, P., A. Mogk, et al. (1999). "Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network." Proc Natl Acad Sci U S A 96(24): 13732-7.

Gradin, K., J. McGuire, et al. (1996). "Functional interference between hypoxia and dioxin signal transduction pathways: competition for recruitment of the Arnt transcription factor." Mol Cell Biol 16(10): 5221-31.

Guillemin, K. and M. A. Krasnow (1997). "The hypoxic response: huffing and HIFing." Cell 89(1): 9-12.

Guo, Y., R. A. Kammerer, et al. (2000). "The unusually stable coiled-coil domain of COMP exhibits cold and heat denaturation in 4-6 M guanidinium chloride." Biophys Chem 85(2-3): 179-86.

Hagg, R., P. Bruckner, et al. (1998). "Cartilage fibrils of mammals are biochemically heterogeneous: differential distribution of decorin and collagen IX." J Cell Biol 142(1): 285-94.

Hammond, C. and A. Helenius (1995). "Quality control in the secretory pathway." Curr Opin Cell Biol 7(4): 523-9.

Handford, P. A., M. Mayhew, et al. (1991). "Key residues involved in calcium-binding motifs in EGF-like domains." Nature 351(6322): 164-7.

Hankinson, O. (1995). "The aryl hydrocarbon receptor complex." Annu Rev Pharmacol Toxicol 35: 307-40.

Hecht, J. T., E. Hayes, et al. (2001). "Calreticulin, PDI, Grp94 and BiP chaperone proteins are associated with retained COMP in pseudoachondroplasia chondrocytes." Matrix Biol 20(4): 251-62.

Hecht, J. T., D. Montufar-Solis, et al. (1998). "Retention of cartilage oligomeric matrix protein (COMP) and cell death in redifferentiated pseudoachondroplasia chondrocytes." Matrix Biol 17(8-9): 625-33.

Hecht, J. T., L. D. Nelson, et al. (1995). "Mutations in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia." Nat Genet 10(3): 325-9.

Hedbom, E., P. Antonsson, et al. (1992). "Cartilage matrix proteins. An acidic oligomeric protein (COMP) detected only in cartilage." J Biol Chem 267(9): 6132-6.

Hedbom, E. and D. Heinegard (1993). "Binding of fibromodulin and decorin to separate sites on fibrillar collagens." J Biol Chem 268(36): 27307-12.

Hochstrasser, M. (1996). "Ubiquitin-dependent protein degradation." Annu Rev Genet 30: 405-39.

Hohfeld, J., D. M. Cyr, et al. (2001). "From the cradle to the grave: molecular chaperones that may choose between folding and degradation." EMBO Rep 2(10): 885-90.

Holden, P., R. S. Meadows, et al. (2001). "Cartilage oligomeric matrix protein interacts with type IX collagen, and disruptions to these interactions identify a pathogenetic mechanism in a bone dysplasia family." J Biol Chem 276(8): 6046-55.Horton, W. A. (1993). Morphology of Connective tissue:Cartilage. Connective Tissue and its Heritable Disorders, Wiley-Liss: 73-84.

Horton, W. A. (1993). Morphology of Connective tissue:Cartilage. Connective Tissue and its Heritable Disorders, Wiley-Liss: 73-84.

Horton, W. and J. Hecht, Eds. (2000). Chondrodysplasias: disorders with defects in structural components of cartilage extracellular matrix. New York, Wiley-Liss.

Horwich, A. (2002). "Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions." J Clin Invest 110(9): 1221-32.

Hosokawa, N. and K. Nagata (2000). "Procollagen binds to both prolyl 4-hydroxylase/protein disulfide isomerase and HSP47 within the endoplasmic reticulum in the absence of ascorbate." FEBS Lett 466(1): 19-25.

Kaartinen, V. and D. Warburton (2003). "Fibrillin controls TGF-beta activation." Nat Genet 33(3): 331-2.

Kaufman, R. J. (2002). "Orchestrating the unfolded protein response in health and disease." J Clin Invest 110(10): 1389-98.

Keene, D. R., G. P. Lunstrum, et al. (1991). "Two type XII-like collagens localize to the surface of banded collagen fibrils." J Cell Biol 113(4): 971-8.

Keene, D. R., J. T. Oxford, et al. (1995). "Ultrastructural localization of collagen types II, IX, and XI in the growth plate of human rib and fetal bovine epiphyseal cartilage: type XI collagen is restricted to thin fibrils." J Histochem Cytochem 43(10): 967-79.

Keene, D. R., C. C. Ridgway, et al. (1998). "Type VI microfilaments interact with a specific region of banded collagen fibrils in skin." J Histochem Cytochem 46(2): 215-20.

Khan, H., K. A. Salman, et al. (2002). "Alpha-1 antitrypsin deficiency in emphysema." J Assoc Physicians India 50: 579-82.

Kim, B. and J. W. Little (1992). "Dimerization of a specific DNA-binding protein on the DNA." Science 255(5041): 203-6.

Kim, P. S. and P. Arvan (1998). "Endocrinopathies in the family of endoplasmic reticulum (ER) storage diseases: disorders of protein trafficking and the role of ER molecular chaperones." Endocr Rev 19(2): 173-202.

Kim, P. S., O. Y. Kwon, et al. (1996). "An endoplasmic reticulum storage disease causing congenital goiter with hypothyroidism." J Cell Biol 133(3): 517-27.

Kopito, R. R. and D. Ron (2000). "Conformational disease." Nat Cell Biol 2(11): E207-9.

Kronenberg, S., C. Esser, et al. (2000). "An aryl hydrocarbon receptor conformation acts as the functional core of nuclear dioxin signaling." Nucleic Acids Res 28(12): 2286-91.

Kunne, A. G. and R. K. Allemann (1997). "Covalently linking BHLH subunits of MASH-1 increases specificity of DNA binding." Biochemistry 36(5): 1085-91.

Kuznetsov, G., L. B. Chen, et al. (1997). "Multiple molecular chaperones complex with misfolded large oligomeric glycoproteins in the endoplasmic reticulum." J Biol Chem 272(5): 3057-63.

Laskey, R. A., B. M. Honda, et al. (1978). "Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA." Nature 275(5679): 416-20.

Lawler, J., F. C. Chao, et al. (1982). "Evidence for calcium-sensitive structure in platelet thrombospondin. Isolation and partial characterization of thrombospondin in the presence of calcium." J Biol Chem 257(20): 12257-65.

Lawler, J., L. H. Derick, et al. (1985). "The structure of human platelet thrombospondin." J Biol Chem 260(6): 3762-72.

Lawler, J. and R. O. Hynes (1986). "The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins." J Cell Biol 103(5): 1635-48.

Lawler, J., K. McHenry, et al. (1995). "Characterization of human thrombospondin-4." J Biol Chem 270(6): 2809-14.

Lawler, J. and E. R. Simons (1983). "Cooperative binding of calcium to thrombospondin. The effect of calcium on the circular dichroism and limited tryptic digestion of thrombospondin." J Biol Chem 258(20): 12098-101.

Le Parc, J. M., S. Molcard, et al. (2000). "Marfan syndrome and fibrillin disorders." Joint Bone Spine 67(5): 401-7.

Liu, Y., P. Choudhury, et al. (1997). "Intracellular disposal of incompletely folded human alpha1-antitrypsin involves release from calnexin and post-translational trimming of asparagine-linked oligosaccharides." J Biol Chem 272(12): 7946-51.

London, J., C. Skrzynia, et al. (1974). "Renaturation of Escherichia coli tryptophanase after exposure to 8 M urea. Evidence for the existence of nucleation centers." Eur J Biochem 47(2): 409-15.

Loughlin, J., C. Irven, et al. (1998). "Identification of five novel mutations in cartilage oligomeric matrix protein gene in pseudoachondroplasia and multiple epiphyseal dysplasia." Hum Mutat Suppl(1): S10-7.

Lundblad, J. R., M. Laurance, et al. (1996). "Fluorescence polarization analysis of protein-DNA and protein-protein interactions." Mol Endocrinol 10(6): 607-12.

Ma, P. C., M. A. Rould, et al. (1994). "Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation." Cell 77(3): 451-9.

Macdonald, J. R. and H. P. Bachinger (2001). "HSP47 binds cooperatively to triple helical type I collagen but has little effect on the thermal stability or rate of refolding." J Biol Chem 276(27): 25399-403.

Maddox, B. K., S. Garofalo, et al. (1997). "Skeletal development in transgenic mice expressing a mutation at Gly574Ser of type II collagen." Dev Dyn 208(2): 170-7.

Maddox, B. K., D. R. Keene, et al. (1997). "The fate of cartilage oligomeric matrix protein is determined by the cell type in the case of a novel mutation in pseudoachondroplasia." J Biol Chem 272(49): 30993-7.

Maddox, B. K., A. Mokashi, et al. (2000). "A cartilage oligomeric matrix protein mutation associated with pseudoachondroplasia changes the structural and functional properties of the type 3 domain." J Biol Chem 275(15): 11412-7.

Malashkevich, V. N., R. A. Kammerer, et al. (1996). "The crystal structure of a five-stranded coiled coil in COMP: a prototype ion channel?" Science 274(5288): 761-5.

Maslen, C. L., G. M. Corson, et al. (1991). "Partial sequence of a candidate gene for the Marfan syndrome." Nature 352(6333): 334-7.

Mathieu, M. C., I. Lapierre, et al. (2001). "Aromatic hydrocarbon receptor (AhR).AhR nuclear translocator- and p53- mediated induction of the murine multidrug resistance mdr1 gene by 3- methylcholanthrene and benzo(a)pyrene in hepatoma cells." J Biol Chem 276(7): 4819-27.

Matlack, K. E., W. Mothes, et al. (1998). "Protein translocation: tunnel vision." Cell 92(3): 381-90.

Maurer, P. and E. Hohenester (1997). "Structural and functional aspects of calcium binding in extracellular matrix proteins." Matrix Biol 15(8-9): 569-80; discussion 581.

Maxwell, P. H., G. U. Dachs, et al. (1997). "Hypoxia-inducible factor-1 modulates gene expression in solid tumors and influences both angiogenesis and tumor growth." Proc Natl Acad Sci U S A 94(15): 8104-9.

Mayer, U., E. Poschl, et al. (1995). "Low nidogen affinity of laminin-5 can be attributed to two serine residues in EGF-like motif gamma 2III4." FEBS Lett 365(2-3): 129-32.

Maynard, J. A., R. R. Cooper, et al. (1972). "A unique rough surfaced endoplasmic reticulum inclusion in pseudoachondroplasia." Lab Invest 26(1): 40-4.

Michalak, M., E. F. Corbett, et al. (1999). "Calreticulin: one protein, one gene, many functions." Biochem J 344 Pt 2: 281-92.

Michel, G., E. Minet, et al. (2000). "A model for the complex between the hypoxia-inducible factor-1 (HIF-1) and its consensus DNA sequence." J Biomol Struct Dyn 18(2): 169-79.

Misenheimer, T. M., A. J. Hahr, et al. (2001). "Disulfide connectivity of recombinant C-terminal region of human thrombospondin 2." J Biol Chem 276(49): 45882-7.

Misenheimer, T. M. and D. F. Mosher (1995). "Calcium ion binding to thrombospondin 1." J Biol Chem 270(4): 1729-33.

Mogk, A., C. Schlieker, et al. (2003). "Roles of individual domains and conserved motifs of the AAA+ protein ClpB in oligomerization, ATP-hydrolysis and chaperone activity." J Biol Chem 6: 6.

Morgelin, M., J. Engel, et al. (1992). "Proteoglycans from the swarm rat chondrosarcoma. Structure of the aggregates extracted with associative and dissociative solvents as revealed by electron microscopy." J Biol Chem 267(20): 14275-84.

Morris, N. P., J. T. Oxford, et al. (2000). "Developmentally regulated alternative splicing of the alpha1(XI) collagen chain: spatial and temporal segregation of isoforms in the cartilage of fetal rat long bones." J Histochem Cytochem 48(6): 725-41.

Muhle-Goll, C., M. Nilges, et al. (1995). "The leucine zippers of the HLH-LZ proteins Max and c-Myc preferentially form heterodimers." Biochemistry 34(41): 13554-64.

Mukhopadhyay, C. K., B. Mazumder, et al. (2000). "Role of hypoxia-inducible factor-1 in transcriptional activation of ceruloplasmin by iron deficiency." J Biol Chem 275(28): 21048-54.

Muragaki, Y., E. C. Mariman, et al. (1996). "A mutation in the gene encoding the alpha 2 chain of the fibril- associated collagen IX, COL9A2, causes multiple epiphyseal dysplasia (EDM2)." Nat Genet 12(1): 103-5.

Nagata, K. (1996). "Hsp47: a collagen-specific molecular chaperone." Trends Biochem Sci 21(1): 22-6.

Newton, G., S. Weremowicz, et al. (1994). "Characterization of human and mouse cartilage oligomeric matrix protein." Genomics 24(3): 435-9.

Noiva, R. (1999). "Protein disulfide isomerase: the multifunctional redox chaperone of the endoplasmic reticulum." Semin Cell Dev Biol 10(5): 481-93.

Oldberg, A., P. Antonsson, et al. (1992). "COMP (cartilage oligomeric matrix protein) is structurally related to the thrombospondins." J Biol Chem 267(31): 22346-50.

Paassilta, P., J. Lohiniva, et al. (1999). "COL9A3: A third locus for multiple epiphyseal dysplasia." Am J Hum Genet 64(4): 1036-44.

Parodi, A. J. (2000). "Protein glucosylation and its role in protein folding." Annu Rev Biochem 69: 69-93.

Parraga, A., L. Bellsolell, et al. (1998). "Co-crystal structure of sterol regulatory element binding protein 1a at 2.3 A resolution." Structure 6(5): 661-72.

Patil, C. and P. Walter (2001). "Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals." Curr Opin Cell Biol 13(3): 349-55.

Perlmutter, D. H. (1996). "Alpha-1-antitrypsin deficiency: biochemistry and clinical manifestations." Ann Med 28(5): 385-94.

Phillips, S. E. (1994). "Built by association: structure and function of helix-loop-helix DNA- binding proteins." Structure 2(1): 1-4.

Plaas, A. H. and S. Wong-Palms (1993). "Biosynthetic mechanisms for the addition of polylactosamine to chondrocyte fibromodulin." J Biol Chem 268(35): 26634-44.

Plemper, R. K. and D. H. Wolf (1999). "Retrograde protein translocation: ERADication of secretory proteins in health and disease." Trends Biochem Sci 24(7): 266-70.

Prabakaran, D., P. S. Kim, et al. (1996). "Oligomeric assembly of thrombospondin in the endoplasmic reticulum of thyroid epithelial cells." Eur J Cell Biol 70(2): 134-41.

Qabar, A., L. Derick, et al. (1995). "Thrombospondin 3 is a pentameric molecule held together by interchain disulfide linkage involving two cysteine residues." J Biol Chem 270(21): 12725-9.

Qu, D., J. H. Teckman, et al. (1996). "Degradation of a mutant secretory protein, alpha1-antitrypsin Z, in the endoplasmic reticulum requires proteasome activity." J Biol Chem 271(37): 22791-5.

Rabouille, C., N. Hui, et al. (1995). "Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides." J Cell Sci 108(Pt 4): 1617-27.

Reinhardt, D. P., D. R. Keene, et al. (1996). "Fibrillin-1: organization in microfibrils and structural properties." J Mol Biol 258(1): 104-16.

Reinhardt, D. P., R. N. Ono, et al. (2000). "Mutations in calcium-binding epidermal growth factor modules render fibrillin-1 susceptible to proteolysis. A potential disease-causing mechanism in Marfan syndrome." J Biol Chem 275(16): 12339-45.

Reinhardt, D. P., R. N. Ono, et al. (1997). "Calcium stabilizes fibrillin-1 against proteolytic degradation." J Biol Chem 272(2): 1231-6.

Reisz-Porszasz, S., M. R. Probst, et al. (1994). "Identification of functional domains of the aryl hydrocarbon receptor nuclear translocator protein (ARNT)." Mol Cell Biol 14(9): 6075-86.

Ritter, C. and A. Helenius (2000). "Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase." Nat Struct Biol 7(4): 278-80.

Rosenberg, K., H. Olsson, et al. (1998). "Cartilage oligomeric matrix protein shows high affinity zinc-dependent interaction with triple helical collagen." J Biol Chem 273(32): 20397-403.

Ruddon, R. W. and E. Bedows (1997). "Assisted protein folding." J Biol Chem 272(6): 3125-8.

Ryan, H. E., J. Lo, et al. (1998). "HIF-1 alpha is required for solid tumor formation and embryonic vascularization." Embo J 17(11): 3005-15.

Saito, Y., Y. Ihara, et al. (1999). "Calreticulin functions in vitro as a molecular chaperone for both glycosylated and non-glycosylated proteins." Embo J **18**(23): 6718-29.

Sakai, L. Y. and D. R. Keene (1994). "Fibrillin: monomers and microfibrils." Methods Enzymol 245: 29-52.

Salomon-Nguyen, F., V. Della-Valle, et al. (2000). "The t(1;12)(q21;p13) translocation of human acute myeloblastic leukemia results in a TEL-ARNT fusion." Proc Natl Acad Sci U S A 97(12): 6757-62.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY., Cold Spring Harbor Laboratory.

Saxne, T. and D. Heinegard (1992). "Cartilage oligomeric matrix protein: a novel marker of cartilage turnover detectable in synovial fluid and blood." Br J Rheumatol 31(9): 583-91.

Schecter, A., L. C. Dai, et al. (1995). "Agent Orange and the Vietnamese: the persistence of elevated dioxin levels in human tissues." Am J Public Health 85(4): 516-22.

Schurr, J. M. (1977). "Dynamic light scattering of biopolymers and biocolloids." CRC Crit Rev Biochem 4(4): 371-431.

Scopes, R. K. (1993). Protein Purification: Principles and Practice. Berlin, Springer-Verlag.

Shen, Z., D. Heinegard, et al. (1995). "Distribution and expression of cartilage oligomeric matrix protein and bone sialoprotein show marked changes during rat femoral head development." Matrix Biol 14(9): 773-81.

Shimizu, T., A. Toumoto, et al. (1997). "Crystal structure of PHO4 bHLH domain-DNA complex: flanking base recognition." Embo J 16(15): 4689-97.

Sidrauski, C., R. Chapman, et al. (1998). "The unfolded protein response: an intracellular signalling pathway with many surprising features." Trends Cell Biol 8(6): 245-9.

Smith, R. K., L. Zunino, et al. (1997). "The distribution of cartilage oligomeric matrix protein (COMP) in tendon and its variation with tendon site, age and load." Matrix Biol 16(5): 255-71.

Sogawa, K., R. Nakano, et al. (1995). "Possible function of Ah receptor nuclear translocator (Arnt) homodimer in transcriptional regulation." Proc Natl Acad Sci U S A 92(6): 1936-40.

Sommer, T. and D. H. Wolf (1997). "Endoplasmic reticulum degradation: reverse protein flow of no return." Faseb J 11(14): 1227-33.

Stanescu, R., V. Stanescu, et al. (1993). "Multiple epiphyseal dysplasia, Fairbank type: morphologic and biochemical study of cartilage." Am J Med Genet 45(4): 501-7.

Stanescu, V., R. Stanescu, et al. (1982). "Pathogenesis of pseudoachondroplasia and diastrophic dysplasia." Prog Clin Biol Res 104: 385-94.

Stanescu, V., R. Stanescu, et al. (1984). "Pathogenic mechanisms in osteochondrodysplasias." J Bone Joint Surg Am 66(6): 817-36.

Stryer, L. (1995). Biochemistry. New York, W. H. Freeman and Company, New York..

Svensson, L., A. Aszodi, et al. (2002). "Cartilage oligomeric matrix protein-deficient mice have normal skeletal development." Mol Cell Biol 22(12): 4366-71.

Swanson, H. I. and J. H. Yang (1998). "The aryl hydrocarbon receptor interacts with transcription factor IIB." Mol Pharmacol 54(4): 671-7.

Swanson, H. I. and J. H. Yang (1999). "Specificity of DNA binding of the c-Myc/Max and ARNT/ARNT dimers at the CACGTG recognition site." Nucleic Acids Res 27(15): 3205-12.

Takechi, H., K. Hirayoshi, et al. (1992). "Molecular cloning of a mouse 47-kDa heat-shock protein (HSP47), a collagen-binding stress protein, and its expression during the differentiation of F9 teratocarcinoma cells." Eur J Biochem 206(2): 323-9.

Tasab, M., M. R. Batten, et al. (2000). "Hsp47: a molecular chaperone that interacts with and stabilizes correctly-folded procollagen." Embo J 19(10): 2204-11.

Tatu, U., C. Hammond, et al. (1995). "Folding and oligomerization of influenza hemagglutinin in the ER and the intermediate compartment." Embo J 14(7): 1340-8.

Tatu, U. and A. Helenius (1997). "Interactions between newly synthesized glycoproteins, calnexin and a network of resident chaperones in the endoplasmic reticulum." J Cell Biol 136(3): 555-65.

Thomson, C. A. and V. S. Ananthanarayanan (2000). "Structure-function studies on hsp47: pH-dependent inhibition of collagen fibril formation in vitro." Biochem J 349 Pt 3: 877-83.

Thur, J., K. Rosenberg, et al. (2001). "Mutations in cartilage oligomeric matrix protein causing pseudoachondroplasia and multiple epiphyseal dysplasia affect binding of calcium and collagen I, II, and IX." J Biol Chem 276(9): 6083-92.

Travers, K. J., C. K. Patil, et al. (2000). "Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation." Cell 101(3): 249-58.

Tucker, R. P., J. C. Adams, et al. (1995). "Thrombospondin-4 is expressed by early osteogenic tissues in the chick embryo." Dev Dyn 203(4): 477-90.

Urade, R. and M. Kito (1992). "Inhibition by acidic phospholipids of protein degradation by ER-60 protease, a novel cysteine protease, of endoplasmic reticulum." FEBS Lett 312(1): 83-6.

van Holde, K. E. (1985). Physical Biochemistry. Upper Saddle River, NJ, Prentice Hall.

van Laar, T., A. J. van der Eb, et al. (2001). "Mif1: a missing link between the unfolded protein response pathway and ER-associated protein degradation?" Curr Protein Pept Sci 2(2): 169-90.

Vashist, S., W. Kim, et al. (2001). "Distinct retrieval and retention mechanisms are required for the quality control of endoplasmic reticulum protein folding." J Cell Biol 155(3): 355-68.

Velloso, L. M., K. Svensson, et al. (2002). "Crystal structure of the carbohydrate recognition domain of p58/ERGIC- 53, a protein involved in glycoprotein export from the endoplasmic reticulum." J Biol Chem 277(18): 15979-84.

Vertel, B. M., A. Velasco, et al. (1989). "Precursors of chondroitin sulfate proteoglycan are segregated within a subcompartment of the chondrocyte endoplasmic reticulum." J Cell Biol 109(4 Pt 1): 1827-36.

Vranka, J., A. Mokashi, et al. (2001). "Selective intracellular retention of extracellular matrix proteins and chaperones associated with pseudoachondroplasia." Matrix Biol 20(7): 439-50.

Wang, C. C. (1998). "Protein disulfide isomerase assists protein folding as both an isomerase and a chaperone." Ann N Y Acad Sci 864: 9-13.

Wendt, H., R. M. Thomas, et al. (1998). "DNA-mediated folding and assembly of MyoD-E47 heterodimers." J Biol Chem 273(10): 5735-43.

Wickner, S., M. R. Maurizi, et al. (1999). "Posttranslational quality control: folding, refolding, and degrading proteins." Science 286(5446): 1888-93.

Wilkinson, K. D. (1999). "Ubiquitin-dependent signaling: the role of ubiquitination in the response of cells to their environment." J Nutr 129(11): 1933-6.

Winston, R. L., D. P. Millar, et al. (1999). "Characterization of the DNA binding properties of the bHLH domain of Deadpan to single and tandem sites." Biochemistry 38(16): 5138-46.

Zaia, J., R. E. Boynton, et al. (1997). "Post-translational modifications in cartilage oligomeric matrix protein. Characterization of the N-linked oligosaccharides by matrix-assisted laser desorption ionization time-of-flight mass spectrometry." J Biol Chem 272(22): 14120-6.

Zaucke, F., R. Dinser, et al. (2001). "Cartilage oligomeric matrix protein (COMP) and collagen IX are sensitive markers for the differentiation state of articular primary chondrocytes." Biochem J 358(Pt 1): 17-24.

Zeng, B., J. R. MacDonald, et al. (1998). "Chicken FK506-binding protein, FKBP65, a member of the FKBP family of peptidylprolyl cis-trans isomerases, is only partially inhibited by FK506." Biochem J 330(Pt 1): 109-14.

Zhang, Y., G. Nijbroek, et al. (2001). "Hsp70 molecular chaperone facilitates endoplasmic reticulum-associated protein degradation of cystic fibrosis transmembrane conductance regulator in yeast." Mol Biol Cell 12(5): 1303-14.

Zhulin, I. B., B. L. Taylor, et al. (1997). "PAS domain S-boxes in Archaea, Bacteria and sensors for oxygen and redox." Trends Biochem Sci 22(9): 331-3.