STRUCTURE AND FUNCTION STUDIES OF HSP47: A COLLAGEN-SPECIFIC MOLECULAR CHAPERONE

By

CHRISTY A. THOMSON, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

STRUCTURE AND FUNCTION STUDIES OF HSP47

DOCTOR OF PHILOSOPHY (2002) McMaster University

(Biochemistry) Hamilton, Ontario

TITLE:	Structure and Function Studies of Hsp47: A Collagen-Specific Molecular Chaperone
AUTHOR:	Christy A. Thomson, B.Sc. (Acadia University)
SUPERVISOR:	Professor V.S. Ananthanarayanan
NUMBER OF PAGES:	xiii, 142

ABSTRACT

Heat shock protein 47 (Hsp47) is an ER resident protein and is considered to be a molecular chaperone specific for collagen. Although Hsp47 has been shown to be necessary for collagen biosynthesis, its exact role(s) as a chaperone for collagen is/are unknown. Available evidence on type I collagen shows that Hsp47 binds to the nascent proα1 and proα2 chains and remains bound during triple-helix formation and subsequent export from the ER. Upon entering the *cis*-Golgi network, Hsp47 is thought to release procollagen and is subsequently returned to the ER.

With a view to understanding the exact role of Hsp47 in collagen biosynthesis, we undertook biophysical and biochemical characterization of Hsp47. This involved purification of recombinant Hsp47 using the IMPACT^{TM} T7 system. Purified Hsp47 existed as both a monomer and a trimer, and underwent pH-induced conformational changes that correlated to its collagen-binding activity. Hsp47 was shown to effectively inhibit type I collagen fibrillogenesis *in vitro*. This inhibition of fibre formation provided a novel assay for Hsp47 activity, as well as support for an *in vivo* role for Hsp47 in preventing collagen aggregation in the ER. Finally, in an attempt to map the binding site(s) on the collagen molecule, assays were designed to analyze the interaction of radio-labelled Hsp47 with CNBr fragments derived from type I and type II collagen. The results showed that Hsp47 preferentially binds to peptides from the N-terminal region of collagen in their triple-helical conformation.

ACKNOWLEDGMENTS

I would like to take this opportunity to thank those individuals who were instrumental in the completion of this thesis. First and foremost, I thank my family for their unconditional support. Mom and dad, your sacrifice will not be forgotten. To Scott Covey, we persevered together and your unwavering dedication to your own studies provided me with needed incentive.

I would like to thank my friends, members of Dr. Ananthanaryanan's laboratory, and the kitties from the SPCA. Our time spent together proved as a reminder that there is life beyond science. In particular I would like to thank Aaron Kerman who served as the Yin to my Yang and thus helped to maintain a working balance in the laboratory.

To Dr. Rosenthal, Dr. Wright and Dr. Yang, as members of my supervisory committee you not only guided me during my studies but also played a key role in motivating my learning. I am grateful for not only this but also for the open door policies and the generous sharing of reagents and equipment.

To my supervisor, Dr. Ananthanarayanan, I would like to thank you for your guidance throughout my studies. You provided an active learning environment that promoted independent thinking. I am grateful to you for allowing me to pursue some of my own interests and for your confidence in my abilities, especially in the early years.

Finally, I would like to thank McMaster University and the Heart and Stroke Foundation of Canada for their financial support. It was muchly needed and greatly appreciated.

PREFACE

Chapters 2 and 3 of this thesis contain material that has been previously published. Chapter 4 contains material that has been submitted for publication prior to completion of this thesis. The purpose of this preface is to draw attention to the previously published material in lieu of externally referencing it in the thesis. All experimental work illustrated in each of these chapters was performed by myself.

Chapter 2 was published as: Christy A. Thomson and Vettai S. Ananthanarayanan (2000). Structure-function studies on Hsp47: pH-dependent inhibition of collagen fibril formation in vitro. Biochem J *349*, 877-83.

Chapter 3 was published as: Christy A. Thomson and Vettai S. Ananthanarayanan (2001). A method for expression and purification of soluble, active Hsp47, a collagen-specific molecular chaperone. Protein Expr Purif *23*, 8-13.

Chapter 4 was submitted to Protein Science (Oct. 2002) as: Christy A. Thomson, Ruggero Tenni and Vettai S. Ananthanarayanan. Mapping Hsp47-binding Site(s) Using CNBr Peptides Derived from Type I and Type II Collagen

The above articles have been reproduced here with the permission of co-authors. Irrevocable, nonexlusive license has been granted to McMaster University and the National Library of Canada for Chapters 2 and 3 by the appropriate publishers. Copies of the granted licences have been given to the School of Graduate Studies at McMaster University.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 4

LIST OF TABLES

ABBREVIATIONS

ANS	-1-anilino-8-naphthalene sulfonate
ATP	-adenosine triphosphate
BiP	-immunoglobulin heavy chain binding protein
BSA	-bovine serum albumin
°C	-degrees, Celsius
CD	-circular dichroism
CNBr	-cyanogen bromide
COP	-coat protein
CsA	-cyclosporin A
dpc	-days postcoitus
DTT	-dithiothreitol
ECM	-extracellular matrix
EDTA	-*N,N,N'*,N'-ethylenediaminetetraacetic acid
ER	-endoplasmic reticulum
FACIT	-fibril-associated collagen with interrupted triple-helical regions
Grp	-glucose-related protein
Hsp47	-47 kDa heat shock protein
Hyp or O	-4-hydroxyproline
IMPACT	-intein mediated purification with an affinity chitin-binding tag
IPTG	-isopropyl-α-D-thiogalactophyranoside

K_d	-dissociation constant
LB	-Luria-Bertani medium
LDM	-low density microsomes
PBS	-phosphate-buffered saline
PCR	-polymerase chain reaction
PDI	-protein disulfide isomerase
PMSF	-phenylmethylsulfonyl fluoride
RSV	-Rous sarcoma virus
serpin	-serine proteinase inhibitor
SV40	-simian virus 40
T_m	-temperature midpoint
TBS	-tris-buffered saline
Tris	-tris-(hydroxymethyl)aminomethane
TTBS	-tris-buffered saline with Tween-20
SDS-PAGE	-sodium dodecyl sulfate polyacrylamide gel electrophoresis

CHAPTER 1: INTRODUCTION

1.1 COLLAGEN

Collagens are the major protein component of the extracellular matrix (ECM) and are composed of three polypeptide chains interwoven together to form a right-handed superhelix. To date, twenty-six types of collagen have been identified based on variability in their α chains and their homotrimeric and heterotrimeric assemblies (Sato et al., 2002). For a given type of collagen, the individual α chains are numbered by Arabic numerals with the collagen type indicated in parentheses by Roman numerals. In tissues, collagens assemble into different polymeric structures and are grouped into classes based on the structures they form. Examples of collagen classes include collagens that form fibrils; fibril-associated collagens with interrupted triple-helical regions (FACIT); and collagens found in basement membranes. Myllyharju and Kivirikko (2001) provide a recent review on the collagen types and the polymeric assemblies they form. Since this thesis focuses predominately on Hsp47's interaction with type I collagen, a fiber forming collagen, the following sections will describe, in detail, its biosynthesis, assembly and structure.

1.1.1 Collagen Biosynthesis

Type I collagen is synthesized as a precursor molecule, termed procollagen, which undergoes extensive post-translational processing before the formation of mature collagen. Figure 1.1 illustrates the structural features of procollagen whose biosynthesis originates with translocation of two pro-α1(I) chains and one pro-α2(I) chain into the rough endoplasmic reticulum (ER) (Beck et al., 1996). Each procollagen chain is

approximately 1450 amino acids in length and consists of a long triple-helical domain flanked by globular N- and C-terminal propeptide domains. Between the propeptide ends and the central triple-helical domain lie short non-helical regions called the N- and C-telopeptides.

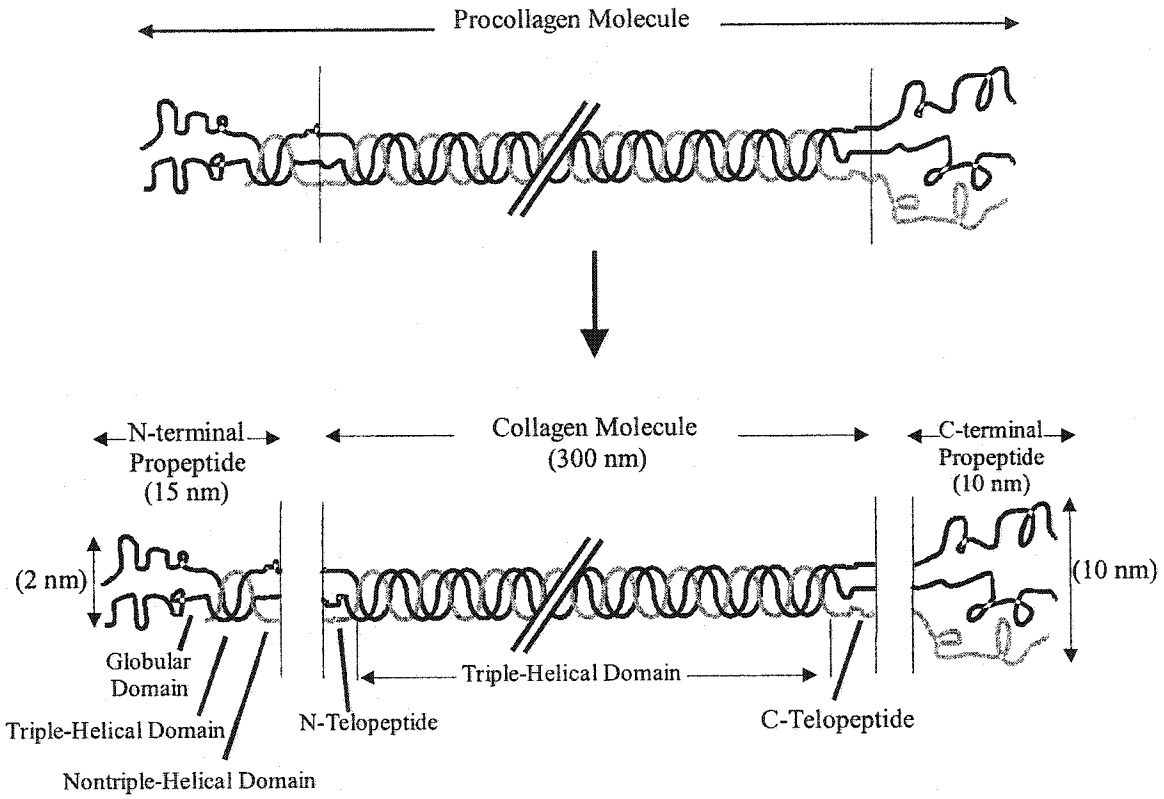


Figure 1.1 A schematic representing the structure of the procollagen molecule.

(Adapted from Prockop et al., 1979)

Before folding into a triple helix, newly translated procollagen chains interact with a variety of molecular chaperones and modifying enzymes (Lamande and Bateman, 1999). It is the interaction with these proteins that allows the polypeptide chains of procollagen to assemble into a triple helix and to be secreted into the extracellular space for incorporation into stable, cross-linked fibrils. The nascent proα chains of procollagen are first bound by the chaperones Grp78, Grp94 and Hsp47 (Ferreira et al., 1994). Grp78 and Grp94 are chaperones of the ER which recognize and bind multiple unfolded proteins, whereas Hsp47 is considered to be a collagen-specific molecular chaperone (Nagata, 1996). Although Hsp47's exact role in collagen biosynthesis is unclear, it is known that Hsp47 interacts with the nascent procollagen chains and remains associated with them after triple helix formation and during export to the cis-Golgi apparatus (Sauk et al., 1994; Smith et al., 1995). A detailed description of Hsp47 and its interactions with collagen will be forthcoming in subsequent sections of the introduction.

Collagen is a unique protein for a variety of reasons. Notable among them is the fact that many of its Pro and Lys residues are post-translationally modified to hyroxyproline (Hyp) and hydroxylysine respectively. The enzymes that catalyze these reactions are prolyl 4-hydroxylase and lysyl hydroxylase (As in Prockop et al., 1979). Hydroxylation of the prolyl and lysyl residues is specific to their location within the collagen polypeptide. For instance, prolyl and lysyl hydroxylase only hydroxylate proline and lysine residues, respectively, in the Y position of the Gly-X-Y sequence (Hutton et al., 1967; Kivirikko et al., 1972). Additionally, hydroxylation of these residues requires that the individual procollagen chains are in a random conformation,

with trimerization of the chains preventing further hydroxylation (Murphy and Rosenbloom, 1973). Once hydroxylated, the lysyl residues are further modified via O-linked glycosylation (Aguilar et al., 1973). These modifications are required for collagen's interactions in the ECM leading to fiber formation.

Hydroxylation of proline residues by prolyl 4-hydroxylase is known to be essential for collagen stability (Berg and Prockop, 1973). The prolyl 4-hydroxylase enzyme is a tetramer composed of two α subunits and two β subunits. The α subunits are the catalytic portion of the protein whereas the β subunits are identical to protein disulfide isomerase (PDI) (Vuori et al., 1992). PDI itself is a multifunctional protein of the ER. As part of prolyl 4-hydroxylase, its role is to prevent the aggregation of the α subunit and to retain the α subunit in the ER (John et al., 1993). More recently it has been proposed that PDI also plays a role as a molecular chaperone which interacts with monomeric procollagen chains to prevent their premature assembly (Wilson et al., 1998).

In order for collagen triple helix formation to take place, it is necessary for its many Pro and Hyp peptide bonds to be in the *trans*-configuration. Because, in nascent procollagen, many of these bonds are in the *cis*-configuration, they must undergo *cis/trans*-isomerization before folding can occur (Engel and Prockop, 1991). The ability of cyclosporin A (CsA) to slow triple helix formation of collagen *in situ*, has implicated members of the cyclophilin family of proteins as being responsible for catalyzing the *cis/trans*-isomerization of collagen peptide bonds (Steinmann et al., 1991). Cyclophilins are peptidyl-prolyl-*cis-trans*-isomerases whose activity is inhibited by CsA, and cyclophilin B has been shown to bind procollagen in the ER (Smith et al., 1995).

Propagation of the collagen triple helix occurs in a 'zipper-like' fashion from the C-terminus towards the N-terminus, with the *cis/trans*-isomerization of its peptide bonds being the rate-limiting step (Bachinger et al., 1980). Inter-chain disulfide bonds between the C-propeptides are required for proper chain alignment, which then serves as a nucleus for triple-helix formation (Lees et al., 1997). Indeed, if mature fully processed type I collagen minus the propeptides is thermally denatured, only 15-20 % of the chains are able to properly align and form full triple helices (Bruckner and Prockop, 1981). Similarly, mutations of collagen that are towards the C-terminus, generally result in more severe phenotypes than those towards the N-terminus (Bachinger et al., 1993).

Once folded into a triple helix in the ER, procollagen exits the cell via the secretory pathway. It could be postulated that procollagen passes through the Golgi network as monomers in transport vesicles and that upon being deposited into the ECM it would be enzymatically processed and self-assembled into native fibrils. However, studies of the transport of procollagen through the Golgi revealed that the procollagen is bundled into condensed aggregates within the Golgi before being discharged into the ECM (Bruns et al., 1979; Trelstad and Hayashi, 1979). As well, it was observed that the Golgi vacuoles were too small in size (~ 70 nm diameter) to accommodate the large procollagen molecule (300 nm) (Trelstad and Hayashi, 1979). Recent studies on the mechanism of procollagen export from the ER through the Golgi network has shed further light on the secretion process. Briefly, it is now thought that procollagen exits the ER in a transport complex distinct from representative cargo proteins and that both COPI and COPII coat proteins are required for export (Stephens and Pepperkok, 2002).

Procollagen is then thought to remain in the lumen of Golgi cisternae and transverse the Golgi via cisternal progression (Bonfanti et al., 1998). Although further investigation is required, the above results point to an exciting new field of collagen research and also to the benefit of exploiting the complex collagen biosynthesis pathway to examine biological processes such as protein secretion.

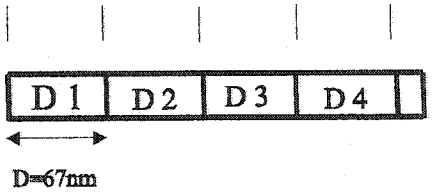
1.1.2 Collagen Assembly

Triple helical procollagen passes through the Golgi network and is secreted onto the cell surface where it undergoes fibrillogenesis (Leblond, 1989). Enzymatic reactions required for the formation of mature fibers, include removal of the C- and N-terminal propeptides by procollagen C-proteinase and procollagen N-proteinase, and the oxidative deamination of lysines and hydroxylysines by lysyl oxidase (Prockop et al., 1998; Siegel, 1976). The removal of the propeptides is thought to dictate the timing of fiber formation as well as the fibril shape and diameter (Hulmes, 2002), whereas the aldehyde groups generated by deamination are required for covalent cross-linking between adjacent collagen molecules (Siegel, 1976).

The formation of collagen fibers is a multistep process which is not completely understood at the molecular level and which remains controversial. *In vitro*, the self-assembly of collagen fibers can be monitored by turbidity measurements recorded as optical density at 313 nm (Gelman et al., 1979). The turbidity curves generated are sigmoidal and can be divided into a lag phase in which nucleation of the collagen fibers is thought to occur, and then a propagation phase of fiber growth which subsequently

reaches a plateau. From EM studies it is known that in the formed fibrils, the individual collagen molecules are staggered by a discrete fibril periodicity of D, where D is composed of 234 residues at a length of 67 nm (As in Kadler et al., 1996). This D period represents a repeating unit in the collagen triple helix based on the distribution of hydrophobic and hydrophilic amino acids, of which there are 4.4 D periods in the mature collagen molecule (Hulmes et al., 1973). As illustrated in Figure 1.2, the individual collagen chains are staggered relative to each other by 1 D, 2 D, 3 D, or 4 D which results in a 0.4 D overlap region and a 0.6 D gap region (Hulmes et al., 1973). Also important is the directionality of the fibers. Fibers can be either N-N bipolar with N-termini at each end of the growing fiber, or unipolar with both C- and N-termini (Graham et al., 2000; Kadler et al., 1996).

A.



B.

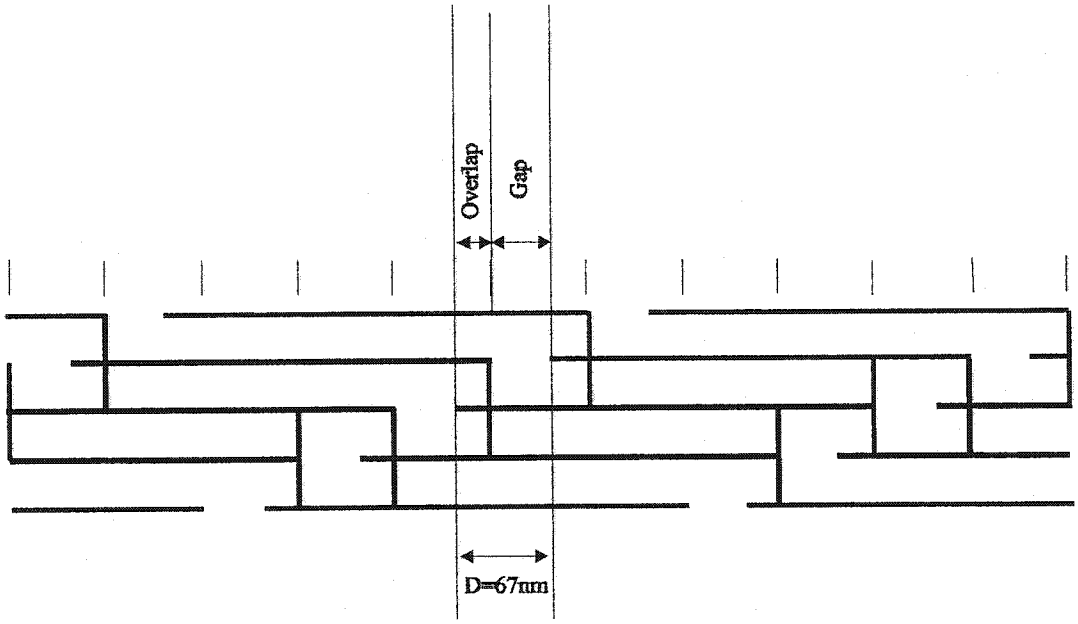


Figure 1.2 D-periodic collagen fibrils showing axial packing arrangement. A) Schematic of a type I collagen molecule divided into D periods. B) Schematic showing the axial packing arrangement of type I collagen molecules in a fibril. (Adapted from Kadler et al., 1996).

In vivo, the ability of collagen to form fibers with D periodicity and with discrete diameter and shape is based on a variety of factors. Included in them is the role of the telopeptides in nucleation of collagen assembly. Telopeptides are short non-helical domains at both the C- and N-terminus of the collagen triple helix that remain after processing of the collagen molecule by the C- and N-proteinases (Figure 1.1). The telopeptides have been shown to be important for catalyzing fiber assembly yet their exact role in molecular recognition remains in debate (Kuznetsova and Leikin, 1999; Prockop and Fertala, 1998). As well, collagen fibers are known to be heterogeneous, with the heterogeneity itself thought to dictate fiber assembly (Burgeson and Nimni, 1992). For instance, the presence of type V collagen has been shown to regulate the diameter of type I collagen fibers (Chanut-Delalande et al., 2001). Other factors that have been proposed to modulate fiber assembly include differences in lysine glycosylation, and the interaction of collagen with small leucine-rich proteoglycans (Graham et al., 2000; Notbohm et al., 1999). Needless to say, with the abundance of collagen types and collagen interacting proteins dictating fiber assembly *in vivo*, fiber assembly is a complex process regulated at multiple levels.

1.1.3 Collagen Structure

Collagen is a rod-like protein that provides tensile strength to tissues and that interacts with a wide variety of molecules. The structure of the collagen trimer has been extensively studied in order to understand at a molecular level its biochemical properties. (See Brodsky and Ramshaw, 1997, for a recent review.) The collagen triple helix is

characterized by a coiled coil structure resulting from three left-handed polyproline II helical polypeptide chains interwoven together to form a right-handed superhelix . The primary structure of collagen can be represented by $(Gly\text{-}X\text{-}Y)_n$ where X and Y can be any amino acid (Ramachandran, 1976). However, examination of the collagen sequence shows that approximately 30 % of the X and Y positions are occupied by imino acids, with Pro being preferentially found in the X position and hydroxyproline (Hyp) being preferentially found in the Y position (van der Rest and Garrone, 1991). The inability of the many Pro residues to freely rotate around the $N\text{-}C^{\alpha}$ bond provides the collagen structure with its rigidity (Ramachandran, 1976). The presence of hydroxyproline in the Y position increases the thermal stability of the collagen triple helix (Engel et al., 1977) via the formation of intra- and inter-triple-helical hydrogen bonds (Brodsky and Ramshaw, 1997).

Collagen model peptides such as $(Pro\text{-}Pro\text{-}Gly)_n$ and $(Pro\text{-}Hyp\text{-}Gly)_n$ have been extensively used as mimics of the collagen structure. Early X-ray studies showed that they could adopt a triple helical structure and exhibit triple helix-to-coil transitions characteristic of the collagen molecule (Engel et al., 1977; Yonath and Traub, 1969). Recently, collagen model peptides have been used in X-ray crystallographic studies to further define the molecular interactions important in maintaining the collagen structure (Brodsky and Ramshaw, 1997). For instance, substitution of an Ala residue in place of the central Gly caused a slight alteration in peptide conformation which could be related to the effects of glycine mutations in collagen diseases such as osteogenesis imperfecta (Bella et al., 1994).

Because Gly is the only amino acid small enough to reside along the center axis of the triple helix, all amino acids in the X and Y positions of collagen are exposed on the surface of the molecule. As a consequence, amino acids that are normally buried in a globular protein are exposed to the solvent in the collagen molecule and can therefore play a role in both the lateral interactions of collagen triple-helices and in interactions with other molecules. Arg is an interesting amino acid in terms of the collagen structure as its nitrogen atoms have been shown to make direct contacts with the peptide backbone (Kramer et al., 2001). This imparts on it both hydrophobic and hydrophilic characteristics and allows Arg to interact with neighboring hydrophobic amino acids (Kramer et al., 2001). As a result, Arg can offer stability to the collagen triple helix that is comparable to that offered by Hyp (Yang et al., 1997). Similarly, studies using the host-quest peptide system showed that the positioning of amino acids in either the X or Y site along the triple helix could influence their potential for inter- and intramolecular interactions with Arg being preferentially found in the Y position (Chan et al., 1997).

As a folded triple helix, the collagen molecule exhibits a circular dichroism (CD) spectrum containing a positive peak at 221 nm. Using CD to monitor collagen conformation *in vitro,* a temperature midpoint (T_m) of 42°C was initially observed for its helix to coil conversion (Hayashi et al., 1979). Although a value of 45°C was found for collagen triple helices *in cellulo* (Bruckner and Eikenberry, 1984), it was considered odd that a protein involved in holding our tissues together could be thermally denatured at a temperature so close to that of the human body. Further investigations into collagen stability using proteases to probe conformation, showed that collagen undergoes

temporary and local unfolding that correlates with temperature (Kadler et al., 1988; Ryhanen et al., 1983). Local relaxation of collagen conformation is described as being "non-denaturational" and it is proposed that these regions of motility or microunfolding are required for collagen assembly (Privalov, 1982). Leikina et al. (2002) recently revisited the equilibrium T_m of monomeric collagen and by taking measurements at slower heating and cooling rates, were able to show that the T_m of human type I collagen is actually less than 36°C. This finding provides new insight into collagen folding and assembly while at the same time raising the question as to how collagen folds *in vivo* if it is unstable at body temperature? Perhaps a collagen-specific molecular chaperone(s) is required to prevent its unfolding in the biosynthetic pathway.

1.2 HSP47

In 1984 Kurkinen et al. reported that a basic glycoprotein with N-linked oligosaccharide side chains and a molecular weight of 47 kDa was found in various cell lines which produce type IV procollagen. They called this protein 'colligin' because it bound native type IV collagen and gelatin. Several other proteins were subsequently found to have similar collagen-binding properties, including rat skeletal myoblast gp46 (Clarke et al., 1991), J6 from mouse teratocarcinoma cells (Wang and Gudas, 1990) and Hsp47 from chick embryo fibroblasts (Hirayoshi et al., 1991). Comparison of the cDNA sequences of these gelatin-binding proteins to human gp46 (cbp1) found a high degree of sequence homology, and they were found to represent the same protein with differences in sequence resulting from species variation (Clarke and Sanwal, 1992). Further analysis

of colligin sequences showed that colligin is a member of the serpin supergene family and that in humans there is both a pseudogene (cbp1) and a functional gene (cbp2) (Clarke et al., 1991; Nagai et al., 1999). To avoid confusion when referring to this protein, the name Hsp47, as found in the recent literature, will be used throughout this thesis (Nagata, 1998).

1.2.1 Hsp47 Expression

Hsp47's synthesis is known to parallel that of collagen. Tissue localization in neonatal rats showed that Hsp47 was found in tissues with a high tensile strength such as muscle, skin, heart, kidney and lungs, and that it is absent in soft tissues such as brain and liver (Miyaishi et al., 1992; Nandan et al., 1990). Hsp47 is produced in collagen-producing cell lines such as fibroblast and myoblast cell lines and is absent in other cells (Nandan et al., 1990; Pak et al., 1996). The expression of Hsp47 is modulated at several cellular events including differentiation (Kurkinen et al., 1984), heat shock (Nagata et al., 1986), transformation (Nagata and Yamada, 1986), and in pathological conditions such as fibrosis (Masuda et al., 1994) and myocardial infarction (Takeda et al., 2000).

Regulation of Hsp47 expression occurs at both the translational and transcriptional levels. The mouse hsp47 gene was found to contain six exons and five introns with heat shock inducing alternative 5′ splicing (Hosokawa et al., 1993; Takechi et al., 1994). The splice variant formed at 42°C was shown to have increased translatability that correlated to the increased expression of Hsp47 observed under heat shock (42°C) conditions (Nagata et al., 1986; Takechi et al., 1994). The mouse hsp47

promoter contains a complete heat-shock element (HSE) with the HSE consensus sequence of 3 repeats of NGAAN and NTTCN (Amin et al., 1988; Hosokawa et al., 1993). The HSE is responsible for Hsp47's transcriptional induction as a result of heat shock whereas the parallel expression of Hsp47 with collagen under normal growth conditions is under the control of two separate elements (Hirata et al., 1999). One is within the 280-bp promoter region that is responsible for the basal expression of Hsp47. The other is in a downstream intron that is responsible for cell-specific expression.

1.2.2 Biochemical Characterization of Hsp47

Hsp47 is a basic glycoprotein (pI = 9) phosphorylated on Ser and Thr residues (Nagata and Yamada, 1986) and containing N-linked oligosaccharide chains of the high mannose type (Hughes et al., 1987). However, binding studies of recombinant Hsp47 demonstrated that neither glycosylation nor phosphorylation is required for procollagen binding (Jain et al., 1994). Hsp47 is an ER-resident protein containing an N-terminal signal sequence for translocation into the ER and a C-terminal RDEL sequence for ER retention (Hirayoshi et al., 1991; Satoh et al., 1996). Confocal microscopy and immunofluorescence experiments revealed that Hsp47 co-localizes with procollagen in the ER and remains associated with it through its transfer to the Golgi apparatus (Satoh et al., 1996; Smith et al., 1995). Hsp47 dissociates from procollagen between the post-ER and cis-Golgi compartments and is recycled back to the ER via its RDEL sequence (Satoh et al., 1996). Figure 1.3 depicts the association of Hsp47 with procollagen along the secretory pathway (Nagata, 1996).

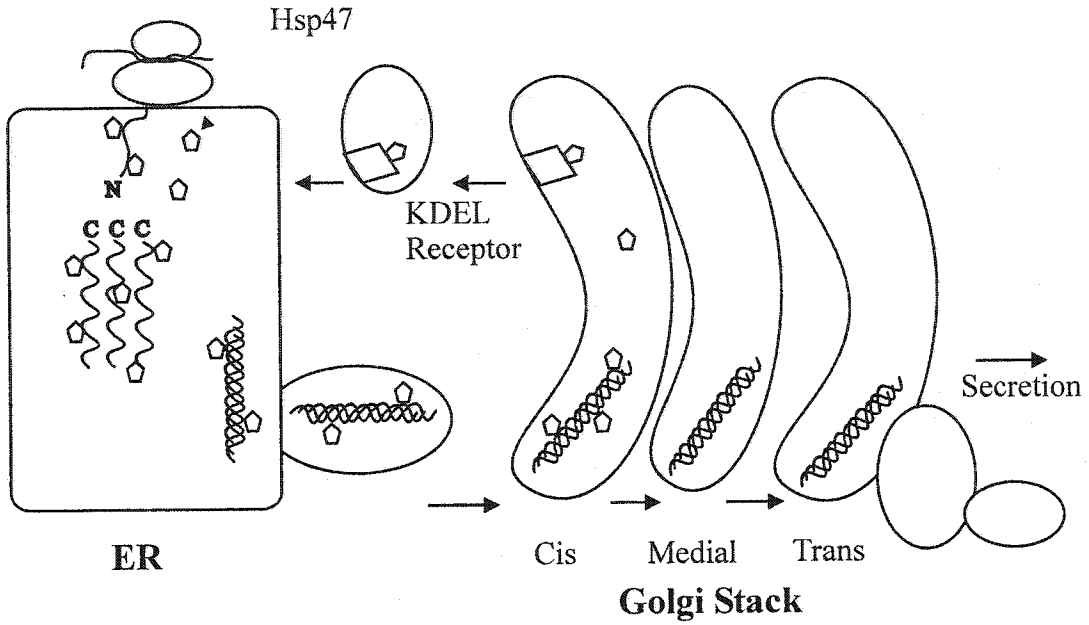


Figure 1.3 Schematic representation of the association of Hsp47 with procollagen in the secretory pathway. Hsp47 binds to nascent procollagen chains and remains associated with it during triple helix formation and export from the ER. Hsp47 dissociates from procollagen upon entering the cis-Golgi where it is bound by the KDEL receptor for subsequent recycling back to the ER. (Adapted from Nagata, 1996).

1.2.3 Hsp47 as a Serpin

Sequence analysis of Hsp47 found no significant sequence homology to other ER proteins (Clarke et al., 1991) or to other collagen-binding proteins (Hirayoshi et al., 1991). Surprisingly, Hsp47 was found to be a member of the serpin supergene family

with the highest similarity to human α_1-antitrypsin and human protein C inhibitor (Clarke et al., 1991; Hirayoshi et al., 1991). Alignment of Hsp47 to the reactive site region of serpins implicated Arg^{365} and Ser^{366} as the P_1 and P'_1 residues respectively, with considerable sequence similarity being found C-terminal of the reactive site between residues P_1-P'_{15} (Clarke et al., 1991). N-terminal of the reactive site, Hsp47 was found to have very little homology to inhibitory serpins, with amino acids containing large side chains (Phe-Asp-Glu) in the P_{12}-P_{10} sites (Clarke et al., 1991). This implicates Hsp47 as having no inhibitory activity because small amino acids such as Ala are required in these positions to allow rapid insertion of the reactive center loop into β-sheet A (Hopkins et al., 1993; Irving et al., 2000).

Homology modeling of Hsp47 against the known structure of human protein C inhibitor revealed a long, deep cleft with both hydrophilic and hydrophobic amino acids being projected inward (Davids et al., 1995). This cleft is a likely candidate for collagen binding due to both its size and the fact that both hydrophilic and hydrophobic interactions have been implicated in Hsp47 binding to collagen (Sauk et al., 2000; Thomson and Ananthanarayanan, 2000). Analysis of a similar Hsp47 model revealed three clusters of three His in regions important to serpin conformational change (Dafforn et al., 2001). His residues have long been implicated as having a role in Hsp47 binding to collagen, as binding is abolished below pH 6.3 (Saga et al., 1987; Thomson and Ananthanarayanan, 2000). In the absence of a crystal structure, a three-dimensional model of Hsp47 may prove useful in future studies of ligand binding and in attempts to derive inhibitors of Hsp47 activity.

1.2.4 Hsp47 Binding to Collagen

Initial studies *in vitro* showed that Hsp47 binds to type IV collagen and gelatin (Kurkinen et al., 1984). Jain et al. (1994) found that Hsp47 binds to type I and IV collagen as well as to the individual proα1(I) and proα2(I) chains of procollagen. Natsume et al. (1994) found that it binds to types I-V collagen with similar dissociation constants ($K_d \approx 10^{-7}$ M). This relatively low K_d was shown to result from an association rate constant of $k_{ass} \approx 2 \times 10^4\ M^{-1}s^{-1}$ and a dissociation rate constant of $k_{diss} \approx 10^{-2}s^{-1}$ (Natsume et al., 1994). These authors propose that Hsp47 may therefore undergo a transient interaction with collagen inside of the cell.

Studies of the *in vivo* association of Hsp47 with collagen revealed that as the nascent polysome associated proα1(I) chains of procollagen are translocated into the ER they become bound by Hsp47 (Hu et al., 1995; Sauk et al., 1994). Hsp47 remains bound to procollagen during triple helix formation and during export from the ER (Nakai et al., 1992). Upon entering the *cis*-Golgi network Hsp47 dissociates from the procollagen and is recycled back to the ER (Satoh et al., 1996). If triple-helix formation of procollagen is inhibited, or if transport of procollagen from the ER is blocked, Hsp47 remains associated with the procollagen for an extended period of time (Satoh et al., 1996). The binding of Hsp47 to collagen was found to be pH dependent, in that binding is abolished at pH 6.3 and below (Saga et al., 1987). The pH dependent binding and release of collagen by Hsp47 correlates to the neutral pH of the ER where Hsp47 binds to procollagen, and to the more acidic Golgi apparatus where procollagen is released (Kim

et al., 1998; Kim et al., 1996). Indeed, if intracellular pH is decreased, the association of Hsp47 with procollagen is disrupted (Nakai et al., 1992).

The structural aspects of the binding interactions between Hsp47 and collagen have not been resolved. It is unclear whether it is simply the conformation of collagen that dictates Hsp47 binding or if there exists a recognition sequence necessary for binding. From the results of research to date, it is likely that it is a combination of both. Hu et al. (1995) were the first to examine the sequence/conformation of collagen necessary for Hsp47 binding. They found that Hsp47 was able to weakly bind the synthetic peptide $(Pro\text{-}Pro\text{-}Gly)_{10}$ but not $(Pro\text{-}Hyp\text{-}Gly)_{10}$. Using peptides derived from the N-terminal portion of the proα1(I) chain, they were able to show that Hsp47 binds to the amino propeptide and the N-terminal triple-helical domain (see Figure 1.1). However, these results did not complete the search for Hsp47 interacting domains on collagen. This is because *in vitro* studies had demonstrated that Hsp47 binds tightly to fully processed collagen and gelatin, neither of which contains these regions.

Due to their ability to mimic the collagen molecule, synthetic model peptides have been extensively used to examine Hsp47's interaction with collagen. In a competition assay similar to that used by Hu et al. (1995), it was again shown that Hsp47 binds to $(Pro\text{-}Pro\text{-}Gly)_{10}$ but not to $(Pro\text{-}Hyp\text{-}Gly)_{10}$ and that $(Pro\text{-}Pro\text{-}Gly)_7$ has the minimal length necessary for Hsp47 binding (Koide et al., 1999). In their subsequent work, these authors developed a yeast two-hybrid screen of two peptide libraries representing (Pro-Pro-Gly) -$(X\text{-}Pro\text{-}Gly)_6$ -(Pro-Pro-Gly) and (Pro-Pro-Gly) -$(Pro\text{-}Y\text{-}Gly)_6$ -(Pro-Pro-Gly) and examined Hsp47 preference for binding to these peptides with

random amino acids inserted into either the X or Y position (Koide et al., 2000). They found no binding of Hsp47 to the randomized (Pro-Pro-Gly) -(X-Pro-Gly)$_6$ -(Pro-Pro-Gly) peptide suggesting that Pro is required in the X position. From their data on the (Pro-Pro-Gly) -(Pro-Y-Gly)$_6$ -(Pro-Pro-Gly) peptides which interacted with Hsp47, Koide et al. (2000) were able to calculate enrichment ratios for amino acids residing in the Y position. Interestingly, they found a direct relationship between the enrichment of an amino acid in the Hsp47-interacting peptides and the thermal stabilities of the corresponding model peptide. In particular, Arg and Pro were highly preferred.

In a similar study, Sauk et al. (2000) examined the binding of Hsp47 to two random peptide display libraries. They found increased binding of Hsp47 to peptides containing the hydrophobic amino acids valine, alanine, tryptophan and leucine and proposed that the Hsp47 binding pocket could accommodate a peptide with large hydrophobic amino acids flanked by charged residues. It is odd that the two different techniques to analyse Hsp47 binding to random peptide libraries generated completely different answers. However, it is important to note that the peptides used by Koide et al. (2000) were more realistic in terms of mimicking collagen's structure including triple-helix formation and are thus more likely to represent the true binding motif.

Using a semi-permeabilized cell system, Tasab et al. (2000) demonstrated that Hsp47 only interacts with triple-helical procollagen and not the individual monomer chains. In further studies, these authors designed procollagen-like molecules for use in their semi-permeablilized cell system and assessed the binding of Hsp47 (Tasab et al., 2002). They found that Hsp47 binds to triple-helical sequences with high thermal

stability and containing at least one Arg residue. Hsp47 bound to these sequences regardless of the presence or absence of Hyp residues. Similarly, recent fluorescence quenching assays show that Hsp47 can bind to both $(Pro\text{-}Hyp\text{-}Gly)_{10}$ and $(Pro\text{-}Pro\text{-}Gly)_{10}$ but that the binding is very weak (Macdonald and Bachinger, 2001).

Needless to say, there are some discrepancies among the studies on the binding of Hsp47 to collagen. Hsp47 binding to triple-helical collagen is generally accepted, however, whether or not Hsp47 binds to the individual procollagen chains *in vivo* is debatable. It is unknown where exactly on the collagen molecule Hsp47 binding takes place and what role sequence and conformation play. Arg has been implicated as being important for Hsp47 binding, but whether or not all of the many Arg residues in the collagen molecule are involved remains unresolved. Mapping Hsp47-binding sites on collagen and collagen-binding sites in Hsp47 will undoubtedly help us to understand Hsp47's role in collagen biosynthesis.

1.2.5 Hsp47 Function

The production of an Hsp47 knockout mouse confirmed the belief that Hsp47 is necessary for collagen biosynthesis (Nagai et al., 2000). The mouse was embryonic lethal at 9.5 dpc with aberrant formation of triple-helical type I collagen molecules and with defects in basement membrane formation. However, the exact role of Hsp47 in collagen biosynthesis is unknown. As reviewed by Nagata (1998) and illustrated in Figure 1.4, multiple functions have been ascribed to Hsp47 in terms of its exact role as a collagen-specific molecular chaperone. These include assisting in procollagen

translation; prevention of aggregation of individual proα-chains or of triple-helical collagen molecules in the ER; quality control; transport of procollagen from the ER to the Golgi apparatus; prevention of procollagen degradation; and/or regulation of procollagen modification. The evidence for each of these proposed roles varies, as summarized below.

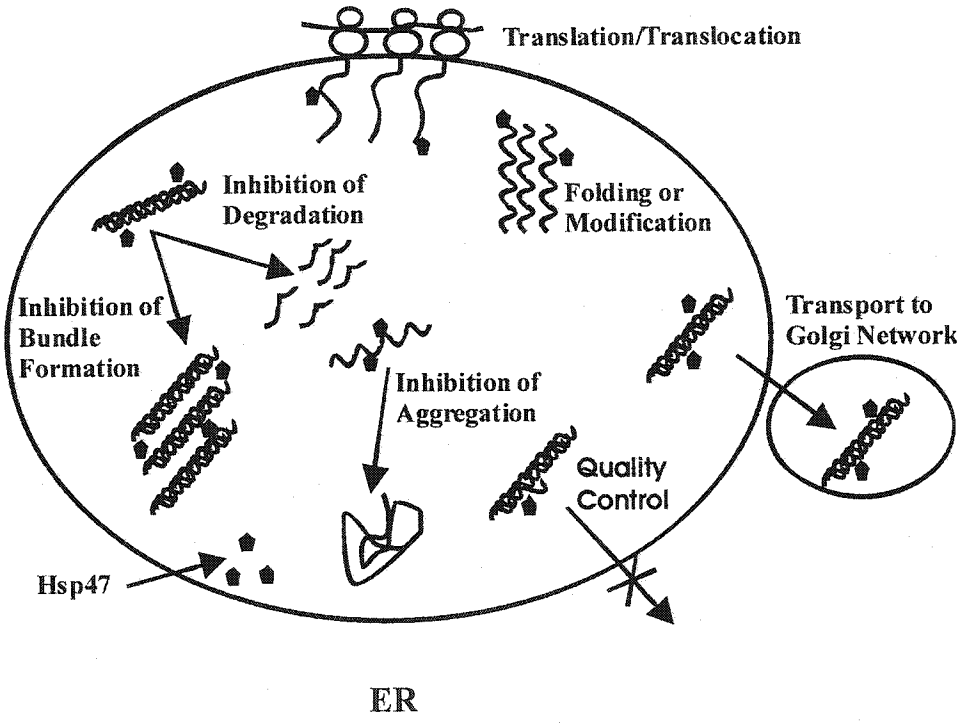


Figure 1.4 Schematic representing the proposed roles of Hsp47 as a collagen-specific molecular chaperone. (Adapted from Nagata, 1998)

1.2.5.1 *Translation/Translocation*

The binding of Hsp47 to nascent procollagen chains during their translocation into the ER implicated Hsp47 in having a role in chaperoning the translocation of the individual procollagen chains (Hu et al., 1995; Sauk et al., 1994). In particular, the binding of Hsp47 to the N-terminal propeptide of the α1(I) chain, led Hu et al. (1995) to propose that Hsp47 may have a role in preventing premature association of the N-propeptides before translation of the C-terminus is completed. It has also been shown that partially elongated proα1(I) chains accumulate when the production of Hsp47 is inhibited by the addition of antisense oligodeoxynucleotides (Sauk et al., 1994). Since Hsp47 is also known to bind to triple-helical procollagen, it is likely that it plays multiple roles in collagen folding.

1.2.5.2 *Regulation of Collagen Modification*

When type III procollagen was stably transfected into a human kidney cell line that does not normally produce collagen or Hsp47, the collagen was secreted but had an altered melting temperature (T_m) (Hosokawa et al., 1998). Conversely, when Hsp47 was cotransfected with the procollagen, the rate of procollagen secretion was slowed down and the procollagen T_m resembled that of normal collagen. The authors of this work concluded that Hsp47 may therefore play a role in collagen modification, such as in regulating levels of hydroxylation. In later studies on the interaction of Hsp47 with collagen, it was shown that Hsp47 and prolyl hydroxylase can compete for non-triple-helical procollagen as a substrate and that Hsp47 can retard the hydroxylation of collagen

model peptides (Asada et al., 1999; Hosokawa and Nagata, 2000). These results suggest a role for Hsp47 in monitoring collagen modification.

1.2.5.3 *Collagen Secretion*

As mentioned previously, Hsp47 associates with procollagen during procollagen export from the ER. Inhibition of ER-Golgi transport results in increased association of Hsp47 with procollagen in the intermediate compartment, whereas association of Hsp47 with procollagen is not observed when intracellular transport within the Golgi is inhibited (Ko and Kay, 1999; Satoh et al., 1996). It is therefore conceivable that Hsp47 may play a role in procollagen transport from the ER. Studies in support of this theory were performed by Tomita et al. (1997), who showed that cotransfection of the α1 and α2 chains of type I collagen into a baculovirus expression system resulted in stable heterotrimer formation but with delayed secretion into the serum. In subsequent studies they showed that by cotransfecting Hsp47 into the expression system, heterotrimeric procollagen was efficiently secreted (Tomita et al. 1999). This work and that of Hosokawa et al. (1998), mentioned earlier, demonstrate the need for Hsp47 in the secretion and formation of stable collagen molecules. However, they also provide conflicting evidence as to the exact role of Hsp47 in collagen biosynthesis.

1.2.5.4 *Prevention of Procollagen Degradation*

As a member of the serpin superfamily it is possible that the role of Hsp47 *in vivo* is to prevent the premature degradation of procollagen. Jain et al. (1994) demonstrated

that Hsp47 does have the ability to prevent the degradation of procollagen in liver microsomes. (However Jain et al. (1994) note that their studies did not discern whether or not Hsp47 directly inhibited protease activity or if it was somehow protecting the procollagen chains from degradation through its binding activity.) By contrast, inhibition studies with recombinant Hsp47 revealed that it has no serpin activity against any of the serine proteases tested (Davids et al., 1995). Even though procollagen does undergo significant intracellular degradation (Ripley et al., 1993) it is unlikely that Hsp47 is directly involved in this process. This is because the degradation of procollagen is not thought to take place in the ER and studies have implicated the procollagen degradative pathway as being distinct from that in which Hsp47 is found (Ko and Kay, 1999; Ripley et al., 1993).

1.2.5.5 *Quality Control*

During the unfolded protein response, the expression of stress proteins such as BiP and Grp94 is induced as a result of the accumulation of unfolded proteins in the ER (Gething and Sambrook, 1992). Similarly, under conditions of heat shock and cell stress, Hsp47 expression is induced (Higashi et al., 1994; Kudo et al., 1994). Examination of fibroblast cells under heat shock conditions (45°C), revealed decreased procollagen secretion coinciding with the increased association of procollagen and Hsp47 (Nakai et al., 1992). If the role of Hsp47 is similar to that of other molecular chaperones, it could be involved in recognizing aberrant procollagen molecules under stress and disease conditions. This theory is supported by studies showing that if proper procollagen

folding is disrupted, it remains bound by Hsp47 (Hosokawa and Nagata, 2000). Similarly, in cells cultured from a patient with lethal osteogenesis imperfecta, the accumulation of abnormal procollagen molecules in the ER results in the increased expression and co-localization of Hsp47 (Kojima et al., 1998). It is possible that studies of Hsp47 expression and procollagen binding under stress conditions may provide the insight necessary to determine the role of Hsp47 in a healthy collagen-producing cell.

1.2.5.6 *Inhibition of Aggregation*

The triple-helical conformation of collagen places the side chains of amino acids on the outer periphery of the helix and exposed to solvent. This necessitates a means of protecting collagen from misfolding and aggregation both before and after triple helix formation. As well, the process of fibrillogenesis, which involves the association of collagen triple-helices, is a spontaneous one that could prove detrimental if allowed to occur in the lumen of the ER. Hsp47 may therefore play a role in preventing the premature association of collagen helices in the ER. This suggestion is not unreasonable, since the prevention of premature polypeptide associations is one that is often attributed to members of the heat shock family of proteins (Blond-Elguindi et al., 1993). Experiments described later in this thesis have demonstrated that Hsp47 effectively inhibits type I collagen fibrillogenesis *in vitro* (Thomson and Ananthanarayanan, 2000). This result was reproduced by Dafforn et al. (2001) using type I collagen and by MacDonald and Bachinger (2001) using type III collagen. Further investigation is

required however, to determine whether or not Hsp47 inhibits collagen associations *in vivo* and whether or not it can prevent the aggregation of individual proα chains.

1.2.6 Hsp47 and Disease

The parallel regulation of Hsp47 and collagen expression is apparent in many disease states in which changes in collagen expression are observed. Included in this list are various fibrotic diseases of the liver, kidney, lung and dermis, as well as atheroslerosis and cancer (Masuda et al., 1994; Nagata et al., 1986; Razzaque and Ahmed, 2002; Razzaque et al., 1998; Razzaque and Taguchi, 1997; Rocnik et al., 2000). The increased expression of Hsp47 has been linked to a variety of factors. In particular transforming growth factor-β1 has been implicated as being a key modulator of this process (Razzaque and Ahmed, 2002). Interestingly, the expression of Hsp47 has been shown to have profound effects on collagen production. For instance, the use of antisense oligodeoxynucleotides to decrease Hsp47 expression was sufficient to attenuate the accumulation of collagen in a glomerulonephritis rat model (Sunamoto et al., 1998). In smooth muscle cells, the overexpression of Hsp47 caused increased procollagen expression and secretion (Rocnik et al., 2002). These results point to Hsp47 as having an important role in regulating collagen production in pathophysiological conditions. Therefore Hsp47 may be a relevant target for drug design in order to combat diseases correlated to collagen deposition.

1.3 SUMMARY

The importance of Hsp47 in collagen biosynthesis is well understood. However, further investigation is required to understand the exact *in vivo* role(s) of Hsp47 as a collagen-specific molecular chaperone. Therefore, the main objective of this thesis is to understand the function of Hsp47 in terms of its structure and the structure of collagen. Specifically, in this thesis I will describe my work on: 1) purification of recombinant Hsp47; 2) biophysical characterization of Hsp47; 3) inhibition of collagen fibrillogenesis by Hsp47; and 4) mapping binding of Hsp47 to specific regions of the collagen molecule. I will then conclude with a brief discussion on the implications of the results obtained, followed by a discussion of the future prospects of Hsp47 related research.

1.4 REFERENCES

Aguilar, J. H., Jacobs, H. G., Butler, W. T., and Cunningham, L. W. (1973). The distribution of carbohydrate groups in rat skin collagen. J Biol Chem *248*, 5106-13.

Amin, J., Ananthan, J., and Voellmy, R. (1988). Key features of heat shock regulatory elements. Mol Cell Biol *8*, 3761-9.

Asada, S., Koide, T., Yasui, H., and Nagata, K. (1999). Effect of HSP47 on prolyl 4-hydroxylation of collagen model peptides. Cell Struct Funct *24*, 187-96.

Bachinger, H. P., Bruckner, P., Timpl, R., Prockop, D. J., and Engel, J. (1980). Folding mechanism of the triple helix in type-III collagen and type-III pN-collagen. Role of disulfide bridges and peptide bond isomerization. Eur J Biochem *106*, 619-32.

Bachinger, H. P., Morris, N. P., and Davis, J. M. (1993). Thermal stability and folding of the collagen triple helix and the effects of mutations in osteogenesis imperfecta on the triple helix of type I collagen. Am J Med Genet *45*, 152-62.

Beck, K., Boswell, B. A., Ridgway, C. C., and Bachinger, H. P. (1996). Triple helix formation of procollagen type I can occur at the rough endoplasmic reticulum membrane. J Biol Chem *271*, 21566-73.

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

Berg, R. A., and Prockop, D. J. (1973). The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple-helix of collagen. Biochem Biophys Res Commun *52*, 115-20.

Blond-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F., and Gething, M. J. (1993). Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. Cell *75*, 717-28.

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

Brodsky, B., and Ramshaw, J. A. (1997). The collagen triple-helix structure. Matrix Biol *15*, 545-54.

Bruckner, P., and Eikenberry, E. F. (1984). Procollagen is more stable in cellulo than in vitro. Eur J Biochem *140*, 397-9.

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

Bruns, R. R., Hulmes, D. J., Therrien, S. F., and Gross, J. (1979). Procollagen segment-long-spacing crystallites: their role in collagen fibrillogenesis. Proc Natl Acad Sci U S A *76*, 313-7.

Burgeson, R. E., and Nimni, M. E. (1992). Collagen types. Molecular structure and tissue distribution. Clin Orthop, 250-72.

Chan, V. C., Ramshaw, J. A., Kirkpatrick, A., Beck, K., and Brodsky, B. (1997). Positional preferences of ionizable residues in Gly-X-Y triplets of the collagen triple-helix. J Biol Chem *272*, 31441-6.

Chanut-Delalande, H., Fichard, A., Bernocco, S., Garrone, R., Hulmes, D. J., and Ruggiero, F. (2001). Control of heterotypic fibril formation by collagen V is determined by chain stoichiometry. J Biol Chem *276*, 24352-9.

Clarke, E. P., Cates, G. A., Ball, E. H., and Sanwal, B. D. (1991). A collagen-binding protein in the endoplasmic reticulum of myoblasts exhibits relationship with serine protease inhibitors. J Biol Chem *266*, 17230-5.

Clarke, E. P., and Sanwal, B. D. (1992). Cloning of a human collagen-binding protein, and its homology with rat gp46, chick hsp47 and mouse J6 proteins. Biochim Biophys Acta *1129*, 246-8.

Dafforn, T. R., Della, M., and Miller, A. D. (2001). The molecular interactions of heat shock protein 47 (Hsp47) and their implications for collagen biosynthesis. J Biol Chem *276*, 49310-9.

Davids, J., El-Thaher, T., Nakai, A., Nagata, K., Miller, A. (1995). Modeling the Three-Dimensional Structure of Serpin/Molecular Chaperone Hsp47. Bioorganic Chemistry *23*, 427-438.

Engel, J., Chen, H. T., Prockop, D. J., and Klump, H. (1977). The triple helix in equilibrium with coil conversion of collagen-like polytripeptides in aqueous and nonaqueous solvents. Comparison of the thermodynamic parameters and the binding of water to (L-Pro-L-Pro-Gly)n and (L-Pro-L-Hyp-Gly)n. Biopolymers *16*, 601-22.

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

Ferreira, L. R., Norris, K., Smith, T., Hebert, C., and Sauk, J. J. (1994). Association of Hsp47, Grp78, and Grp94 with procollagen supports the successive or coupled action of molecular chaperones. J Cell Biochem *56*, 518-26.

Gelman, R. A., Williams, B. R., and Piez, K. A. (1979). Collagen fibril formation. Evidence for a multistep process. J Biol Chem *254*, 180-6.

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

Graham, H. K., Holmes, D. F., Watson, R. B., and Kadler, K. E. (2000). Identification of collagen fibril fusion during vertebrate tendon morphogenesis. The process relies on unipolar fibrils and is regulated by collagen-proteoglycan interaction. J Mol Biol *295*, 891-902.

Higashi, T., Takechi, H., Uemura, Y., Kikuchi, H., and Nagata, K. (1994). Differential induction of mRNA species encoding several classes of stress proteins following focal cerebral ischemia in rats. Brain Res *650*, 239-48.

Hirata, H., Yamamura, I., Yasuda, K., Kobayashi, A., Tada, N., Suzuki, M., Hirayoshi, K., Hosokawa, N., and Nagata, K. (1999). Separate cis-acting DNA elements control cell type- and tissue-specific expression of collagen binding molecular chaperone HSP47. J Biol Chem *274*, 35703-10.

Hirayoshi, K., Kudo, H., Takechi, H., Nakai, A., Iwamatsu, A., Yamada, K. M., and Nagata, K. (1991). HSP47: a tissue-specific, transformation-sensitive, collagen-binding heat shock protein of chicken embryo fibroblasts. Mol Cell Biol *11*, 4036-44.

Hopkins, P. C., Carrell, R. W., and Stone, S. R. (1993). Effects of mutations in the hinge region of serpins. Biochemistry *32*, 7650-7.

Hosokawa, N., Hohenadl, C., Satoh, M., Kuhn, K., and Nagata, K. (1998). HSP47, a collagen-specific molecular chaperone, delays the secretion of type III procollagen transfected in human embryonic kidney cell line 293: a possible role for HSP47 in collagen modification. J Biochem (Tokyo) *124*, 654-62.

Hosokawa, N., and Nagata, K. (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*, 19-25.

Hosokawa, N., Takechi, H., Yokota, S., Hirayoshi, K., and Nagata, K. (1993). Structure of the gene encoding the mouse 47-kDa heat-shock protein (HSP47). Gene *126*, 187-93.

Hu, G., Gura, T., Sabsay, B., Sauk, J., Dixit, S. N., and Veis, A. (1995). Endoplasmic reticulum protein Hsp47 binds specifically to the N- terminal globular domain of the amino-propeptide of the procollagen I alpha 1 (I)-chain. J Cell Biochem *59*, 350-67.

Hughes, R. C., Taylor, A., Sage, H., and Hogan, B. L. (1987). Distinct patterns of glycosylation of colligin, a collagen-binding glycoprotein, and SPARC (osteonectin), a secreted Ca2+-binding glycoprotein. Evidence for the localisation of colligin in the endoplasmic reticulum. Eur J Biochem *163*, 57-65.

Hulmes, D. J. (2002). Building collagen molecules, fibrils, and suprafibrillar structures. J Struct Biol *137*, 2-10.

Hulmes, D. J., Miller, A., Parry, D. A., Piez, K. A., and Woodhead-Galloway, J. (1973). Analysis of the primary structure of collagen for the origins of molecular packing. J Mol Biol *79*, 137-48.

Hutton, J. J., Jr., Kaplan, A., and Udenfriend, S. (1967). Conversion of the amino acid sequence gly-pro-pro in protein to gly-pro- hyp by collagen proline hydroxylase. Arch Biochem Biophys *121*, 384-91.

Hayashi, T., Curran-Patel, S., and Prockop, D. J. (1979). Thermal stability of the triple helix of type I procollagen and collagen. Precautions for minimizing ultraviolet damage to proteins during circular dichroism studies. Biochemistry *18*, 4182-7.

Irving, J. A., Pike, R. N., Lesk, A. M., and Whisstock, J. C. (2000). Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. Genome Res *10*, 1845-64.

Jain, N., Brickenden, A., Ball, E. H., and Sanwal, B. D. (1994). Inhibition of procollagen I degradation by colligin: a collagen-binding serpin. Arch Biochem Biophys *314*, 23-30.

Jain, N., Brickenden, A., Lorimer, I., Ball, E. H., and Sanwal, B. D. (1994). Interaction of procollagen I and other collagens with colligin. Biochem J *304*, 61-8.

John, D. C., Grant, M. E., and Bulleid, N. J. (1993). Cell-free synthesis and assembly of prolyl 4-hydroxylase: the role of the beta-subunit (PDI) in preventing misfolding and aggregation of the alpha-subunit. Embo J *12*, 1587-95.

Kadler, K. E., Hojima, Y., and Prockop, D. J. (1988). Assembly of type I collagen fibrils de novo. Between 37 and 41 degrees C the process is limited by micro-unfolding of monomers. J Biol Chem *263*, 10517-23.

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

Kim, J. H., Johannes, L., Goud, B., Antony, C., Lingwood, C. A., Daneman, R., and Grinstein, S. (1998). Noninvasive measurement of the pH of the endoplasmic reticulum at rest and during calcium release. Proc Natl Acad Sci U S A *95*, 2997-3002.

Kim, J. H., Lingwood, C. A., Williams, D. B., Furuya, W., Manolson, M. F., and Grinstein, S. (1996). Dynamic measurement of the pH of the Golgi complex in living cells using retrograde transport of the verotoxin receptor. J Cell Biol *134*, 1387-99.

Kivirikko, K. I., Kishida, Y., Sakakibara, S., and Prockop, D. J. (1972). Hydroxylation of (X-Pro-Gly)n by protocollagen proline hydroxylase. Effect of chain length, helical conformation and amino acid sequence in the substrate. Biochim Biophys Acta *271*, 347-56.

Ko, M. K., and Kay, E. P. (1999). Hsp47-dependent and -independent intracellular trafficking of type I collagen in corneal endothelial cells. Mol Vis *5*, 17.

Koide, T., Asada, S., and Nagata, K. (1999). Substrate recognition of collagen-specific molecular chaperone HSP47. Structural requirements and binding regulation. J Biol Chem *274*, 34523-6.

Koide, T., Aso, A., Yorihuzi, T., and Nagata, K. (2000). Conformational requirements of collagenous peptides for recognition by the chaperone protein HSP47. J Biol Chem *275*, 27957-63.

Kojima, T., Miyaishi, O., Saga, S., Ishiguro, N., Tsutsui, Y., and Iwata, H. (1998). The retention of abnormal type I procollagen and correlated expression of HSP 47 in fibroblasts from a patient with lethal osteogenesis imperfecta. J Pathol *184*, 212-8.

Kramer, R. Z., Bella, J., Brodsky, B., and Berman, H. M. (2001). The crystal and molecular structure of a collagen-like peptide with a biologically relevant sequence. J Mol Biol *311*, 131-47.

Kudo, H., Hirayoshi, K., Kitagawa, Y., Imamura, S., and Nagata, K. (1994). Two collagen-binding proteins, osteonectin and HSP47, are coordinately induced in transformed keratinocytes by heat and other stresses. Exp Cell Res *212*, 219-24.

Kurkinen, M., Taylor, A., Garrels, J. I., and Hogan, B. L. (1984). Cell surface-associated proteins which bind native type IV collagen or gelatin. J Biol Chem *259*, 5915-22.

Kuznetsova, N., and Leikin, S. (1999). Does the triple helical domain of type I collagen encode molecular recognition and fiber assembly while telopeptides serve as catalytic domains? Effect of proteolytic cleavage on fibrillogenesis and on collagen-collagen interaction in fibers. J Biol Chem *274*, 36083-8.

Lamande, S. R., and Bateman, J. F. (1999). Procollagen folding and assembly: the role of endoplasmic reticulum enzymes and molecular chaperones. Semin Cell Dev Biol *10*, 455-64.

Leblond, C. P. (1989). Synthesis and secretion of collagen by cells of connective tissue, bone, and dentin. Anat Rec *224*, 123-38.

Lees, J. F., Tasab, M., and Bulleid, N. J. (1997). Identification of the molecular recognition sequence which determines the type-specific assembly of procollagen. Embo J *16*, 908-16.

Leikina, E., Mertts, M. V., Kuznetsova, N., and Leikin, S. (2002). Type I collagen is thermally unstable at body temperature. Proc Natl Acad Sci U S A *99*, 1314-8.

Macdonald, J. R., and Bachinger, H. P. (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*, 25399-403.

Masuda, H., Fukumoto, M., Hirayoshi, K., and Nagata, K. (1994). Coexpression of the collagen-binding stress protein HSP47 gene and the alpha 1(I) and alpha 1(III) collagen genes in carbon tetrachloride- induced rat liver fibrosis. J Clin Invest *94*, 2481-8.

Miyaishi, O., Sakata, K., Matsuyama, M., and Saga, S. (1992). Distribution of the collagen binding heat-shock protein in chicken tissues. J Histochem Cytochem *40*, 1021-9.

Murphy, L., and Rosenbloom, J. (1973). Evidence that chick tendon procollagen must be denatured to serve as substrate for proline hydroxylase. Biochem J *135*, 249-51.

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

Nagai, N., Hosokawa, M., Itohara, S., Adachi, E., Matsushita, T., Hosokawa, N., and Nagata, K. (2000). Embryonic lethality of molecular chaperone hsp47 knockout mice is associated with defects in collagen biosynthesis. J Cell Biol *150*, 1499-506.

Nagai, N., Tetuya, Y., Hosokawa, N., and Nagata, K. (1999). The human genome has only one functional hsp47 gene (CBP2) and a pseudogene (pshsp47). Gene *227*, 241-8.

Nagata, K. (1998). Expression and function of heat shock protein 47: a collagen-specific molecular chaperone in the endoplasmic reticulum. Matrix Biol *16*, 379-86.

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

Nagata, K., Saga, S., and Yamada, K. M. (1986). A major collagen-binding protein of chick embryo fibroblasts is a novel heat shock protein. J Cell Biol *103*, 223-9.

Nagata, K., and Yamada, K. M. (1986). Phosphorylation and transformation sensitivity of a major collagen- binding protein of fibroblasts. J Biol Chem *261*, 7531-6.

Nakai, A., Satoh, M., Hirayoshi, K., and Nagata, K. (1992). Involvement of the stress protein HSP47 in procollagen processing in the endoplasmic reticulum. J Cell Biol *117*, 903-14.

Nandan, D., Cates, G. A., Ball, E. H., and Sanwal, B. D. (1990). Partial characterization of a collagen-binding, differentiation-related glycoprotein from skeletal myoblasts. Arch Biochem Biophys *278*, 291-6.

Natsume, T., Koide, T., Yokota, S., Hirayoshi, K., and Nagata, K. (1994). Interactions between collagen-binding stress protein HSP47 and collagen. Analysis of kinetic parameters by surface plasmon resonance biosensor. J Biol Chem *269*, 31224-8.

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

Pak, B. J., Vanhorne, J. B., and Pang, S. C. (1996). Tissue distribution and immunohistochemical localization of the collagen-binding heat-shock protein gp46 in neonatal rats. Histochem J *28*, 895-902.

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

Prockop, D. J., and Fertala, A. (1998). Inhibition of the self-assembly of collagen I into fibrils with synthetic peptides. Demonstration that assembly is driven by specific binding sites on the monomers. J Biol Chem *273*, 15598-604.

Prockop, D. J., Kivirikko, K. I., Tuderman, L., and Guzman, N. A. (1979). The biosynthesis of collagen and its disorders (first of two parts). N Engl J Med *301*, 13-23.

Prockop, D. J., Sieron, A. L., and Li, S. W. (1998). Procollagen N-proteinase and procollagen C-proteinase. Two unusual metalloproteinases that are essential for procollagen processing probably have important roles in development and cell signaling. Matrix Biol *16*, 399-408.

Ramachandran, G., and Ramakrishnan, C. (1976). Molecular Structure. In Biochemistry of Collagen, G. Ramachandran, and Reddi, A., ed. (New York and London: Plenum Press), pp. 45-84.

Razzaque, M. S., and Ahmed, A. R. (2002). Collagens, collagen-binding heat shock protein 47 and transforming growth factor-beta 1 are induced in cicatricial pemphigoid: possible role(s) in dermal fibrosis. Cytokine *17*, 311-6.

Razzaque, M. S., Nazneen, A., and Taguchi, T. (1998). Immunolocalization of collagen and collagen-binding heat shock protein 47 in fibrotic lung diseases. Mod Pathol *11*, 1183-8.

Razzaque, M. S., and Taguchi, T. (1997). Collagen-binding heat shock protein (HSP) 47 expression in anti- thymocyte serum (ATS)-induced glomerulonephritis. J Pathol *183*, 24-9.

Ripley, C. R., Fant, J., and Bienkowski, R. S. (1993). Brefeldin A inhibits degradation as well as production and secretion of collagen in human lung fibroblasts. J Biol Chem *268*, 3677-82.

Rocnik, E., Chow, L. H., and Pickering, J. G. (2000). Heat shock protein 47 is expressed in fibrous regions of human atheroma and Is regulated by growth factors and oxidized low-density lipoprotein. Circulation *101*, 1229-33.

Rocnik, E. F., van der Veer, E., Cao, H., Hegele, R. A., and Pickering, J. G. (2002). Functional linkage between the endoplasmic reticulum protein Hsp47 and procollagen expression in human vascular smooth muscle cells. J Biol Chem *277*, 38571-8.

Ryhanen, L., Zaragoza, E. J., and Uitto, J. (1983). Conformational stability of type I collagen triple helix: evidence for temporary and local relaxation of the protein conformation using a proteolytic probe. Arch Biochem Biophys *223*, 562-71.

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

Sato, K., Yomogida, K., Wada, T., Yorihuzi, T., Nishimune, Y., Hosokawa, N., and Nagata, K. (2002). Type XXVI collagen, a new member of the collagen family, is specifically expressed in the testis and ovary. J Biol Chem *277*, 37678-84.

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

Sauk, J. J., Coletta, R. D., Norris, K., and Hebert, C. (2000). Binding motifs of CBP2 a potential cell surface target for carcinoma cells. J Cell Biochem *78*, 251-63.

Sauk, J. J., Smith, T., Norris, K., and Ferreira, L. (1994). Hsp47 and the translation-translocation machinery cooperate in the production of alpha 1(I) chains of type I procollagen. J Biol Chem *269*, 3941-6.

Siegel, R. C. (1976). Collagen cross-linking. Synthesis of collagen cross-links in vitro with highly purified lysyl oxidase. J Biol Chem *251*, 5786-92.

Smith, T., Ferreira, L. R., Hebert, C., Norris, K., and Sauk, J. J. (1995). Hsp47 and cyclophilin B traverse the endoplasmic reticulum with procollagen into pre-Golgi intermediate vesicles. A role for Hsp47 and cyclophilin B in the export of procollagen from the endoplasmic reticulum. J Biol Chem *270*, 18323-8.

Steinmann, B., Bruckner, P., and Superti-Furga, A. (1991). Cyclosporin A slows collagen triple-helix formation in vivo: indirect evidence for a physiologic role of peptidyl-prolyl cis-trans-isomerase. J Biol Chem *266*, 1299-303.

Stephens, D. J., and Pepperkok, R. (2002). Imaging of procollagen transport reveals COPI-dependent cargo sorting during ER-to-Golgi transport in mammalian cells. J Cell Sci *115*, 1149-60.

Sunamoto, M., Kuze, K., Tsuji, H., Ohishi, N., Yagi, K., Nagata, K., Kita, T., and Doi, T. (1998). Antisense oligonucleotides against collagen-binding stress protein HSP47 suppress collagen accumulation in experimental glomerulonephritis. Lab Invest *78*, 967-72.

Takechi, H., Hosokawa, N., Hirayoshi, K., and Nagata, K. (1994). Alternative 5' splice site selection induced by heat shock. Mol Cell Biol *14*, 567-75.

Takeda, K., Kusachi, S., Ohnishi, H., Nakahama, M., Murakami, M., Komatsubara, I., Oka, T., Doi, M., Ninomiya, Y., and Tsuji, T. (2000). Greater than normal expression of the collagen-binding stress protein heat-shock protein-47 in the infarct zone in rats after experimentally- induced myocardial infarction. Coron Artery Dis *11*, 57-68.

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

Tasab, M., Jenkinson, L., and Bulleid, N. J. (2002). Sequence-specific recognition of collagen triple helices by the collagen-specific molecular chaperone HSP47. J Biol Chem *277*, 35007-12.

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

Tomita, M., Kitajima, T., and Yoshizato, K. (1997). Formation of recombinant human procollagen I heterotrimers in a baculovirus expression system. J Biochem (Tokyo) *121*, 1061-9.

Tomita, M., Yoshizato, K., Nagata, K., and Kitajima, T. (1999). Enhancement of secretion of human procollagen I in mouse HSP47- expressing insect cells. J Biochem (Tokyo) *126*, 1118-26.

Trelstad, R. L., and Hayashi, K. (1979). Tendon collagen fibrillogenesis: intracellular subassemblies and cell surface changes associated with fibril growth. Dev Biol *71*, 228-42.

van der Rest, M., and Garrone, R. (1991). Collagen family of proteins. Faseb J *5*, 2814-23.

Vuori, K., Pihlajaniemi, T., Myllyla, R., and Kivirikko, K. I. (1992). Site-directed mutagenesis of human protein disulphide isomerase: effect on the assembly, activity and endoplasmic reticulum retention of human prolyl 4-hydroxylase in Spodoptera frugiperda insect cells. Embo J *11*, 4213-7.

Wang, S. Y., and Gudas, L. J. (1990). A retinoic acid-inducible mRNA from F9 teratocarcinoma cells encodes a novel protease inhibitor homologue. J Biol Chem *265*, 15818-22.

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

Yang, W., Chan, V. C., Kirkpatrick, A., Ramshaw, J. A., and Brodsky, B. (1997). Gly-Pro-Arg confers stability similar to Gly-Pro-Hyp in the collagen triple-helix of host-guest peptides. J Biol Chem *272*, 28837-40.

Yonath, A., and Traub, W. (1969). Polymers of tripeptides as collagen models. IV. Structure analysis of poly(L-proly-glycyl-L-proline). J Mol Biol *43*, 461-77.

CHAPTER 2: STRUCTURE-FUNCTION STUDIES ON HSP47: pH-DEPENDENT INHIBITION OF COLLAGEN FIBRIL FORMATION *IN VITRO*

Christy A. THOMSON and Vettai S. ANANTHANARAYANAN (2000) *Biochemical Journal*, **349**, 877-883

2.1 PREFACE

Although Hsp47 had long been understood to be a collagen-specific molecular chaperone, no *in vitro* evidence had been generated in support of its proposed *in vivo* functions. Similarly, no detailed biophysical characterization of Hsp47 had been carried out. Using a novel Hsp47 expression system for producing pure recombinant Hsp47 in *E. coli* we have shown that Hsp47 inhibits type I collagen fibrillogenesis *in vitro*. It had been previously proposed that Hsp47 might function to prevent premature aggregation of collagen in the ER. The inhibition of collagen fibre formation by Hsp47 provides *in vitro* evidence for this proposed role. It also provides an *in vitro* assay for Hsp47 activity.

Circular dichroism (CD) and fluorescence spectroscopy were used to structurally characterize the recombinant Hsp47. Hsp47 was shown to undergo pH-induced conformational changes that correlated well to its pH-dependent binding and release of collagen. Hsp47 also demonstrated binding to ANS. ANS is a fluorescent dye that binds to exposed hydrophobic regions of proteins.

I performed all of the experimental work presented in this publication during the course of my Ph.D. research. Professor N. Greenfield, Rutgers University, NJ, assisted in the secondary structure analysis presented in Table 1.

2.2 PAPER

Structure-function studies on Hsp47. pH-dependent inhibition of collagen fibril formation in vitro

Short Title: Hsp47 inhibits *in vitro* collagen fibril formation

Christy A. THOMSON* and Vettai S. ANANTHANARAYANAN *[1]

***Department of Biochemistry, McMaster University, Hamilton, Ontario, Canada L8N 3Z5**

[1]To whom correspondence should be addressed. (Tel: 905-525-9140, x-22783; Fax: 905-522-9033; E-mail: ananth@fhs.mcmaster.ca)

Key Words: collagen chaperone, collagen aggregation, functional assay, circular dichroism, fluorescence

Abbreviations used: Hsp47, 47-kDa heat-shock protein; ER, endoplasmic reticulum; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; CD, circular dichroism; IPTG, isopropyl-β-D-thiogalactopyranoside; ANS, 1-anilino-8-naphthalene sulfonate; RSV, Rous sarcoma virus; SV40, simian virus 40.

2.3 ABSTRACT

Hsp47, a 47 kDa protein whose expression level parallels that of collagen, has been regarded as a collagen-specific molecular chaperone. Studies from other laboratories have established the association of Hsp47 with the nascent as well as triple helical procollagen molecule in the ER and its dissociation from procollagen in the Golgi. One of the several roles suggested for Hsp47 in collagen biosynthesis is the prevention of aggregation of procollagen in the ER. No experimental evidence has, however, been available to verify this suggestion. In this study, we have followed the aggregation of mature triple helical collagen molecules into fibrils using turbidimetric measurements in the absence and presence of Hsp47. In the pH range of 6 to 7, fibril formation of type I collagen, as monitored by turbidimetry, proceeds with a lag of about 10 min and levels off by about 60 min. Addition of Hsp47 at pH 7 effectively inhibits fibril formation at and above a 1:1 molar ratio of Hsp47 to triple helical collagen. This inhibition is markedly pH-dependent, being significantly diminished at pH 6. Circular dichroism and fluorescence spectral data of Hsp47 in the pH range between 4.2 and 7.4 reveal a significant alteration in its structure at pH values below 6.2, with a reduction in α-helix and an increase in β-structure. This conformational change is likely to be the basis of the decreased binding of Hsp47 to collagen *in vitro* at pH 6.3 as well as its inability to inhibit collagen fibril formation at this pH. Our results also provide a functional assay for Hsp47 which may be used in studies on collagen and Hsp47 interactions.

2.4 INTRODUCTION

Before its final emergence as the major fibrillar component of the connective tissues, the collagen molecule undergoes many posttranslational modifications. Most, if not all, of these modifications involve specific enzymes such as peptidyl proline isomerase , proline hydroxylase, lysine hydroxylase, protein disulfide isomerase, lysine oxidase and procollagen peptidase (Prockop and Kivirikko, 1995). These modifications provide several control points along collagen's biosynthetic pathway. The assembly of the three procollagen chains, with all the peptide bonds in the *trans* conformation, into the triple-helical conformation in the endoplasmic reticulum (ER) and its subsequent secretion of the tropocollagen molecule into the extracellular space have long been regarded as spontaneous processes that do not require the help of other proteins. This view has changed since discovery of the collagen-binding heat-shock protein Hsp47 which is regarded as a collagen-specific chaperone (Kurkinen et al., 1984; Nagata, 1998; Nagata, 1996). Hsp47 has been shown to be associated with procollagen very early in the translocation of the nascent chains (Cates et al., 1987; Sauk et al., 1994) and it interacts with the triple-helical procollagen in the ER and during its export via the secretary pathway (Smith et al., 1995). Once in the *cis*-Golgi compartment, Hsp47 disscociates from procollagen and is recycled back to the ER (Satoh et al., 1996).

The following studies have lent additional support for considering Hsp47 as a collagen-specific chaperone. Hsp47 is expressed only in cells which produce any of a variety of collagen molecules and its content correlates with the abundance of collagen in

a given tissue (Nagata, 1996). Conditions which alter collagen levels affect Hsp47 expression levels as well, the only exception being heat shock when Hsp47 levels rise independent of collagen (Hirayoshi et al., 1991). Thus: when the *in vivo* folding of collagen is retarded, the rates of Hsp47 release from procollagen and procollagen secretion are together delayed (Smith et al., 1995). The transformation of cells by RSV and SV40 causes decreased collagen and Hsp47 syntheses (Nagata and Yamada, 1986; Nakai et al., 1990). Similarly, an increase in the levels of both the proteins has been observed during the process of fibrosis (Masuda et al., 1994). Increase in Hsp47 levels along with increased retention of collagen in the ER has been noted in osteogenesis imperfecta fibroblasts (Kojima et al., 1998). Administration of antisense Hsp47 oligonucleotides suppresses collagen production in glomerulonephritis (Sunamoto et al., 1998).

On the basis of the above results, several roles for Hsp47 as a collagen chaperone in the ER have been proposed (Nagata, 1998). These include quality control, collagen transport to the Golgi, inhibition of collagen degradation and inhibition of aggregation and bundle formation. Whether Hsp47 is involved in any or all of these chaperone activities is, as yet, unclear since no clear-cut experimental evidence is available to substantiate any of its suggested roles. Furthermore, there is no direct correlation between the binding of Hsp47 to collagen and its chaperone activity.

The absence of knowledge about a definitive function for Hsp47 in collagen biosynthesis has stood in the way of detailed structure-function correlation studies on this

protein. With a view to understanding the structural aspects of its interaction with collagen, we have initiated a systematic structural study on purified recombinant Hsp47 using physical chemical techniques (Thomson and Ananthanarayanan, 1999). It is known that Hsp47 binds to both native and denatured collagen (Nagata, 1998). Lowering the pH to 6.3 from neutral pH abolishes this binding (Saga et al., 1987). In this paper, we present data that show that Hsp47 effectively prevents the *in vitro* aggregation of collagen molecules into fibrils at neutral but not acidic pH. The pH-dependence of the fibril inhibition function of Hsp47 correlates with the pH-dependence of its structure as monitored by spectroscopic methods. Our findings provide a functional assay for Hsp47. In addition, they could also be regarded as *in vitro* experimental support for a suggested chaperone role for this protein in preventing the premature association of procollagen chains within the ER and during its transport from the ER to the Golgi.

2.5 MATERIALS AND METHODS

2.5.1 Cloning and Purification of Hsp47

The cDNA for rat gp46/Hsp47, obtained from Dr. Sanwal, was cloned into the pTYB4 vector of the IMPACT^{TM} T7 system from New England Biolabs (Mississauga, Canada) and transformed into the BL21 *E. coli* strain. The pTYB4 vector adds a single glycine residue to the C-terminus of the targeted protein. The IMPACT^{TM} system uses a protein splicing element, intein, which undergoes a self-cleavage reaction at its N-terminus in the presence of thiols such as β-mercaptoethanol. The pTYB4 vector creates

a fusion between the C-terminus of Hsp47 and the N-terminus of the intein. The intein, in turn, is fused to a 5 kDa chitin-binding domain which allows affinity purification of the three-part fusion protein. Recombinant *E. coli* BL21/pTYB4 was grown at 37°C in 3 litres of LB-medium, containing 100 μg/ml ampicillin. At an $OD_{600nm} > 0.4$, the bacterial cells were induced with IPTG (1 mM) and left shaking overnight at 14° C. The cells were harvested by centrifugation at 4000g for 12 min and resuspended in 40 ml of lysis buffer (100 mM Tris, pH 7.5, 500 mM NaCl, 0.5 mM EDTA, 0.1% Triton X-100 and four complete mini-protease tablets (Boehringer, Laval, Canada). Cells were lysed using a French press and the debris was spun down at 12000g for 12 min. The cell lysate supernatant was mixed with 20 ml of chitin beads for 1 h at 4° C, and the chitin beads washed extensively in 50 mM Tris, pH 7.4, 1M NaCl, 0.5% Triton X-100 followed by 20 mM sodium phosphate buffer, pH 7.4, 50 mM NaCl. To the final wash, 100 mM β-mercaptoethanol was added and the chitin beads were incubated at 4° C for 48 h to effectively cleave Hsp47 from the intein element. To purify the resulting Hsp47 from smaller molecular weight degradation products, the fractions containing Hsp47 in the eluant were collected and loaded on to a 3 ml hydroxylapatite column. The bound Hsp47 was eluted using an increasing phosphate gradient and the fraction containing pure Hsp47 was dialyzed in 20 mM sodium phosphate buffer, pH 7.4 containing 50 mM NaCl.

2.5.2 CD and Fluorescence Spectroscopy

Circular dichroism (CD) spectra were recorded using a Jasco J-600

spectropolarimeter. Data were collected from 195-250 nm at 0.2 nm intervals and a time constant of 0.5 seconds and were corrected for buffer contribution. A 1.0 mm quartz cell was used. The protein concentration was about 300 μg/ml and was calculated using an extinction coefficient of 42634 $M^{-1}cm^{-1}$ as estimated from the amino acid composition (Mach, 1998). The CD data were expressed in terms of mean residue ellipticity, [θ]. Secondary structure estimates were performed using the computer program Selcon (Sreerama and Woody, 1994).

Fluorescence spectra were recorded using a Perkin Elmer LS 50 luminescence spectrometer using a 1.0 mm cell at ~ 24° C. The excitation wavelength was set at 295 nm so as to monitor tryptophan emission between 320-350 nm. The protein concentration used was ~100 μg/ml. In experiments involving ANS (Sigma, St. Louis, USA), fluorescence measurements were made on a 50 μM ANS solution in 20 mM sodium phosphate, pH 7.4, 50 mM NaCl in the absence and presence of ~80 μg/ml Hsp47. A 1 cm cell was used. The samples were excited at 350 nm. The dye was equilibrated with the protein for 5 min prior to data collection.

2.5.3 pH Titrations

These experiments were normally started with Hsp47 in the pH 7.4 phosphate buffer. The pH was then progressively decreased by adding a few μl of 1 M HCl at a time. The solution was mixed thoroughly and the pH was measured using a Corning pH meter 125 with a Corning extra-slender electrode.

2.5.4 Turbidity Measurements

Acid-soluble calf skin collagen (type I) was purchased from Worthington Biochemical Corporation (Freehold, USA) as a 6 mg/ml liquid preparation in 75 mM sodium citrate, pH 3.6 - 4.0, containing 0.01% merthiolate. This solution was brought to a pH of 7.0 and to a final protein concentration of 0.2 mg/ml by the addition of 20 mM sodium phosphate buffer, 50 mM NaCl, pH 7.4. The protocol used for following collagen self-assembly was essentially similar to that described (Williams et al., 1978). Turbidity measurements at 313 nm were made using a Perkin Elmer Lambda 6 UV/VIS spectrophotometer over a 120 min period at a constant temperature of 34° C maintained by circulating water from a Lauda thermostated water bath. A 1-cm quartz cuvette was used. In experiments on the effect of Hsp47 on collagen aggregation, the collagen diluting buffer contained the appropriate amounts of Hsp47. All additions were made with thorough mixing. In experiments at pH 6.0, the buffer containing Hsp47 was first adjusted to pH 6.3 with subsequent addition of collagen solution to give a final pH of 6.0.

Fibril formation with the protease-treated collagen was performed in the same manner as described above except that the final collagen concentration was 1 μM. Pronase and pepsin treated collagen fragments were obtained using the recently described protocols (Kuznetsova and Leikin, 1999). Briefly, acid-soluble collagen was dialyzed against 0.1 M calcium acetate, pH 7.2. Pronase (Calbiochem, La Jolla, USA) was added to the collagen at a ratio of 1:100 and mixed at room temperature. After 20 h an equal volume of 0.5 M acetic acid was added to stop the reaction. Pepsin was added directly to

the acid soluble type I collagen at a ratio of 10 mg pepsin/100 mg collagen and mixed end over end for 48 h at 4°C. After 24 h an additional 10 mg of pepsin was added. 1 M NaCl (final concentration) was mixed with the digested collagen for salt precipitation, thus separating it from the enzymes. SDS-PAGE was used to identify the protease products (Kuznetsova and Leikin, 1999).

2.6 RESULTS

2.6.1 Purification of Recombinant Hsp47

The structural work described below required the availability of pure Hsp47 in sufficient amounts. This was accomplished by devising suitable protocols to overexpress Hsp47 as a recombinant protein in *E. coli* (Figure 2.1, *lane 3*). Growing the bacterial cells at 14° C overnight (instead of 37° C) helped minimize improper folding and deposition into inclusion bodies. The use of the hydroxylapatite column helped eliminate the degradation products that co-eluted with Hsp47 after cleavage from the intein element (Figure 2.1, *lanes 4 and 5*). These protocols yielded an adequate supply of the pure protein. However, Hsp47 has a tendency to aggregate in solution with time. In our hands, the protein also was structurally unstable after 24-36 h, as evidenced by changes in its CD spectral features corresponding to reduction in the α-helical content as well as the onset of aggregation. This necessitated the immediate use of the recombinant Hsp47 in the spectroscopic and fibril formation experiments described below.

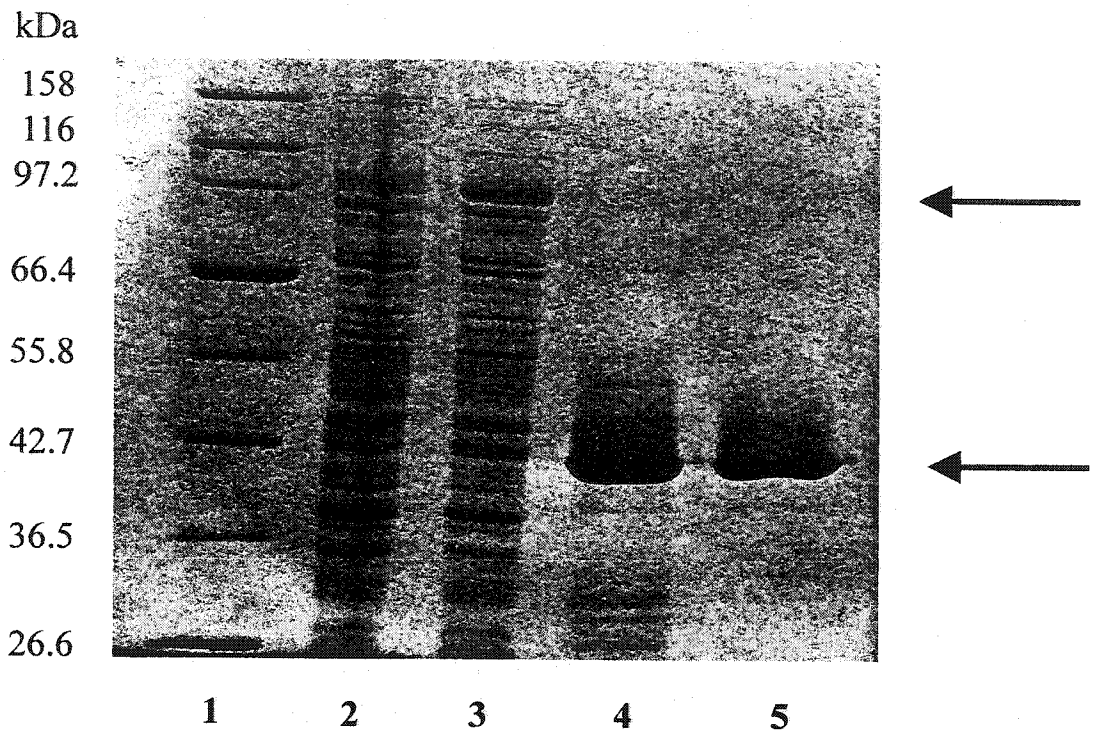


Figure 2.1 Purification of recombinant rat Hsp47 using the IMPACT T7 system. SDS-PAGE of Hsp47 purification run under reducing conditions and stained with Coomassie Brilliant Blue. Lane 1, protein marker, broad range (NEB #7701); Lane 2, uninduced cells; Lane 3, cells induced with IPTG (1 mM); Lane 4, semi-pure Hsp47 eluted from a chitin column; Lane 5, pure Hsp47 eluted from an hydroxylapatite column. The upper arrow indicates the position of Hsp47 plus intein; the lower arrow indicates that of Hsp47 alone.

2.6.2 Inhibition of Collagen Association

The association of collagen molecules *in vitro* into fibrillar structures has been studied by many workers (Holmes et al., 1996; Kadler et al., 1996; Kuznetsova and Leikin, 1999; Notbohm et al., 1999; Prockop and Fertala, 1998; Suarez et al., 1980; Wallace, 1985; Ward et al., 1986; Williams et al., 1978). This process can be studied either by warming a neutral pH solution of collagen extracted from tissues (Kuznetsova and Leikin, 1999; Notbohm et al., 1999; Suarez et al., 1980; Wallace, 1985; Ward et al., 1986; Williams et al., 1978) or, by using collagen generated *de novo* by enzymatic cleavage of the propeptides (Kadler et al., 1988; Prockop and Fertala, 1998). While the packing arrangement and shape of the resulting fibrils is dependent on the procedure used (Holmes et al., 1996), the fibril formation process in both cases is essentially similar. It is an entropy-driven self-assembly process involving a lag phase, a propagation phase and an equilibrium state between fibrils and monomers (Kadler et al., 1988; Wallace, 1985). In order to examine if Hsp47 may be involved in modulating this process, we studied the aggregation of type I collagen at 34°C by monitoring its turbidity as reflected in absorbance increase at 313 nm (Kuznetsova and Leikin, 1999; Notbohm et al., 1999; Suarez et al., 1980; Wallace, 1985; Ward et al., 1986; Williams et al., 1978).

A plot of turbidity versus time of collagen solution in a neutral pH buffer showed a lag of approximately 10 minutes and a subsequent increase in turbidity which leveled off after 40 minutes (Figure 2.2, upper panel). Addition of the buffer containing a 2 molar excess of Hsp47 to the collagen solution at a final pH of 7.0 caused the absorbance of the

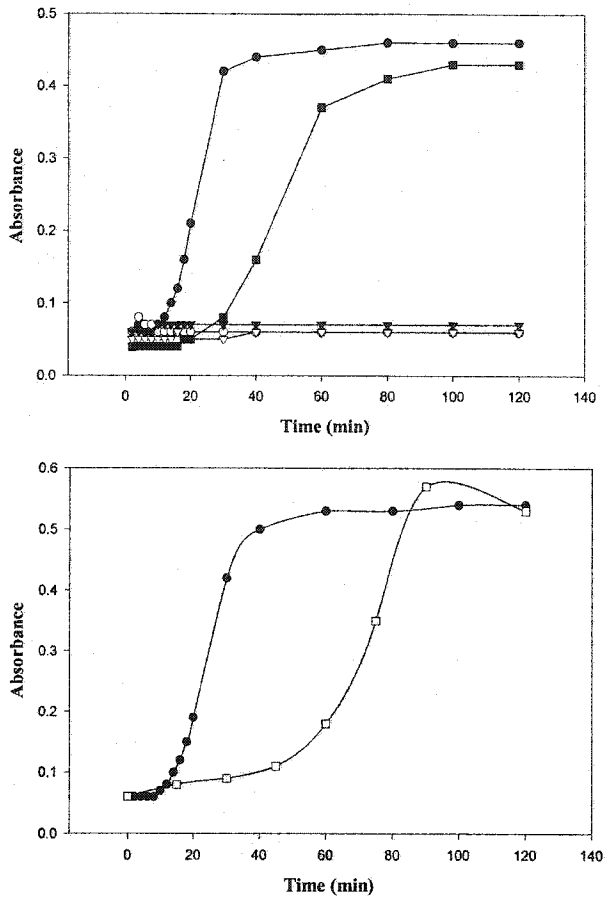


Figure 2.2 Effect of Hsp47 on collagen fibril formation. Collagen association into fibrils was monitored at 313 nm in the presence or absence of Hsp47. Upper panel: Acid-soluble type I collagen was diluted in 20 mM sodium phosphate buffer, pH 7.4, 50 mM NaCl to a final concentration of 200 μg/ml (0.6 μM). Molar ratio of Hsp47 to collagen triple helix: (●) 0:1; (■) 0.5:1, (▽) 1:1, (▼) 2:1 and (o) 6:1. Lower panel: Acid-soluble type I collagen was diluted in 20 mM sodium phosphate buffer, pH 6.3, 50 mM NaCl to a final concentration of 200 μg/ml (0.6μM). Molar ratio of Hsp47 to collagen triple helix: (●) 0:1 and (□) 3.5:1.

latter to remain at the basal value for 2 h with no onset of turbidity (Fig. 2.2, upper panel). As a control, the addition of 2 molar excess of bovine serum albumin to collagen at pH 7.0 did not cause an inhibition of collagen association (data not shown), thus suggesting that the inhibition of collagen fibril formation is specific to Hsp47.

Experiments were performed to study the concentration-, pH- and time-dependence of the ability of Hsp47 to prevent collagen fibril formation. The mole ratio of Hsp47 to collagen triple helix required to prevent collagen association was determined by turbidity measurements taken at a constant collagen concentration of 0.6 μM and varying amounts of Hsp47. As shown in Fig. 2.2, upper panel, a minimum of one Hsp47 to one collagen triple helix was found necessary to effectively inhibit collagen association at pH 7.0. Studies of the pH dependence of fibril formation showed that at pH 6.0, the time course of the turbidity increase for collagen alone in the absence of Hsp47 was comparable to that obtained at pH 7.0 (Fig. 2.2, lower panel). However, unlike the results at pH 7.0, fibril formation did occur at pH 6.0 even in the presence of a 3.5 molar excess of Hsp47, albeit after a markedly increased lag period (Fig. 2.2, lower panel). To find out whether, at pH 7.0, Hsp47 affects collagen fibril assembly during the lag and/or the propagation phase, Hsp47 was added to the collagen solution at a 2:1 mole ratio at selected time points. As shown in Fig. 2.3, when Hsp47 was added after 5 minutes, the absorbance of the collagen solution remained at basal levels. However, when Hsp47 was added after 10 minutes, fibrils continued to form with the absorbance reaching about half of the value obtained in the absence of Hsp47. The above results show that Hsp47

effectively inhibits fibril formation at neutral pH when added during the lag phase with at least one mole of Hsp47 per mole of collagen triple helix.

2.6.3 Role of Telopeptides

To examine whether Hsp47 exerts its effects on collagen fibre formation by binding to the telopeptide domains of the collagen monomer, we performed turbidity measurements on collagen which had been digested by proteases. Removal of the extra helical regions of collagen by proteases has been shown to alter its fibre assembly kinetics, not its ability to form normal fibres (Kuznetsova and Leikin, 1999). In our experiments collagen was treated with pronase which removes both the N- and C-telopeptides. Higher collagen concentrations (1 μM) were used to accelerate fibril formation in the absence of the telopeptides, with the ratio of Hsp47 to collagen remaining at 2:1. Figure 2.4 illustrates that the pronase-treated collagen is capable of forming fibrils at pH 7. However, in the presence of Hsp47 it is no longer able to do so. Similar results were found using pepsin-treated collagen (data not shown). Therefore, we concluded that because Hsp47 is still able to prevent fibril formation in the absence of the telopeptides it must bind within the triple helical region of collagen and not to the telopeptides themselves.

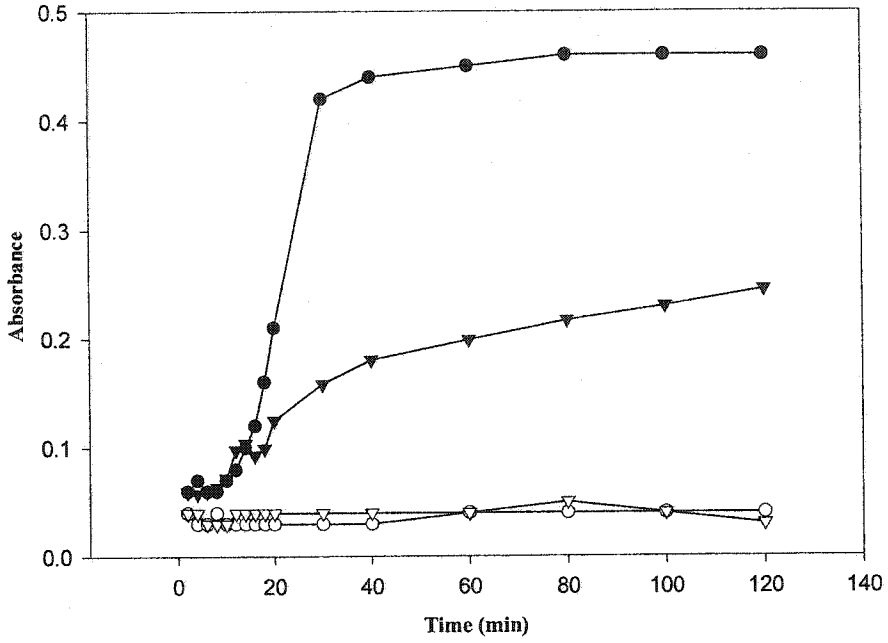


Figure 2.3 Effect of addition of Hsp47 after the start of collagen fibril formation. Fibril assays were done at pH 7.0 as in Figure 2.2. Hsp47 was either not added (●) or added in a 2:1 ratio of Hsp47 to collagen triple helix at (○) 0 min; (▽) 5 min and (▼) 10 min after initiation of fibril formation.

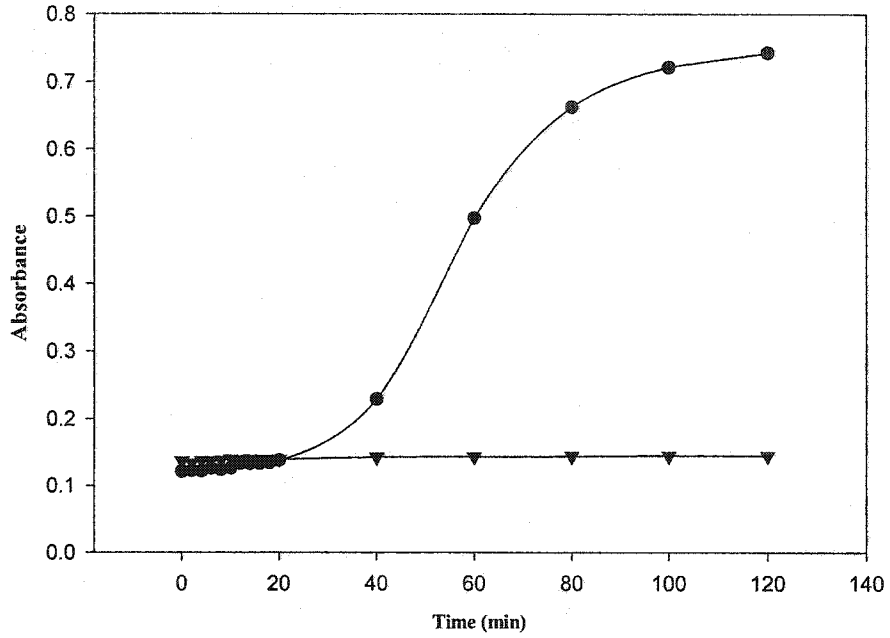


Figure 2.4 Effect of pronase treatment on the ability of Hsp47 to prevent collagen fibril formation. Pronase treatment, as described in Materials and Methods, was used to remove the C- and N-terminal telopeptides from collagen, type I. Fibril formation assays were performed with the pronase treated collagen as in Figure 2.2 but with a final collagen concentration of 370 μg/ml (1 μM). Molar ratio of Hsp47 to collagen triple helix was 0:1 (●) or 2:1 (▼).

2.6.4 pH-Induced Conformational Changes

Hsp47 binding to collagen is known to be ATP-independent (Nakai et al., 1992) and pH-dependent (Thomson and Ananthanarayanan, 1999). The latter finding correlates with the observation that bound Hsp47 can be eluted from a collagen affinity column by lowering the pH from neutral to 6.3 (Saga et al., 1987). It is highly likely that this pH-dependence of Hsp47 binding to collagen is responsible for the effect of pH on the inhibition of collagen fibril formation by Hsp47. To know the effect of pH on Hsp47 structure, we resorted to CD and fluorescence spectroscopic methods.

The CD data are shown in Figure 2.5. At pH 7.4, the spectrum of Hsp47 displayed well-defined minima near 222 and 208 nm which are indicative of the presence of the α-helical structure in Hsp47 at this pH where it binds to collagen. Titration down to a pH of 6.3 resulted in a decrease in the magnitudes of the ellipticities at 222 and 208 nm indicating a decrease in α-helical content. Further reduction of the pH caused a progressive loss of the α-helix until at pH 4.2, there was no further change in the CD spectrum of Hsp47 (Figure 2.5). An estimate of the relative amounts of secondary structures in Hsp47 at the different pHs showed that, on going from pH 7.4 to pH 4.2, the α-helical content decreased from about 28% to about 12% while the amount of β-structure increased from about 23% to 39%. The changes in the β-turn and random structures were relatively small over this pH range. The changes in the α-helix and β-structure contents on reducing the pH from 7.4 to 6.2 although modest, are significant and reproducible (Table 2.1).

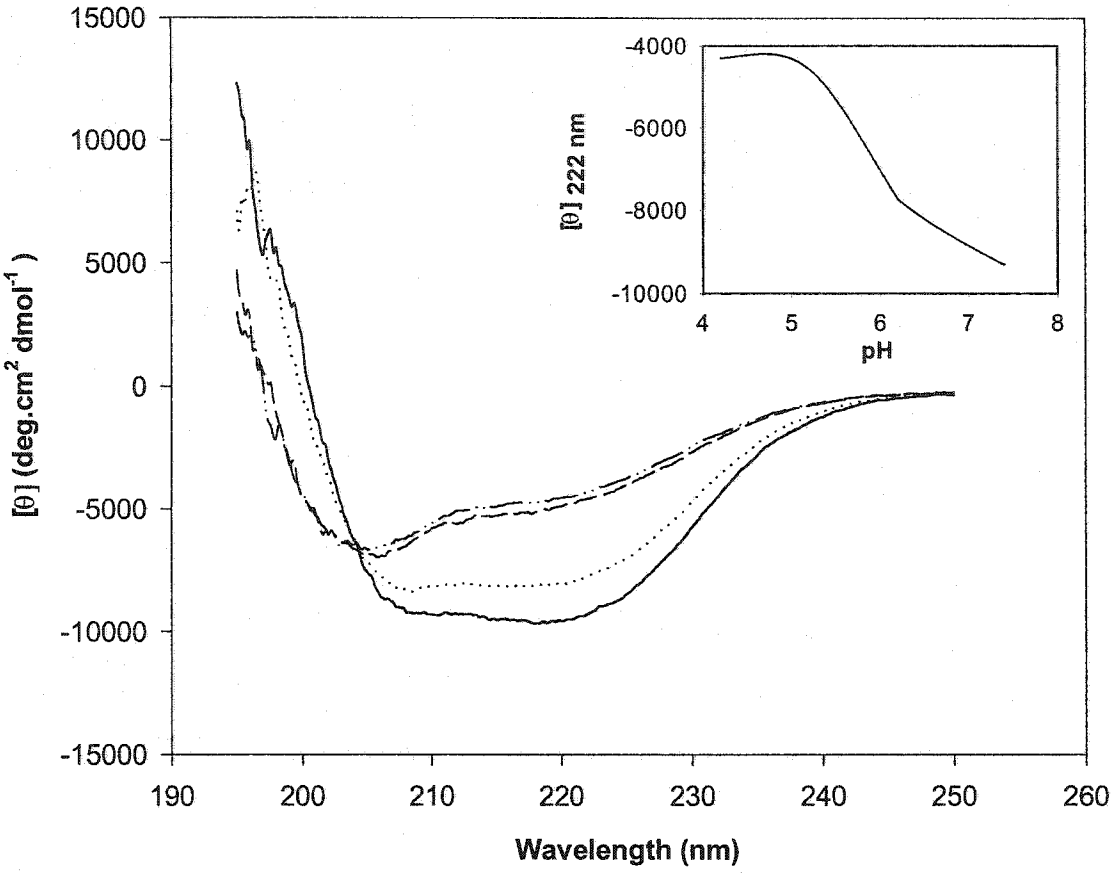


Figure 2.5 Effect of pH on the mean residue ellipticity of Hsp47. CD spectra were recorded of Hsp47 (6.5 μM) as a result of titration from pH 7.4 (—) to pH 6.2 (·····), pH 5.2 (— —) and pH 4.2 (— ··). Inset: Changes in mean residue ellipticity of Hsp47 at 222 nm as a function of pH.

Table 2.1

pH-Induced Changes in Secondary Structure of Hsp47

Contents of secondary structure were estimated with the Selcon program of Sreerama and Woody, 1994. Figures for β-structure include both parallel and anti-parallel β-structure.

	Proportion of structure as:			
pH	α-Helix	β-Structure	β-Turn	Other
7.4	0.28	0.23	0.26	0.24
6.2	0.23	0.26	0.25	0.24
5.2	0.13	0.37	0.25	0.24
4.2	0.12	0.39	0.26	0.24

Further examination of the pH-dependent conformational changes in Hsp47 was made by monitoring changes in the fluorescence emission spectra of tryptophan residues in Hsp47. At pH 7.4, the fluorescence spectrum of Hsp47 showed a broad peak with a maximum around 336 nm (Figure 2.6). It is known that the emission maximum of tryptophan buried in the protein interior lies well below 330 nm while that of solvent-exposed tryptophan is observed around 335 nm and above (Chen, 1969). The emission maximum of 336 nm observed in native Hsp47 would thus suggest the exposure to the solvent of one or more of the five tryptophan residues in the protein at pH 7.4. Decreasing pH to 6.2 from 7.4 caused a decrease in fluorescence intensity and a slight shift to higher wavelength. Titration of the pH to 5.3 and 4.3 caused further decreases in fluorescence intensity and a red shift in the emission maximum (Figure 2.6). These changes correlate well with the unfolding of Hsp47 at lower pHs inferred from the CD data.

El-Thaher et al. (1996) have reported pH-dependent spectral changes in Hsp47 and suggested their relevance to the pH-dependent binding of the protein to collagen. Our spectral data agree well with these authors' data. However, our estimated secondary structural content at different pHs differs significantly from that reported by El-Taher et al. This is very likely due to the different programs used. The one used by El-Taher et al. is known to yield a relatively high β-structure content (Greenfield, 1996). Also, while we found substantial pH-dependent fluorescence changes that correlated well with the CD changes at protein concentrations of 0.1- 0.3 $mg.ml^{-1}$ (Figures 2.5 and 2.6), El-Thaher et

al. 1996, rather intriguingly, found only a relatively minor change in Hsp47 fluorescence at the same protein concentration (1.2 $mg.ml^{-1}$) that was used in their CD experiments. This contradicts the significant tertiary structural changes detected by these authors from the near-uv CD data.

2.6.5 Binding of ANS

1-Anilino-8-naphthalene sulfonate is a fluorescent dye which can be used to assess the surface hydrophobicity of proteins (Cardamone and Puri, 1992). Transfer of the dye from a polar aqueous environment by binding to a non-polar binding site on a protein causes a substantial increase in quantum yield and a large blue shift in emission maximum (Stryer, 1965). To obtain information on the surface hydrophobicity of Hsp47, we examined ANS binding to the protein. As shown in Figure 2.7, addition of Hsp47 at pH 7.4 led to enhancement of the fluorescence intensity of the ANS along with a large blue shift of the initial emission maximum at 519 nm. This suggests that native Hsp47 may have exposed hydrophobic patches on its surface at neutral pH. A similar experiment performed at pH 6 (Figure 2.7) showed a further increase in ANS fluorescence upon interaction with Hsp47 at this pH. This is indicative of increased exposure of hydrophobic groups at the lower pH and is consistent with the fluorescence data in Figure 2.6.

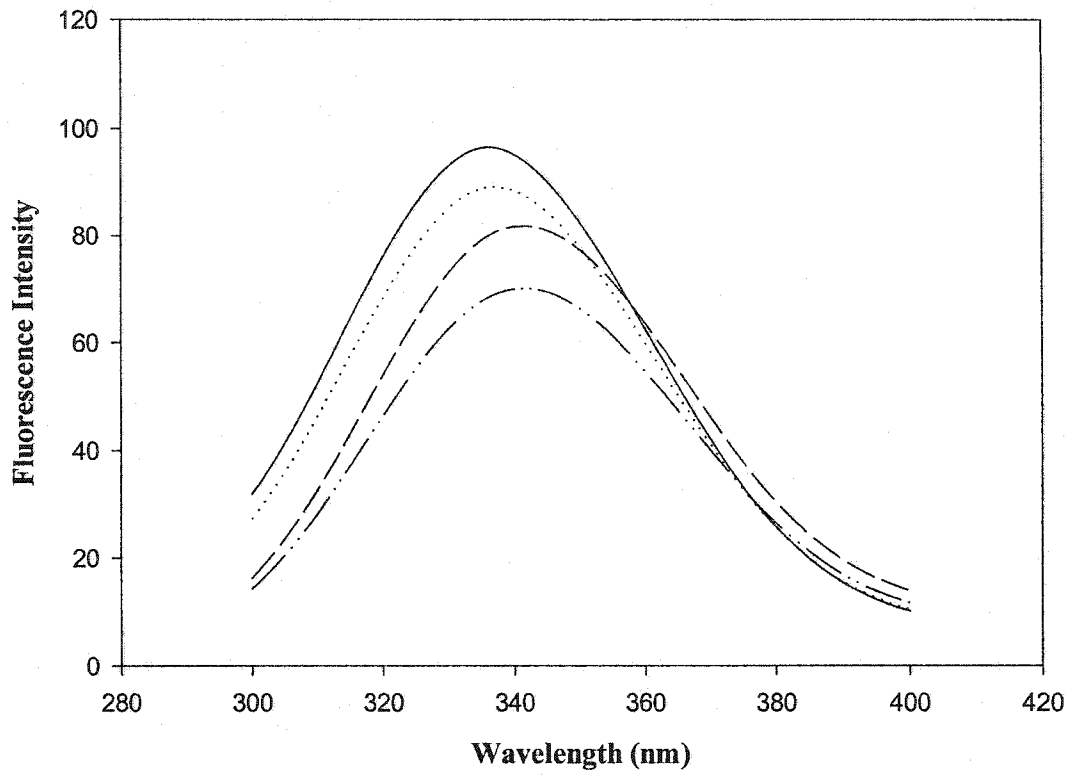


Figure 2.6 Effect of decreasing pH on Hsp47 tryptophan fluorescence. Fluorescence emission spectra were recorded with an excitation wavelength of 295 nm. Changes in Hsp47 fluorescence are shown as a result of titration from pH 7.4 (——) to pH 6.2 (·····), pH 5.2 (— —) and pH 4.3 (— ··).

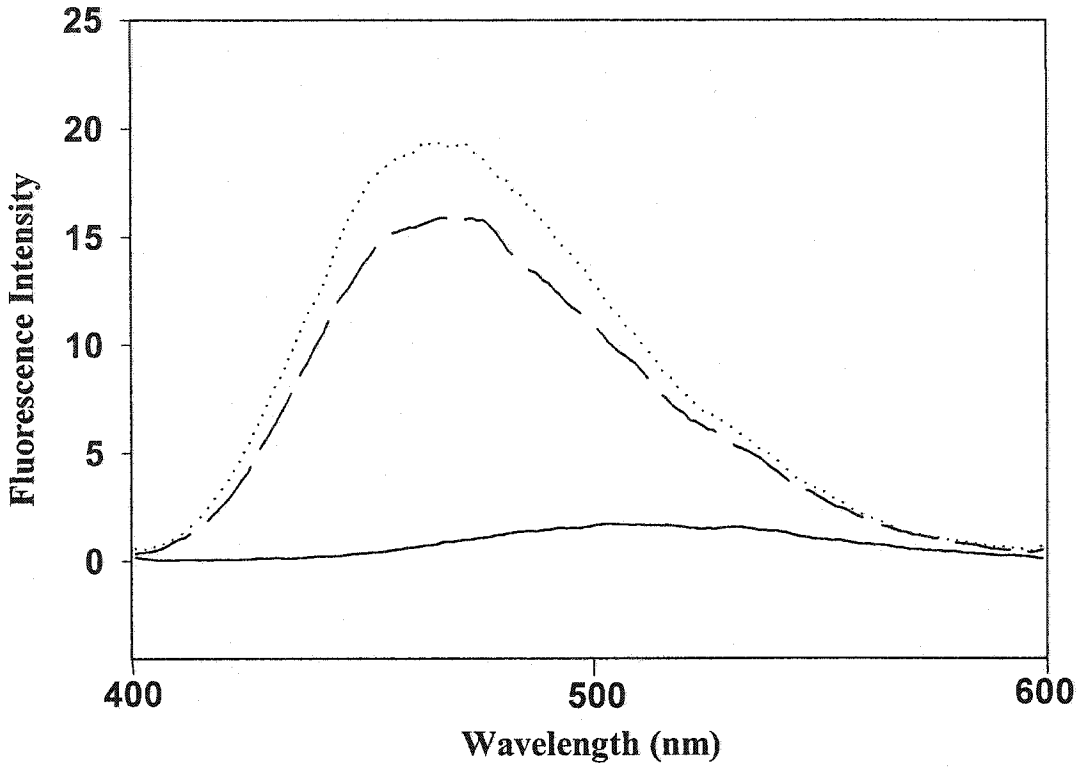


Figure 2.7 Enhancement of ANS fluorescence on binding to Hsp47. Fluorescence emission spectra were recorded with the excitation wavelength set at 350 nm. Shown is the fluorescence of 50 μM ANS (20 mM sodium phosphate buffer, pH 7.4, 50 mM NaCl) alone (—) and in the presence of 1.9 μM Hsp47 at pH 7.4 (— —) and at pH 6.3 (·····).

2.7 DISCUSSION

Studies over the past several years have established that Hsp47 is a collagen-specific molecular chaperone and that it plays an important role in collagen biosynthesis in normal and disease states (Kojima et al., 1998; Masuda et al., 1994; Nagata, 1998; Nagata, 1996; Sunamoto et al., 1998). Unlike the chaperones for globular proteins which bind to the hydrophobic regions of the unfolded proteins and do not interact with their folded conformations (Ellis and Hartl, 1999), Hsp47 binds both to the unfolded and triple helical forms of collagen (Satoh et al., 1996). This is likely due to the fact that the hydrophobic groups in the side chains are exposed in both these forms. Based on observations of chaperone proteins that act on globular proteins (Ellis and Hartl, 1999), earlier workers have proposed the possibility of Hsp47 binding to the single polypeptide chains of nascent procollagen and protecting them from being folded or aggregated until the translation is complete (Nagata, 1998; Nagata, 1996; Satoh et al., 1996). There is, however, little experimental evidence for this suggestion.

The results presented in this paper show that Hsp47 is an effective inhibitor of collagen fibril formation *in vitro* at neutral but not lower pHs. Examination of the fibrils formed in our buffer system at pH 7 in the absence of Hsp47 by atomic force field microscopy, which has earlier been used to study collagen fibrils (Gale et al., 1995), showed the typical fibrillar pattern formed by the packing of the triple-helical collagen monomers (data not shown). This means that in our experiments Hsp47 prevents the association of collagen monomers into fibrils rather than into a nonspecific aggregated

form which could also lead to an increase in the turbidity of the collagen solution. Our CD and fluorescence spectral data show that Hsp47 undergoes significant secondary as well as tertiary structural changes on going from neutral to lower pHs (Figures 2.5 and 2.6). Therefore, the pH dependence of fibril inhibition appears to arise from changes in the structure of Hsp47 rather than in the structure of the collagen monomer since the latter is capable of fibril formation at both pHs. Fibril formation by collagen has been regarded as involving hydrophobic interactions as evidenced by the inhibition of the process by hydrophobic alkyl ureas (Suarez et al., 1980). However, dihydroxy alcohols can also affect fibril formation so that additional interactions such as hydrogen bonding have also been suggested (Kuznetsova et al., 1998). Both types of interactions are possible between Hsp47 and the triple helical collagen molecule. On the one hand, Hsp47 seems to have a significant proportion of its hydrophobic groups exposed to the solvent as suggested by the tryptophan fluorescence data in Figure 2.6 and the ANS fluorescence data in Figure 2.7. These could interact with the nonpolar residues in collagen most of which are exposed to the solvent in the triple helical structure. The ANS fluorescence data show a greater extent of exposure of hydrophobic groups at pH 6.3 than at pH 7.4. Since collagen fibril formation inhibition by Hsp47 is absent at the lower pH, hydrophobic interaction alone is apparently not sufficient for Hsp47 action. Additionally, there is certainly an involvement of electrostatic interaction (possibly involving one or more histidine residues) in the binding as seen from the pH-dependence of the binding of Hsp47 to collagen (Saga et al., 1987) and its inhibition of fibril formation (Figure 2.2).

A major outcome of the present study is that it provides for a conformation-dependent *in vitro* functional assay for Hsp47. The structural details of how Hsp47 could inhibit fibril formation are, however, not clear at the moment. In a recent study, Prockop and Fertala (Prockop and Fertala, 1998) have examined the effect of several synthetic peptides on fibril formation of collagen monomers. They found that peptides corresponding to the N- and C-telopeptide regions, which lie outside the central triple helical part of collagen, inhibited collagen fibril formation if added during the initial lag phase but not later. (It should be noted that the inhibitory action of Hsp47 at pH 7.4 occurs at low (10^{-7} M), equimolar concentration of the protein unlike the much higher (10^{-3} M) concentrations needed for the synthetic peptides.) Prockop and Fertala interpreted their data to suggest that collagen fibrillar assembly occurs via specific binding sites close to the N- and C-termini of the monomers. Because a similar behavior is exhibited by Hsp47 at pH 7.4, we hypothesized that Hsp47 may be binding collagen at either the N- or C-terminus and thereby accomplish its inhibitory action on fibril formation. To test this hypothesis we examined Hsp47's ability to prevent fiber formation using collagen in which the telopeptides had been removed (Figure 2.4). Surprisingly, Hsp47 was still able to prevent fiber formation in the absence of the N- and C-termini. We can therefore conclude that Hsp47 may bind within the triple-helical portion of the collagen molecule and prevent fibril formation. Experiments are currently underway in our laboratory to map the exact site/sites of Hsp47 binding to collagen.

Recently studies have shown that the pH of the ER is near neutral (Kim et al., 1998) and that the average pH of the Golgi apparatus is 6.4 (Kim et al., 1996). It is known that Hsp47 binds nascent procollagen in the ER and releases it as triple-helical procollagen in the cis-Golgi network (Satoh et al., 1996). It is tempting to suggest that this may be correlated with the pH-dependence of the structure of Hsp47 as well as its collagen binding and fibril formation inhibition. Admittedly, more structure-function experiments are necessary to fully unravel the collagen chaperone role(s) of Hsp47. Further details regarding how Hsp47 acts to prevent collagen fibril formation are likely to provide insight into the initial stages of fibrillogenesis, *in vivo*.

2.8 ACKNOWLEDGEMENTS

This work was supported, in part, by a grant from the Canadian Heart Foundation. We thank Professor B. Sanwal of the University of Western Ontario for providing the cDNA clone of and monoclonal antibodies to rat gp46. We also thank Professor M.C. Goh and Matthew Paige of the University of Toronto for the examination of collagen fibrils using atomic force field microscopy and Professor N. Greenfield for her help with the secondary structure analysis of the CD data.

2.9 REFERENCES

Cardamone, M., and Puri, N. K. (1992). Spectrofluorimetric assessment of the surface hydrophobicity of proteins. Biochem J *282*, 589-93.

Cates, G. A., Litchfield, D. W., Narindrasorasak, S., Nandan, D., Ball, E. H., and Sanwal, B. D. (1987). Phosphorylation of a gelatin-binding protein from L6 myoblasts by protein kinase C. FEBS Lett *218*, 195-9.

Chen, R. F., Edelhoch, H., and Steiner, R. F. (1969). Fluorescence of proteins. In Physical Principles and Techniques of Protein Chemistry, S. Leach, ed. (New York, NY: Academic Press), pp. 217-231.

Ellis, R. J., and Hartl, F. U. (1999). Principles of protein folding in the cellular environment. Curr Opin Struct Biol *9*, 102-10.

El-Thaher, T., Drake, A., Yokota, S., Nakai, A., Nagata, K., and Miller, A. (1996). The pH-dependent , ATP-independent interaction of collagen specific serpin/stress protein Hsp47. Protein Peptide Lett. *3*, 1-8.

Gale, M., Pollanen, M. S., Markiewicz, P., and Goh, M. C. (1995). Sequential assembly of collagen revealed by atomic force microscopy. Biophys J *68*, 2124-8.

Greenfield, N. J. (1996). Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. Anal Biochem *235*, 1-10.

Hirayoshi, K., Kudo, H., Takechi, H., Nakai, A., Iwamatsu, A., Yamada, K. M., and Nagata, K. (1991). HSP47: a tissue-specific, transformation-sensitive, collagen-binding heat shock protein of chicken embryo fibroblasts. Mol Cell Biol *11*, 4036-44.

Holmes, D. F., Watson, R. B., Chapman, J. A., and Kadler, K. E. (1996). Enzymic control of collagen fibril shape. J Mol Biol *261*, 93-7.

Kadler, K. E., Hojima, Y., and Prockop, D. J. (1988). Assembly of type I collagen fibrils de novo. Between 37 and 41 degrees C the process is limited by micro-unfolding of monomers. J Biol Chem *263*, 10517-23.

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

Kim, J. H., Johannes, L., Goud, B., Antony, C., Lingwood, C. A., Daneman, R., and Grinstein, S. (1998). Noninvasive measurement of the pH of the endoplasmic reticulum at rest and during calcium release. Proc Natl Acad Sci U S A *95*, 2997-3002.

Kim, J. H., Lingwood, C. A., Williams, D. B., Furuya, W., Manolson, M. F., and Grinstein, S. (1996). Dynamic measurement of the pH of the Golgi complex in living cells using retrograde transport of the verotoxin receptor. J Cell Biol *134*, 1387-99.

Kojima, T., Miyaishi, O., Saga, S., Ishiguro, N., Tsutsui, Y., and Iwata, H. (1998). The retention of abnormal type I procollagen and correlated expression of HSP 47 in fibroblasts from a patient with lethal osteogenesis imperfecta. J Pathol *184*, 212-8.

Kurkinen, M., Taylor, A., Garrels, J. I., and Hogan, B. L. (1984). Cell surface-associated proteins which bind native type IV collagen or gelatin. J Biol Chem *259*, 5915-22.

Kuznetsova, N., Chi, S. L., and Leikin, S. (1998). Sugars and polyols inhibit fibrillogenesis of type I collagen by disrupting hydrogen-bonded water bridges between the helices. Biochemistry *37*, 11888-95.

Kuznetsova, N., and Leikin, S. (1999). Does the triple helical domain of type I collagen encode molecular recognition and fiber assembly while telopeptides serve as catalytic domains? Effect of proteolytic cleavage on fibrillogenesis and on collagen-collagen interaction in fibers. J Biol Chem *274*, 36083-8.

Mach, H., Middaugh, C., Denslow, N. (1998). Determining the identity and purity of recombinant proteins by UV absorption spectroscopy. In Current Protocols in Protein Science, J. Coligan, Dunn, B., Ploegh, H., Speicher, D. and Wingfield, P., ed. (New York: John Wiley and Sons), pp. 7.2.8.

Masuda, H., Fukumoto, M., Hirayoshi, K., and Nagata, K. (1994). Coexpression of the collagen-binding stress protein HSP47 gene and the alpha 1(I) and alpha 1(III) collagen genes in carbon tetrachloride- induced rat liver fibrosis. J Clin Invest *94*, 2481-8.

Nagata, K. (1998). Expression and function of heat shock protein 47: a collagen-specific molecular chaperone in the endoplasmic reticulum. Matrix Biol *16*, 379-86.

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

Nagata, K., and Yamada, K. M. (1986). Phosphorylation and transformation sensitivity of a major collagen- binding protein of fibroblasts. J Biol Chem *261*, 7531-6.

Nakai, A., Hirayoshi, K., and Nagata, K. (1990). Transformation of BALB/3T3 cells by simian virus 40 causes a decreased synthesis of a collagen-binding heat-shock protein (hsp47). J Biol Chem *265*, 992-9.

Nakai, A., Satoh, M., Hirayoshi, K., and Nagata, K. (1992). Involvement of the stress protein HSP47 in procollagen processing in the endoplasmic reticulum. J Cell Biol *117*, 903-14.

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

Prockop, D. J., and Fertala, A. (1998). Inhibition of the self-assembly of collagen I into fibrils with synthetic peptides. Demonstration that assembly is driven by specific binding sites on the monomers. J Biol Chem *273*, 15598-604.

Prockop, D. J., and Kivirikko, K. I. (1995). Collagens: molecular biology, diseases, and potentials for therapy. Annu Rev Biochem *64*, 403-34.

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

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

Sauk, J. J., Smith, T., Norris, K., and Ferreira, L. (1994). Hsp47 and the translation-translocation machinery cooperate in the production of alpha 1(I) chains of type I procollagen. J Biol Chem *269*, 3941-6.

Smith, T., Ferreira, L. R., Hebert, C., Norris, K., and Sauk, J. J. (1995). Hsp47 and cyclophilin B traverse the endoplasmic reticulum with procollagen into pre-Golgi intermediate vesicles. A role for Hsp47 and cyclophilin B in the export of procollagen from the endoplasmic reticulum. J Biol Chem *270*, 18323-8.

Sreerama, N., and Woody, R. W. (1994). Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods. J Mol Biol *242*, 497-507.

Stryer, L. (1965). The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe of non-polar binding sites. J Mol Biol *13*, 482-95.

Suarez, G., Veliz, M., and Nagel, R. L. (1980). Role of hydrophobic interactions in collagen fibril formation: effect of alkylureas in vitro. Arch Biochem Biophys *205*, 422-7.

Sunamoto, M., Kuze, K., Tsuji, H., Ohishi, N., Yagi, K., Nagata, K., Kita, T., and Doi, T. (1998). Antisense oligonucleotides against collagen-binding stress protein HSP47 suppress collagen accumulation in experimental glomerulonephritis. Lab Invest *78*, 967-72.

Thomson, C., and Ananthanarayanan, V (1999). Evidence for a role of Hsp47 preventing premature collagen aggregation. FASEB J. (Abstracts) *12*, A1399.

Wallace, D. (1985). The role of hydrophobic bonding in collagen fibril formation: a quantitative model. Biopolymers *24*, 1705-20.

Ward, N. P., Hulmes, D. J., and Chapman, J. A. (1986). Collagen self-assembly in vitro: electron microscopy of initial aggregates formed during the lag phase. J Mol Biol *190*, 107-12.

Williams, B. R., Gelman, R. A., Poppke, D. C., and Piez, K. A. (1978). Collagen fibril formation. Optimal in vitro conditions and preliminary kinetic results. J Biol Chem *253*, 6578-85.

CHAPTER 3: A METHOD FOR EXPRESSION AND PURIFICATION OF SOLUBLE, ACTIVE HSP47, A COLLAGEN-SPECIFIC MOLECULAR CHAPERONE

Christy A. Thomson and Vettai S. Ananthanarayanan (2001) *Protein Expression and Purification*, **23**, 8-13

3.1 PREFACE

The limited number of *in vitro* studies characterizing Hsp47 can be related to difficulties in its purification. Subsequent to the publication of our *Biochemical Journal* paper on Hsp47, we have received numerous requests for the vector containg Hsp47. Therefore, the main goal of the work outlined in this chapter was to provide a detailed description of the cloning, expression and purification of recombinant Hsp47 using the IMPACT^{TM} T7 system from New England Biolabs. Use of this system to purify Hsp47 provided us the opportunity to radio-label Hsp47 with ^{35}S-cysteine. The radio-labelled Hsp47 was then used to demonstrate Hsp47 binding to the individual proα chains of type I collagen and to map Hsp47 binding sites in collagen (Chapter 4).

During the purification of Hsp47 it was noted that there existed a trimeric form of the protein. This trimeric species was shown to exist *in vivo* and disruption of the trimer was shown to have a negative effect on Hsp47 inhibition of fibrillogenesis. This was the first evidence that oligomerization of Hsp47 might have an effect on its collagen-binding activity.

I performed all of the experimental work presented in this publication during the course of my Ph.D. research. The isolated low density microsomes used in the Western analysis, were kindly provided by Scott Covey, McMaster University.

3.2 PAPER

A Method for Expression and Purification of Soluble, Active Hsp47, a Collagen-Specific Molecular Chaperone

Christy A. Thomson and Vettai S. Ananthanarayanan[1]

Department of Biochemistry, McMaster University, Hamilton, Ontario, Canada L8N 3Z5

[1]To whom correspondence should be addressed. (Tel: 905-525-9140, x-22783; Fax: 905-522-9033; E-mail: ananth@mcmaster.ca)

[2]Abbreviations used: Hsp47, 47-kDa heat-shock protein; LDM, low-density microsomes; PBS, phosphate-buffered saline; TBS, tris-buffered saline; TTBS, tris-buffered saline with Tween-20; PMSF, phenylmethylsulfonyl fluoride; IPTG, isopropyl-α-D-thiogalactopyranoside; LB medium, Luria-Bertani medium; DTT, dithiothreitol; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; CD, circular dichroism.

3.3 ABSTRACT

Hsp47 is regarded as a collagen-specific chaperone with several suggested roles in collagen biosynthesis under normal and disease conditions. We describe here a procedure for the expression and purification of Hsp47 in *E.coli* using the IMPACT^{TM} expression system (New England Biolabs) where the guest gene is fused to the adduct, intein, with a chitin-binding domain. Use of this system resulted in relatively high levels of soluble Hsp47 compared to other available protocols, especially when the bacterial cells were induced at 14° C instead of 37° C. The cell lysate was passed through a chitin-Sepharose affinity column and Hsp47 was cleaved from intein using β-mercaptoethanol. Minor degradation products were subsequently removed using a hydroxylapatite column to yield milligram amounts of pure and active protein suitable for structural studies. Gel electrophoretic analysis of the purified protein indicated the presence of a small proportion of trimeric species when non-reducing conditions were used. The ability to form a trimer may be important for its role as a chaperone. The IMPACT^{TM} system allows for radio-labelling of purified Hsp47 with ^{35}S for use in binding experiments. Illustrative data on collagen binding by ^{35}S-Hsp47 are shown.

3.4 INTRODUCTION

Molecular chaperones recognize protein polypeptide chains at various stages of their biosynthesis and help their proper folding, sorting, assembly and disassembly (Gething and Sambrook, 1992). Hsp47, a 47-kDa heat-shock protein, is a recent addition to the list of proteins that act as molecular chaperones during the biosynthesis of collagen, the most abundant protein in vertebrates. The synthesis of Hsp47 is seen to parallel that of collagen in many tissues (Nagata, 1996). This, and the association of Hsp47 with procollagen at different stages of its biosynthesis, point to its functioning as a collagen-specific chaperone (Nagata, 1998; Nagata, 1996). Hsp47 from chick embryo fibroblasts shows a high degree of sequence identity to human and rat gp46 and to J6 from mouse teratocarcinoma cells (Clarke and Sanwal, 1992). The term Hsp47 is now used to represent this group of collagen-binding proteins, the differences being accounted for by species variation (Nagai et al., 1999).

With an ever-increasing number of reports on the clinical (Kojima et al., 1998; Rocnik, 2000) and biochemical (Nagata, 1998; Nagata, 1996) aspects of the protein, there has been much interest in understanding the structure and function of Hsp47. Available data point to the involvement of Hsp47 at several stages of collagen biosynthesis. However, the details of the modulation of collagen folding and assembly by Hsp47 and the nature of its interactions with a variety of substrates including procollagen, collagen and gelatin, are still unclear (Nagata, 1998; Nagata, 1996). In particular, there is no direct experimental evidence for any of the several suggested roles of this protein in collagen

synthesis. We have undertaken a study of the role of Hsp47 as a collagen-specific molecular chaperone from a structural perspective. This, naturally, requires an adequate supply (several mg) of purified Hsp47 in a soluble and active form. Our attempts at obtaining pure recombinant Hsp47 from *E. coli* using published protocols that use either gelatin (Saga et al., 1987) or antibody affinity (Jain et al., 1994) chromatography, met with problems of low yields, contamination by a higher-molecular weight (~70 kDa) protein, or aggregation.

As an alternative to the above methods, we describe here the cloning of recombinant rat *hsp47* into the IMPACT^{TM} expression system (offered by New England Biolabs), which yielded high levels of soluble Hsp47 particularly when cells were induced at 14° C instead of 37° C. This system also allows for radio-labelling of the purified protein with ^{35}S for use in binding experiments. The activity of the purified protein preparations was assessed by its inhibition of the *in vitro* association of type I collagen into fibrils (Thomson and Ananthanarayanan, 2000). The method for isolating pure and active Hsp47 described here is convenient as it yields sufficient amounts of very pure, soluble and active protein with a relatively mild and simple purification protocol.

3.5 MATERIALS AND METHODS

The cDNA for rat *gp46* and monoclonal antibody were a gift from Professor Sanwal of the University of Western Ontario in London, Ontario, Canada. Enzymes, protein markers (broad range NEB #7701) and reagents for cloning were obtained from

New England Biolabs (Mississauga, Ontario). pCYB3, pTYB3 and pTYB4 plasmids were also from New England Biolabs. Hydroxylapatite resin was bought from Bio-Rad (Mississauga, Ontario). ^{35}S-L-cysteine was purchased from NEN Life Science Products (Boston, MA). Nitrocelluose membrane (Hybond ECL) and PD-10 desalting column were from Amersham Pharmacia Biotech (Quebec, Canada). X-ray imaging film (X-OMAT AR) and intensifying screen (BioMax TranScreen-LE) were purchased from Kodak Laboratories (Rochester, New York). Protease inhibitor cocktail (Complete Mini) was obtained from Roche Molecular Biochemicals (Laval, Quebec). Other chemicals and reagents were from Sigma Chemical Co. (Missouri, USA) and GIBCO BRL (Burlington, Ontario)

PCR amplification of the *hsp47* insert was performed in a Techne PHC-3 Thermal cycler from Mandel (Guelph, Canada). Nucleotide sequencing was done at the MOBIX Central facility at McMaster University.

Isolation of the low density microsomal (LDM) fraction of 3T3-F442A cells was performed using a modified protocol (Lange and Brandt, 1990). Briefly, 3T3-F442A cells were grown 2 days past confluence in Dulbecco's Modified Eagle Media (GIBCO BRL) with 10 % calf serum, 2 mM L-glutamine, 50 U/ml penicillin and streptomycin. Cells were washed in PBS, collected in a wash buffer (1 mM Hepes, 1 mM EDTA, 0.25 M sucrose, pH 7.4) and pelleted at 16000g for 10 min. The cells were then resuspended in wash buffer, homogenized in the presence of 50 µg/ml DNase, RNase and 0.1 M PMSF and incubated at 37°C for 10 min. The cell lysate was pelleted at 11000g for 20 min and

the supernatant removed. LDM was isolated from the supernatant by centrifugation at 436000g for 20 min and resuspended in PBS with 0.1 M PMSF to a final concentration of 5.8 mg/ml.

For Western analysis, 25 μg of the isolated LDM from 3T3-F442A cells was run in 10 % SDS-PAGE under reducing and non-reducing conditions and transferred to nitrocellulose at 100 V for 1 h. The blot was blocked with 3 % gelatin in TBS buffer (20 mM Tris, pH 7.6, 150 mM NaCl). Anti-gp46 was added in TTBS (TBS + 0.05 % Tween) and incubated for 1.5 h. After washing in TTBS, the blot was probed with goat-anti mouse IgG from Santa Cruz Biotechnology, INC. (Santa Cruz, California) for 45 min. The blot was subsequently washed 3 times in TTBS and once in TBS before performing chemiluminescence with Luminol reagent from Santa Cruz Biotechnology, INC.

CD spectral measurements were made using a Jasco J-600 spectropolarimeter as described earlier (Thomson and Ananthanarayanan, 2000). Data were obtained from 195 to 250 nm and were corrected for buffer contribution. A 2 mm quartz cell was used. Turbidity measurements were performed using a Cary 300 Bio UV-Visible spectrophotometer as previously described (Thomson and Ananthanarayanan, 2000).

3.6 RESULTS AND DISCUSSION

3.6.1 Cloning and Expression of Hsp47 in pTYB4

The IMPACT^{TM} system was chosen for expressing Hsp47 as it allows for an one-step protein purification as well as for subsequent radio-labelling of the target protein

(Evans and Ming-Qun, 2000). This system exploits the facile cleavage of the guest protein from its fusion adduct, intein, in the presence of thiols. Initial attempts to clone the cDNA of *hsp47* into pCYB3 or pTYB3 plasmids were unsuccessful as they either gave low levels of expression or insufficient cleavage of the intein-bound protein, most likely due to steric hindrance from the C-terminal leucine in Hsp47. To circumvent both these problems, the pTYB4 vector was chosen as it has a T7 promoter that allows for high levels of expression and adds a single glycine to the C-terminus of the target protein which allows for sterically facile cleavage of Hsp47 from the intein.

The cloning of *hsp47* into pTYB4 required PCR amplification of mature *hsp47* cDNA from the original expression vector pTrc46K (Jain et al., 1994). A major deterrent to using pTrc46K to purify Hsp47 was the low expression level of the recombinant protein. The PCR primers were designed to place an Nco I site at the N-terminus and a Sma I site at the C-terminus of the *hsp47* cDNA. Once amplified, the PCR product was digested by Nco I and Sma I and ligated into the corresponding sites of pTYB4 (Figure 3.1). The ligation products were then transformed into the competent *E. coli* strain DH5α and screened for ampicillin resistance. Restriction mapping with Nco I and Kpn I yielded the 1.2 kb insert of *hsp47*. A positive clone was verified by sequencing of the full insert and was subsequently transformed into the BL21(DE3) strain of *E. coli*.

To express the recombinant Hsp47, BL21(DE3) cells containing the plasmid were grown overnight at 37°C in LB with 100 mM ampicillin and were used to inoculate a culture at 1% (v/v). At an OD_{600nm} of 0.5-0.8, the cells were induced with 1 mM IPTG

for 3 hours. In contrast to the low expression levels of the pTrc46K vector, we found significant overexpression of the pTYB4/*hsp47* construct (Figure 3.2). Western blot analysis using monoclonal antibody to gp46 verified the expression of Hsp47 (data not shown).

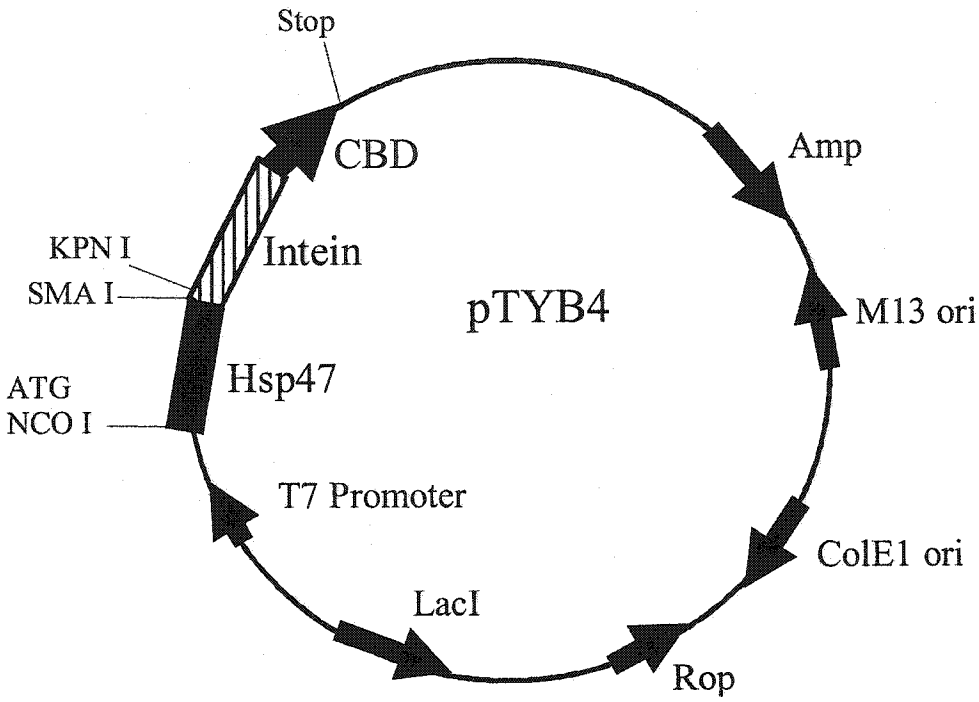


Figure 3.1 pTYB4 expression vector with *hsp47* insert. The cDNA for *hsp47*, minus the signal sequence but including the start methionine, was cloned into the pTYB4 vector at the Nco I and Sma I sites. This yields Hsp47 with a single glycine residue added to its C-terminus following cleavage from the intein.

3.6.2 Purification of Hsp47

One litre cultures of *E. coli* were grown as described above and the cells were collected by centrifugation at 5000g for 10 min. The pellet was resuspended in 20 ml of lysis buffer [0.1 M Tris, pH 7.6, 500 mM NaCl, 0.5 mM EDTA, 0.1% Triton X-100] treated with two protease inhibitor tablets, and lysed using a French press. Cell debris was spun down at 12000g for 10 min and the supernatant and pellet analyzed for Hsp47 by SDS-PAGE. Unfortunately, the majority of the Hsp47 was found in the insoluble fraction, most likely as inclusion bodies. Resolublization of Hsp47 in strong denaturants was unsuccessful in previous trials. Therefore, an attempt to prevent inclusion body formation was made by slowing down the *E. coli* growth rate. One litre flasks of LB/amp were again inoculated at 1% (v/v) and allowed to grow at 37°C with shaking, until an OD_{600nm} of 0.5-0.8 was reached. They were then induced with 1 mM IPTG and placed in a shaker at 14°C overnight. Under these conditions a high proportion of Hsp47 was now found in the soluble fraction (Figure 3.3) so that the purification could proceed. Lowering the temperature thus appeared to have slowed down cell growth which, in turn, allowed more time for the Hsp47 to fold correctly.

Cloning of *hsp47* into pTYB4 created a three-part fusion protein consisting of Hsp47 fused to the N-terminus of the intein and a chitin-binding domain fused to the intein's C-terminus. This allowed for affinity purification of the fusion protein via the chitin-binding domain. To do so, approximately 5-10 ml of chitin-containing affinity beads were equilibrated in the lysis buffer. The cell lysate supernatant was mixed, end

over end, with the beads for 1 h at 4°C and the unbound material was then collected. The beads were poured into a column (1 cm x 10 cm) and washed extensively with 20 mM Tris, pH 7.4, 1.0 M NaCl, 0.5% Triton X-100 followed by a wash with 20 mM sodium phosphate buffer, pH 7.4 containing 50 mM NaCl. The self-cleavage reaction of the intein at its N-terminus was initiated by the addition of 100 mM β-mercaptoethanol to the final wash buffer. The solution was left at 4°C overnight to allow complete cleavage of Hsp47 from the intein. After the cleavage reaction, Hsp47 was collected as an eluant from the chitin column (Figure 3.3) using 20 mM sodium phosphate, pH 7.4 with 50 mM NaCl as running buffer.

Even with the addition of protease inhibitors in the lysis buffer, Hsp47 still underwent some degradation. To remove the degradation products, the eluted Hsp47 was bound to a 2 ml hydroxylapatite column pre-equilibrated in 20 mM sodium phosphate buffer, pH 7.4, 50 mM NaCl. An elution gradient going from 50 mM sodium phosphate, pH 7.4, 50 mM NaCl to 100 mM sodium phosphate, pH 7.4, 50 mM NaCl, was used to separate the full-length Hsp47 from the contaminating proteins (Figure 3.3). This protocol yielded approximately 1-1.5 mg of pure Hsp47 per litre of culture (Table 3.1).

It should be noted that Hsp47 has a tendency to aggregate in solution with time. CD spectral features of the protein changed with time showing that the protein is also structurally unstable after 24-36 hrs (data not shown). This necessitates the immediate use of the recombinant protein for any subsequent structure-function studies. Attempts to stabilize Hsp47 for an extended period of time have, as yet, been unsuccessful.

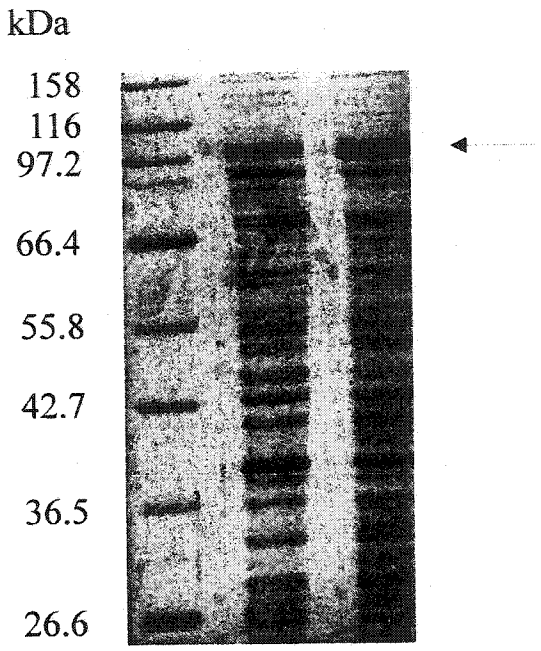


Figure 3.2 Overexpression of Hsp47 in pTYB4 upon induction with IPTG. SDS-PAGE analysis, using 10 % crosslinking, of BL21(DE3) cells run under reducing conditions and stained with Coomassie Brilliant blue. Lane 1, protein markers; Lane 2, uninduced cells; Lane 3, cells induced with IPTG (1 mM). Arrow shows the induced band corresponding to the size of Hsp47 + intein.

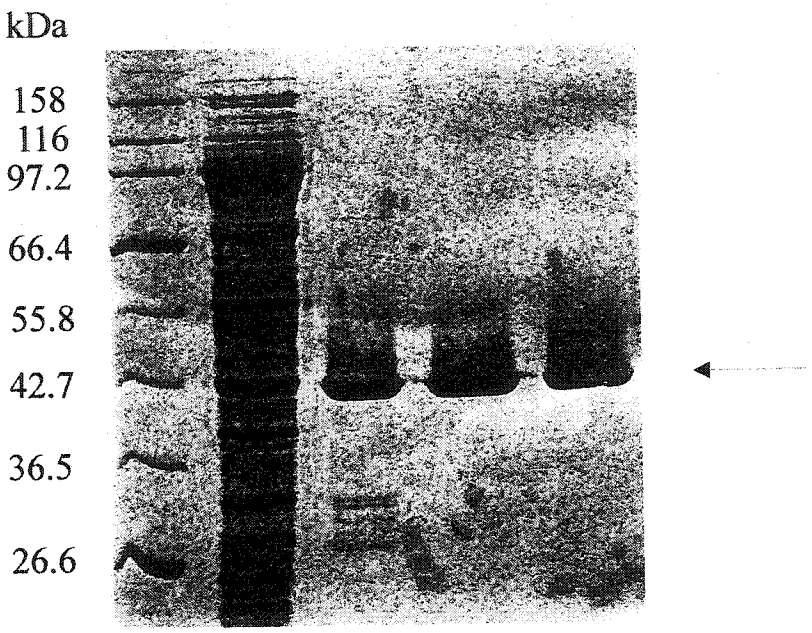


Figure 3.3 Purification of recombinant Hsp47 using the IMPACT T7 system with cells grown at 14°C. 10 % SDS-PAGE analysis of samples under reducing conditions. Staining was done with Coomassie Brilliant blue. Lane 1, protein markers; Lane 2, cell lysate supernatant after removal of cell debris by centrifugation at 12000g; Lane 3, semi-pure Hsp47 eluted from the chitin column with reducing agent; Lanes 4 & 5, purified Hsp47 eluted from hydroxylapatite column. Arrow shows band corresponding to cleaved Hsp47.

TABLE 3.1

Purification of Hsp47 Isolated from 1.5 L of Bacterial Culture

Fraction	Volume (ml)	Protein Concentration (mg/ml)	Total Protein (mg)
Cell lysate	25	26.4	660
Chitin	15	0.16	2.4
Hydroxylapatite	10	0.14	1.4

3.6.3 Trimerization of Hsp47

SDS-PAGE analysis of Hsp47 under reducing conditions showed the recombinant protein running as a monomeric species at 44 kDa (Figure 3.3). However, when the protein was run under non-reducing conditions, a proportion of the Hsp47 shifted to a higher molecular weight species of about 130 kDa (Figure 3.4a). To determine whether or not the trimerization of recombinant Hsp47 is an artifact of the purification procedure, we examined the *in vivo* oligomerization states of Hsp47 using a fibroblast cell line. Western analysis with anti-gp46 was performed on the LDM fraction (which include the endoplasmic reticulum) isolated from 3T3-F442A cells. Figure 3.4b shows that Hsp47 from these cells is monomeric when run in SDS-PAGE under reducing conditions and that a small proportion of the protein again runs as a trimeric species under non-reducing conditions. These results confirm that Hsp47 trimerization is not a result of the *in vitro*

purification procedure, as the trimeric species is also found *in vivo*. The trimerization of Hsp47 cannot be the result of intermolecular disulfide bond formation since the recombinant protein contains only a single cysteine residue. Other causes like hydrophobic association may therefore prevail. Oligomerization of a variety of heat-shock proteins and molecular chaperones has been observed in the past. For instance, the ATPase activity of the molecular chaperone BiP is thought to be the result of conversion of oligomers to active monomers (Blond-Elguindi et al., 1993). The small heat-shock proteins Hsp12.2 and Hsp12.3 both form tetramers (Kokke et al., 1998). Erp29, an ER stress protein, self-associates predominantly into homodimers (Mkrtchian et al., 1998).

It has been previously shown that Hsp47 exists in more than one polymerization state (Jain et al., 1994). However, whether or not the trimerization of Hsp47 has a role in its ability to act as a molecular chaperone for collagen, is as yet unknown. It is interesting to note that the chaperoned collagen molecule also exists as a trimer in its mature form. We have recently shown that Hsp47 is able to prevent the self-association of collagen molecules into fibrils *in vitro* (Thomson and Ananthanarayanan, 2000). We therefore investigated if the presence of reducing agents affected its ability to prevent fibril formation. As shown in Figure 3.4c, in the presence of 10 mM DTT, collagen is able to form fibrils as monitored by turbidity. In the absence of DTT, Hsp47 effectively prevents fibril formation for a 2 hour time period. However, when 10 mM DTT is present its ability to do so is drastically reduced. We therefore hypothesize that Hsp47 may be required to oligomerize to function as a chaperone for collagen.

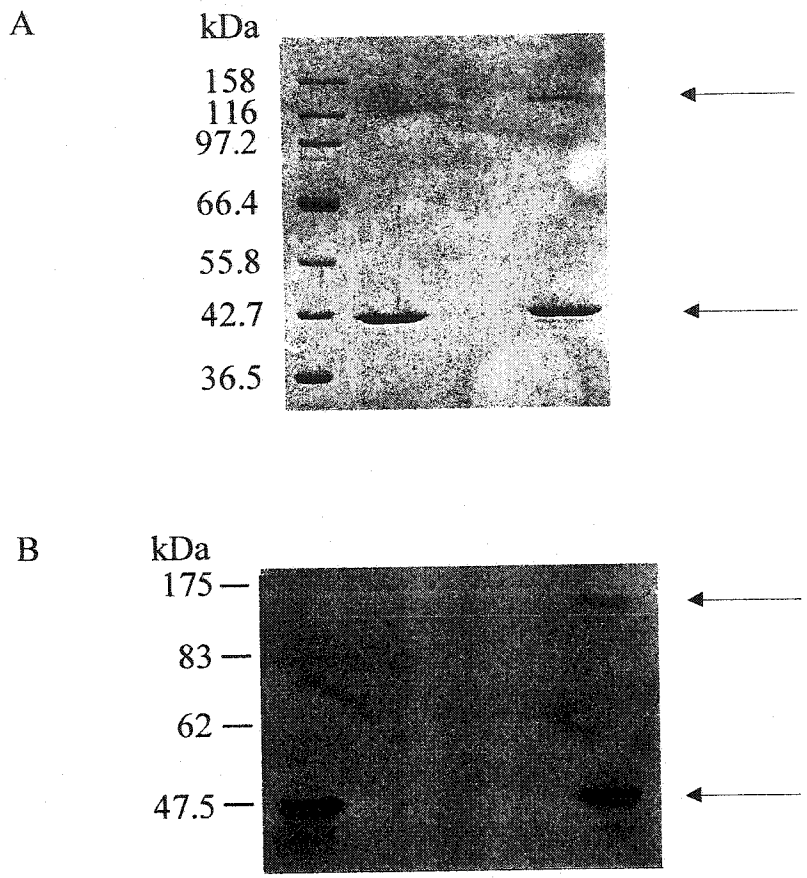


Figure 3.4 Trimerization of recombinant Hsp47 and the effect of reducing conditions on its ability to prevent collagen fibril formation. a) SDS-PAGE analysis of purified recombinant Hsp47 stained with Coomassie brilliant blue. Lane 1, protein markers; Lane 2, Hsp47 run under reducing conditions; Lane 3, empty; Lane 4, Hsp47 run under non-reducing conditions. b) Western analysis of Hsp47 from the LDM fraction of 3T3-F442A cells using anti-gp46 monoclonal antibody. Lane 1, LDM fraction run under reducing conditions; Lane 3 & 4, empty; Lane 5, LDM fraction run under non-reducing conditions. Top and bottom arrows for both a & b correspond to trimerized and monomeric Hsp47, respectively.

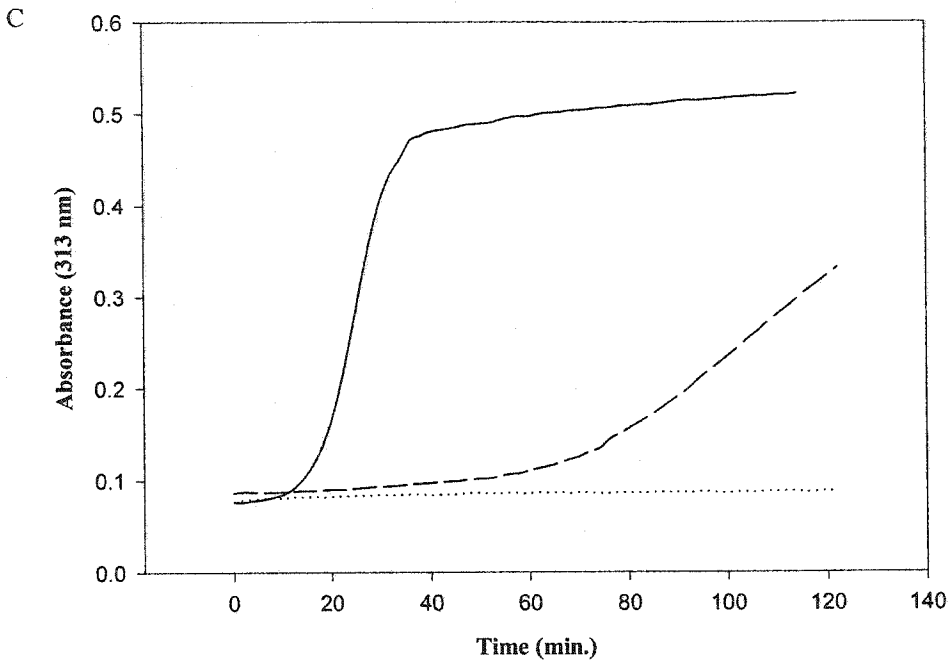


Figure 3.4c Trimerization of recombinant Hsp47 and the effect of reducing conditions on its ability to prevent collagen fibril formation. Solid line, collagen fibril formation under reducing conditions (10 mM DTT); Dotted line, collagen fibril formation under non-reducing conditions and in the presence of Hsp47 (2:1 ratio of Hsp47 to collagen); Dashed line, collagen fibril formation under reducing conditions (10 mM DTT) and in the presence of Hsp47 (2:1 ratio of Hsp47 to collagen).

3.6.4 ^{35}S-Cysteine Labelling of Hsp47

One of the advantages of using the IMPACT^{TM} system is the ease with which the overexpressed protein can be radio-labelled with ^{35}S. This, in turn, can be used in studying the interaction of the target protein with other molecules. We used the IMPACT^{TM} system to covalently label the C-terminus of Hsp47 using L-^{35}S-cysteine. The sulfhydryl group of cysteine can initiate attack at the thioester bond left at the C-terminus of Hsp47 when it is treated with DTT during cleavage from intein. The cysteine forms a stable peptide bond with the C-terminus of the target protein by a spontaneous S-N shift (Evans and Ming-Qun, 2000). For experiments involving ^{35}S-Hsp47, the cleavage reaction was performed using 100 mM DTT rather than β-mercaptoethanol. Upon collection of the purified Hsp47 from the hydroxylapatite column, 1.4 ml of the protein (~150 μg/ml) was mixed with 10 μl of L-^{35}S-cysteine at 4°C for 8 hrs. ^{35}S-Hsp47 was separated from free L-^{35}S-cysteine using a disposable PD-10 desalting column. Figure 3.5a is an autoradiograph showing the successful radio-labelling of ^{35}S-Hsp47 run in 10% SDS-PAGE.

3.6.5 ^{35}S-Hsp47 Binding to Collagen

Labelling of Hsp47 with L-^{35}S-cysteine can be used to assess the activity of the purified protein through binding studies with collagen immobilized on a membrane. To perform such an experiment, acid-soluble calf skin type I collagen, was run under reducing conditions in 7.5% SDS-PAGE and then transferred on to a nitrocellulose

membrane for 4 h at 100 V. The membrane was blocked with 3% bovine serum albumin in TBS (20 mM Tris, pH 7.5, 500 mM NaCl) to prevent non-specific binding of ^{35}S-Hsp47 to the nitrocellulose. The collagen blot was then mixed with ^{35}S-Hsp47 in TTBS (20 mM Tris, pH 7.5, 500 mM NaCl, 0.05% Tween-20) and 1% BSA for 2 h at room temperature to allow binding of Hsp47 to the immobilized collagen. After binding, the ^{35}S-Hsp47 was decanted and the blot washed in TTBS for 5 min followed by TBS for 5 min. The blot was then dried and exposed to x-ray film for autoradiography. The data obtained shows that Hsp47 is capable of binding both the α1 and α2 chains of type I collagen (Figure 3.5b). Previous studies of Hsp47 binding to the α1 and α2 chains of type I collagen by other groups have not been in agreement. Jain et al. (1994) showed that Hsp47 is capable of binding both the proα1(I) and proα2(I) chains, with higher affinity binding to the proα2(I) chain. On the other hand, others have shown that binding only occurs between Hsp47 and the proα1(I) chain (Hu et al., 1995). Our data provides experimental support for Hsp47 binding to both chains of type I collagen.

3.7 ACKNOWLEDGEMENTS

We thank Professor Sanwal and Anne Brickenden, University of Western Ontario, for valuable advice and for providing *gp46* cDNA and monoclonal antibody. We also thank Scott Covey, McMaster University, for supplying the isolated LDM from the 3T3-F442A cells. This work was supported by a grant from the Canadian Heart Foundation.

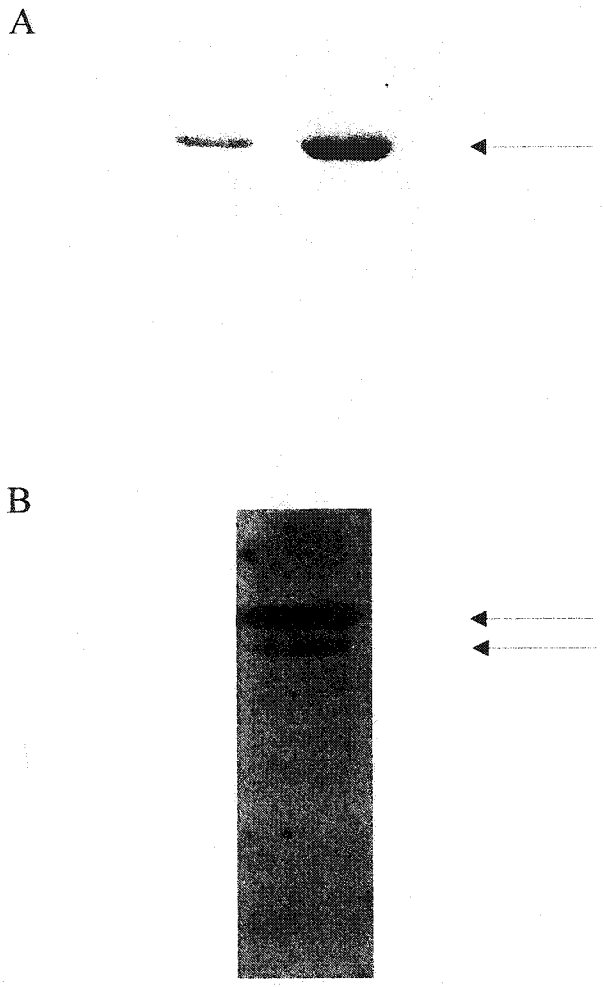


Figure 3.5 ^{35}S-labelling of Hsp47 and its binding to collagen. a.) Hsp47 was labeled with ^{35}S-cysteine as described in the Results section. The labeled protein was subjected to 10% SDS-PAGE and autoradiographed. Lanes 1 & 2, ^{35}S-Hsp47. b) Binding of ^{35}S-Hsp47 to the α1 & α2 chains of collagen. Acid soluble type I collagen was run in 7.5 % SDS-PAGE under reducing conditions, transferred to nitrocellulose and probed with ^{35}S-Hsp47. Top arrow corresponds to the α1 chain of collagen and bottom arrow corresponds to the α2 chain.

3.8 REFERENCES

Blond-Elguindi, S., Fourie, A. M., Sambrook, J. F., and Gething, M. J. (1993). Peptide-dependent stimulation of the ATPase activity of the molecular chaperone BiP is the result of conversion of oligomers to active monomers. J Biol Chem *268*, 12730-5.

Clarke, E. P., and Sanwal, B. D. (1992). Cloning of a human collagen-binding protein, and its homology with rat gp46, chick hsp47 and mouse J6 proteins. Biochim Biophys Acta *1129*, 246-8.

Evans, T., and Ming-Qun, X. (2000). Intein-mediated protein ligation: Harnessing nature's escape artists. Biopolymers (Pep. Sci.) *51*, 355-362.

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

Hu, G., Gura, T., Sabsay, B., Sauk, J., Dixit, S. N., and Veis, A. (1995). Endoplasmic reticulum protein Hsp47 binds specifically to the N- terminal globular domain of the amino-propeptide of the procollagen I alpha 1 (I)-chain. J Cell Biochem *59*, 350-67.

Jain, N., Brickenden, A., Lorimer, I., Ball, E. H., and Sanwal, B. D. (1994). Interaction of procollagen I and other collagens with colligin. Biochem J *304*, 61-8.

Kojima, T., Miyaishi, O., Saga, S., Ishiguro, N., Tsutsui, Y., and Iwata, H. (1998). The retention of abnormal type I procollagen and correlated expression of HSP 47 in fibroblasts from a patient with lethal osteogenesis imperfecta. J Pathol *184*, 212-8.

Kokke, B. P., Leroux, M. R., Candido, E. P., Boelens, W. C., and de Jong, W. W. (1998). Caenorhabditis elegans small heat-shock proteins Hsp12.2 and Hsp12.3 form tetramers and have no chaperone-like activity. FEBS Lett *433*, 228-32.

Lange, K., and Brandt, U. (1990). Insulin-responsive glucose transporters are concentrated in a cell surface-derived membrane fraction of 3T3-L1 adipocytes. FEBS Lett *261*, 459-63.

Mkrtchian, S., Baryshev, M., Matvijenko, O., Sharipo, A., Sandalova, T., Schneider, G., Ingelman-Sundberg, M., and Mkrtchiana, S. (1998). Oligomerization properties of ERp29, an endoplasmic reticulum stress protein. FEBS Lett *431*, 322-6.

Nagai, N., Tetuya, Y., Hosokawa, N., and Nagata, K. (1999). The human genome has only one functional hsp47 gene (CBP2) and a pseudogene (pshsp47). Gene *227*, 241-8.

Nagata, K. (1998). Expression and function of heat shock protein 47: a collagen-specific molecular chaperone in the endoplasmic reticulum. Matrix Biol *16*, 379-86.
Nagata, K. (1996). Hsp47: a collagen-specific molecular chaperone. Trends Biochem Sci *21*, 22-6.

Rocnik, E., Chow, L. and Pickering G. (2000). Heat shock protein 47 is expressed in fibrous regions of human atheroma and is regulated by growth factors and oxidized low-density lipoprotein. Circulation *101*, 1229-1233.

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

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

CHAPTER 4: MAPPING HSP47-BINDING SITE(S) USING CNBr PEPTIDES DERIVED FROM TYPE I AND TYPE II COLLAGEN

Christy A. Thomson, Ruggero Tenni and Vettai S. Ananthanarayanan, submitted to Protein Science

4.1 PREFACE

In our previous studies of Hsp47, we showed that Hsp47 prevented type I collagen fibrillogenesis. However, the site on the collagen molecule where Hsp47 exerted its effects was not known. A scan of the literature revealed discrepancies in describing the specific interactions between Hsp47 and collagen. We therefore designed experiments to map the binding of Hsp47 on the collagen molecule. This was achieved using Hsp47 radio-labelled with ^{35}S-cysteine, as described in Chapter 3, along with CNBr peptides derived from types I and II collagen. The results of this work showed that Hsp47 preferentially binds to discrete regions of the collagen molecule and that the triple helix is the preferred conformation for Hsp47 binding. Further investigation is necessary to narrow down the binding region in terms of its size and specific amino acids, if any, that are required for Hsp47 binding.

I performed all of the experimental work presented in this publication during the course of my Ph.D. research. The CNBr peptides used in the experiments were kindly provided by Professor R. Tenni, University of Pavia, Italy.

4.2 PAPER

Mapping Hsp47-binding Site(s) Using CNBr Peptides Derived from Type I and Type II Collagen*

Running Title: Hsp47-binding Sites on Collagen

Christy A. Thomson[1], Ruggero Tenni[2] and Vettai S. Ananthanarayanan[1,3]

From the [1]Department of Biochemistry, McMaster University, Hamilton, Ontario, Canada, L8N 3Z5 and the [2]Department of Biochemistry, University of Pavia, 27100 Pavia, Italy

[3]To whom correspondence should be addressed: Dept. of Biochemistry, McMaster University, 1200 Main Street West, Hamilton, Ontario, Canada, L8N 3Z5. Tel.: 905-525-9140, ext. 22783; Fax: 905-522-9033; E-mail: ananth@mcmaster.ca.

*The abbreviations used are: ER, endoplasmic reticulum; CD, circular dichroism; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; Hyp or O, 4-hydroxyproline; BSA, bovine serum albumin.

4.3 ABSTRACT

As a crucial molecular chaperone in collagen biosynthesis, Hsp47 interacts with the nascent form as well as the mature triple-helical form of procollagen. The location(s) of Hsp47-binding sites on the collagen molecule are, as yet, unknown. We have examined the substrate specificity of Hsp47 *in vitro* using well-characterized CNBr peptide fragments of type I and type II collagen along with radio-labelled, recombinant Hsp47. Interaction of these peptides with Hsp47 bound to collagen-coated microtiter wells showed several binding sites for Hsp47 along the length of the α1 and α2 chains of type I collagen and the α1 chain of type II collagen, with the N-terminal regions showing the strongest affinities. The latter observation was also supported by the results of a ligand blot assay. Except for two peptides in the α2(I) chain, peptides which showed substantial binding to Hsp47 did so in their triple-helical and not random-coil form. Unlike earlier studies which used peptide models for collagen, the results obtained here on fragments of type I and type II collagen identify, for the first time, binding of Hsp47 to specific regions of the collagen molecule. They also point to additional structural requirements for Hsp47 binding besides the known preference for third-position Arg residues and the triple-helical conformation.

Key Words: Hsp47; molecular chaperone; CNBr peptides of collagen; radio-ligand competition assay

4.4 INTRODUCTION

The endoplasmic reticulum (ER*) is responsible for the folding and export of proteins destined for intracellular organelles and the cell surface. To accomplish this task, the ER contains several folding enzymes and molecular chaperones that assist in the proper folding of proteins and in the retention of misfolded proteins (Helenius et al., 1992). Hsp47 is one such ER-resident heat-shock protein that acts as a chaperone specific to collagen, the most abundant mammalian protein. The expression pattern of Hsp47 is highly correlated with that of collagen both in diseased and normal states (Nagata, 1996). *In situ* experiments have shown that Hsp47 binds to nascent procollagen chains and remains associated with them as they form the triple-helical structure and are exported via the secretary pathway (Sauk et al., 1994; Smith et al., 1995). Once in the cis-Golgi compartment, Hsp47 releases the bound procollagen and returns to the ER via the ER-retention sequence found on its C-terminus (Satoh et al., 1996). The importance of Hsp47 in collagen biosynthesis is illustrated by the retention in the ER of abnormal procollagen molecules produced in the genetic disorder osteogenesis imperfecta, through enhanced expression and co-localization of Hsp47 with procollagen in fibroblasts (Kojima et al., 1998). Furthermore, Hsp47 knock-out mice show embryonic lethality and defects in collagen biosynthesis (Nagai et al., 2000).

The promiscuous binding of Hsp47 to both the unfolded and fully folded procollagen chains makes it difficult to understand the exact stage(s) in collagen biosynthesis where this protein might carry out its chaperone function. In a recent study, we have shown that Hsp47 prevents collagen fibril formation *in vitro* (Thomson and

Ananthanarayanan, 2000). The pH-dependence of this inhibition correlated with pH-induced conformational changes in Hsp47 (Thomson and Ananthanarayanan, 2000). This led us to suggest the possible *in vivo* role for Hsp47 of keeping procollagen from forming fibrils until it reaches the relatively acidic Golgi compartment where Hsp47 would dissociate from procollagen. This does not, however, rule out other possible roles for Hsp47, including its involvement in preventing aggregation of the procollagen chains before they form a triple helix. Conceivably, Hsp47 may have multiple roles similar to the ER chaperone BiP which is involved not only in translocation but also in folding and retention (Gething and Sambrook, 1992).

Studies on the substrate specificity of Hsp47 would be useful in understanding the nature of the Hsp47-collagen interaction and could thereby provide clues to Hsp47 function, including its inhibitory effect on collagen fibril formation. An early study suggested that the N-propeptide region of the proα1(I) chain of collagen may be the predominant domain for binding Hsp47 (Hu et al., 1995). This, however, does not account for the known binding of Hsp47 to processed collagen lacking the propeptide region. More recent studies have employed synthetic peptide models for collagen as well as peptide libraries to delineate Hsp47 binding sites in collagen. Nagata and coworkers found that the triple-helical conformation in the substrate has a much higher affinity for Hsp47 than the random coil form (Koide et al., 1999; Koide et al., 2000; Koide et al., 2002). At the primary structural level, the presence of a Hyp residue was found not to be conducive to Hsp47 binding. Particularly strong binding to Hsp47 was displayed by the Xaa-Arg-Gly sequence embedded within a $(Pro\text{-}Pro\text{-}Gly)_n$ peptide. In contrast, Sauk et

al. have found Arg to be among amino acid residues that disfavour Hsp47 binding (Sauk et al., 2000). These authors also found certain short, non-natural peptides to be very effective in binding Hsp47 (Hebert et al., 2001). So far, no data are available on Hsp47 binding to peptides containing naturally occurring collagen sequences.

In this study, we have examined the Hsp47-binding site(s) on collagen using well-characterized peptide fragments generated by CNBr treatment of type I and type II collagen. In conjunction with sensitive binding assays that employ radio-labelled recombinant Hsp47, we were able to identify regions towards the N-terminus in the triple-helical portion of collagen as the dominant sites of interaction with Hsp47.

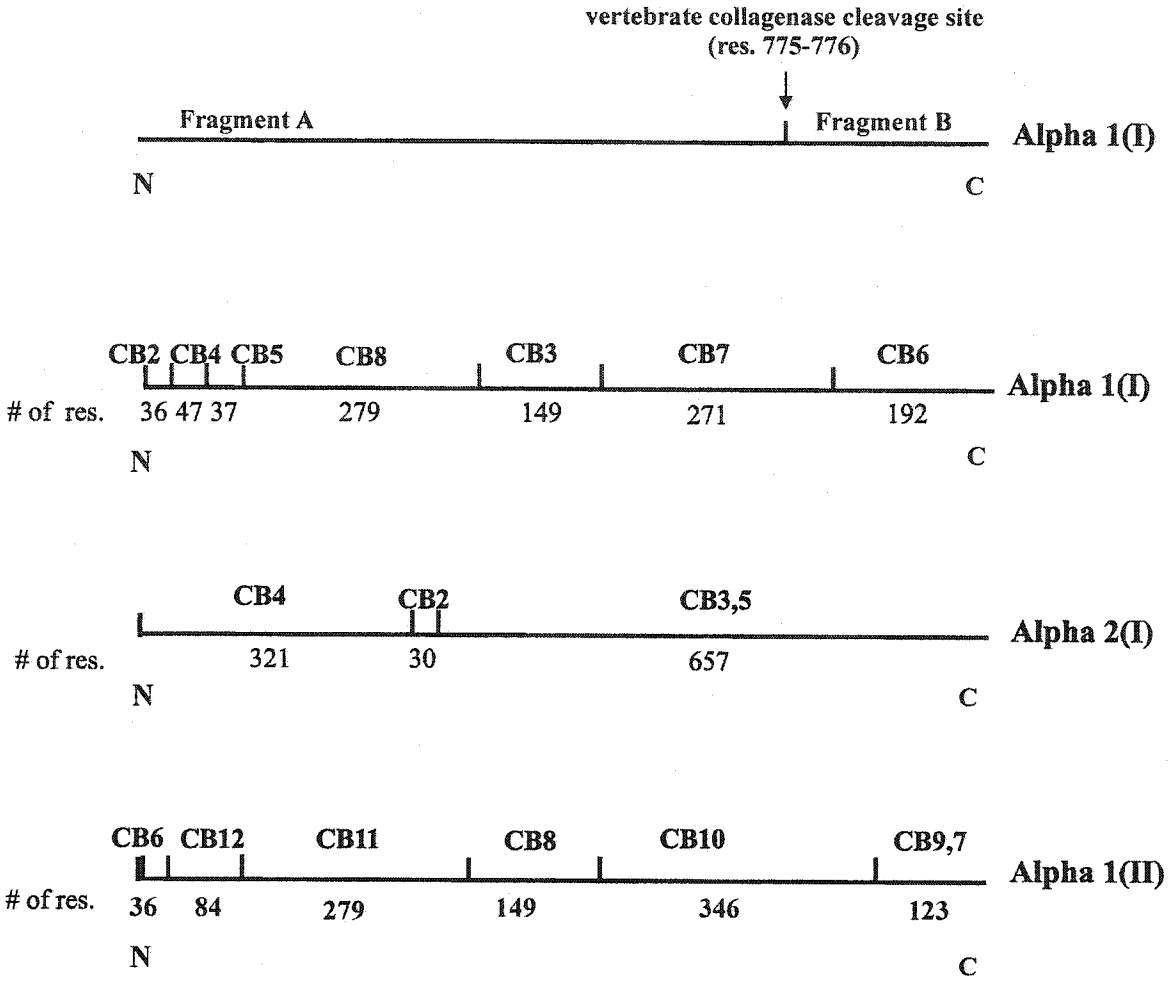


Figure 4.1 Schematic diagram of the α1 and α2 chains of type I collagen and the α1 chain of type II collagen. The nomenclature and positions of the CNBr fragments are indicated. Also shown are the fragments generated by the cleavage of type I collagen by vertebrate collagenase.

4.5 RESULTS

4.5.1 Binding of Hsp47 to Collagenase Cleavage Fragments

Using radio-labelled Hsp47, we have previously shown that this protein binds to both the α1 and the α2 chains of type I collagen immobilized on nitrocellulose (Thomson and Ananthanarayanan, 2001). To map the regions of collagen involved in Hsp47 binding, we first investigated the binding of Hsp47 to peptide fragments generated by digestion of type I collagen by vertebrate collagenase. The enzyme cleaves both the α1 and α2 chains at a site corresponding to Gly^{775}-Xaa^{776} (Gross et al., 1974) to generate an N-terminal fragment A and a C-terminal fragment B of each chain (Figure 4.1). Incubation of ^{35}S-Hsp47 with these fragments immobilized on nitrocellulose membrane after transfer from SDS-PAGE gels, revealed that Hsp47 was bound only to the A fragment of both the α1 and α2 chains of type I collagen (Figure 4.2).

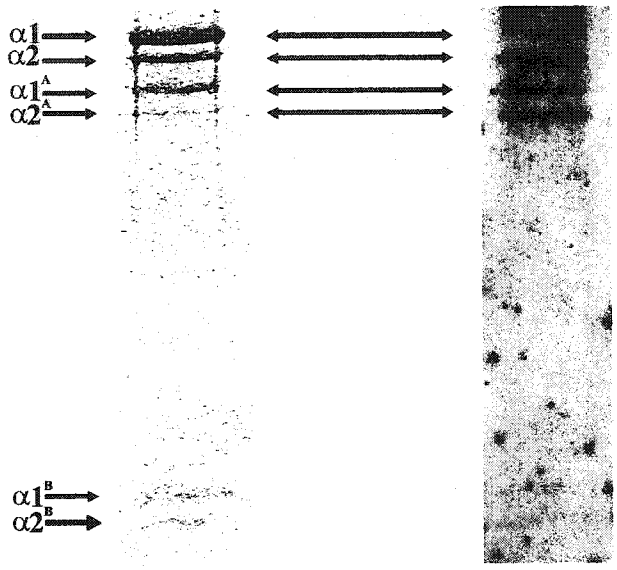


Figure 4.2 Binding of Hsp47 to the α1 and α2 chains and to collagenase fragments of type I collagen. The collagenase-digested collagen fragments were run in SDS-PAGE and were either stained with Coomassie Brilliant Blue (*left panel*) or transferred to nitrocellulose membrane and probed with ^{35}S-Hsp47 and autoradiographed (*right panel*).

4.5.2 Use of the Competition Assay

To further map the binding site(s) of Hsp47 on collagen, we developed an assay that involved the interaction of ^{35}S-Hsp47 with collagen-coated plastic microtiter wells. This assay was used to examine the ability of synthetic peptides and CNBr fragments of collagen to compete for Hsp47 binding to the wells. As illustrated in Figure 4.3, incubation of collagen-coated microtiter wells with ^{35}S-Hsp47 produced a saturation-binding curve with respect to Hsp47 concentration. However, it soon became apparent that non-specific binding of ^{35}S-Hsp47 to the polystyrene wells was also occurring alongside. This problem was circumvented by measuring ^{35}S-Hsp47 binding in the presence and absence of gelatin which when added in excess, effectively binds to and displaces Hsp47 bound to the microtiter wells. Any remaining ^{35}S radioactivity after gelatin treatment would then represent the non-specific binding of Hsp47 to the walls of the microtiter wells. Gelatin was chosen as the competing ligand, since Hsp47 has been shown to bind to gelatin with higher affinity than to collagen types I, II or IV (Jain et al., 1994) and since it is soluble in our buffer system. Figure 4.3 shows the total, nonspecific and specific binding of ^{35}S-Hsp47 to the collagen microtiter wells. Since the focus of this work is to map the binding sites for Hsp47 in the collagen molecule, we did not determine the K_d values of the peptides used in this study. Instead, in presenting and analyzing the binding data, we have made a semi-quantitative comparison of the strengths of binding of the peptides to a fixed amount of Hsp47 using gelatin as the reference compound.

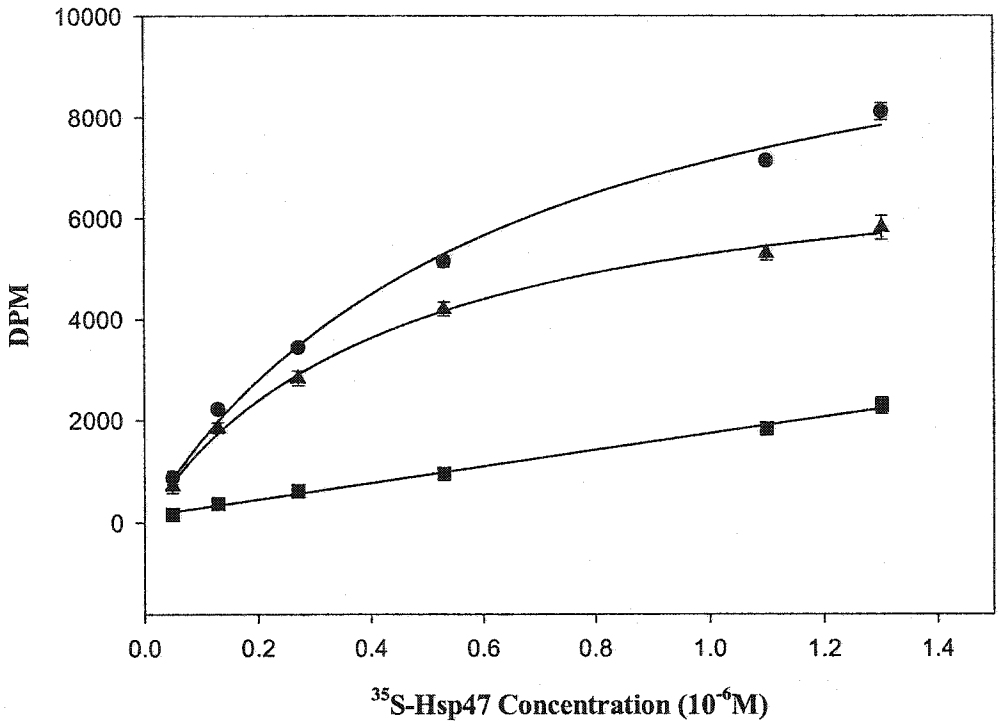


Figure 4.3 Binding of Hsp47 to collagen-coated microtiter wells. Curve 1 (*circles*) Binding to wells in the absence of gelatin; Curve 2 (*squares*) Binding in the presence of 0.1 % gelatin; Curve 3 (*triangles*) Specific binding calculated from subtracting the data in curve 2 from curve 1. Data shown are the mean (±S.E.) of triplicate measurements.

4.5.3 Binding of Hsp47 to Synthetic Collagen Model Peptides

It has been previously shown that Hsp47 binds to synthetic peptides that mimic the collagen triple helix. Koide et al. (1999) showed that Hsp47 binds to triple-helical $(PPG)_{10}$ but not to triple-helical $(POG)_{10}$, whereas others have shown that Hsp47 binding by both these peptides is relatively very weak compared to collagen types I-III and therefore the data do not warrant a comparison between the two peptides (Macdonald and Bachinger, 2001). Our results with these peptides in their triple-helical conformation (as judged from their CD spectra) are shown in Figure 4.4. Both $(POG)_{10}$ and $(PPG)_{10}$ caused only a slight decrease in ^{35}S-Hsp47 binding to the type I collagen-coated microtiter wells, with $(PPG)_{10}$ showing relatively more binding affinity, albeit still weak.

4.5.4 Binding of Hsp47 to CNBr Peptides of Type I Collagen

The α1 and α2 chains of collagen can be cleaved with CNBr to generate discrete peptides representing various regions of the collagen molecule based on the distribution of Met residues. The properties of these peptides have been investigated in earlier studies by one of the authors (Consonni et al., 2000; Rossi et al., 1996; Tenni et al., 2002; Zanaboni et al., 2000). Many, but not all, form soluble, triple-helical homotrimers in solution at room temperature as assessed from CD data. Their random coil forms may be generated by heat denaturation as described in Experimental Procedures. The interactions of Hsp47 with the CNBr peptides in both these conformations were investigated using our radio-ligand competition assay system.

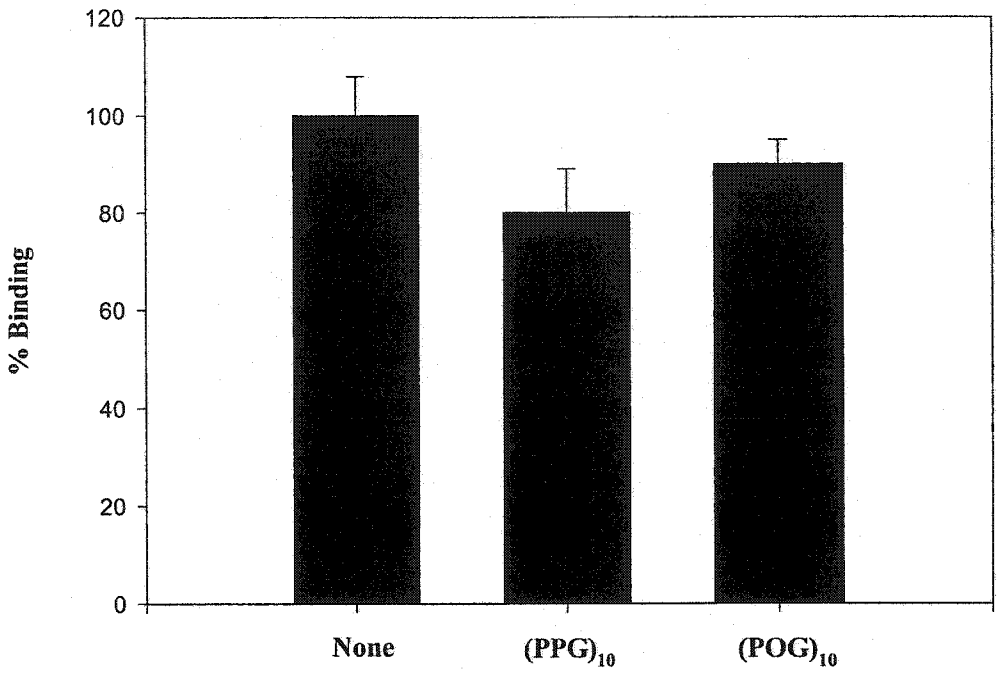


Figure 4.4 Inhibition of Hsp47 binding to collagen-coated microtiter wells by collagen model peptides. Binding of ^{35}S-Hsp47 (0.3 μM) to wells in the presence and absence of the synthetic collagen peptides $(PPG)_{10}$ and $(POG)_{10}$, at 240 μM. Data shown are the mean (±S.E.) of triplicate measurements.

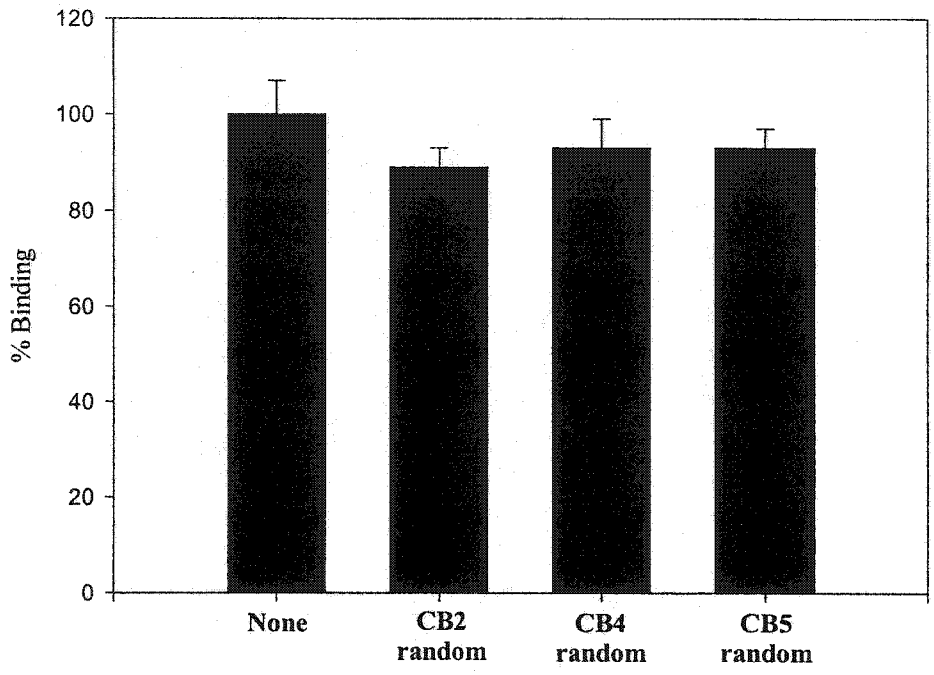


Figure 4.5a Inhibition of Hsp47 binding to collagen microtiter wells by CNBr fragments of type I collagen. Competition of ^{35}S-Hsp47 (0.1 μM) binding to collagen wells by CNBr peptides (6 μM) generated from the N-terminal 120 residues of the α1 chain of type I collagen. The peptides were in a random coil conformation as determined by CD. Data shown are the mean (±S.E.) of triplicate measurements.

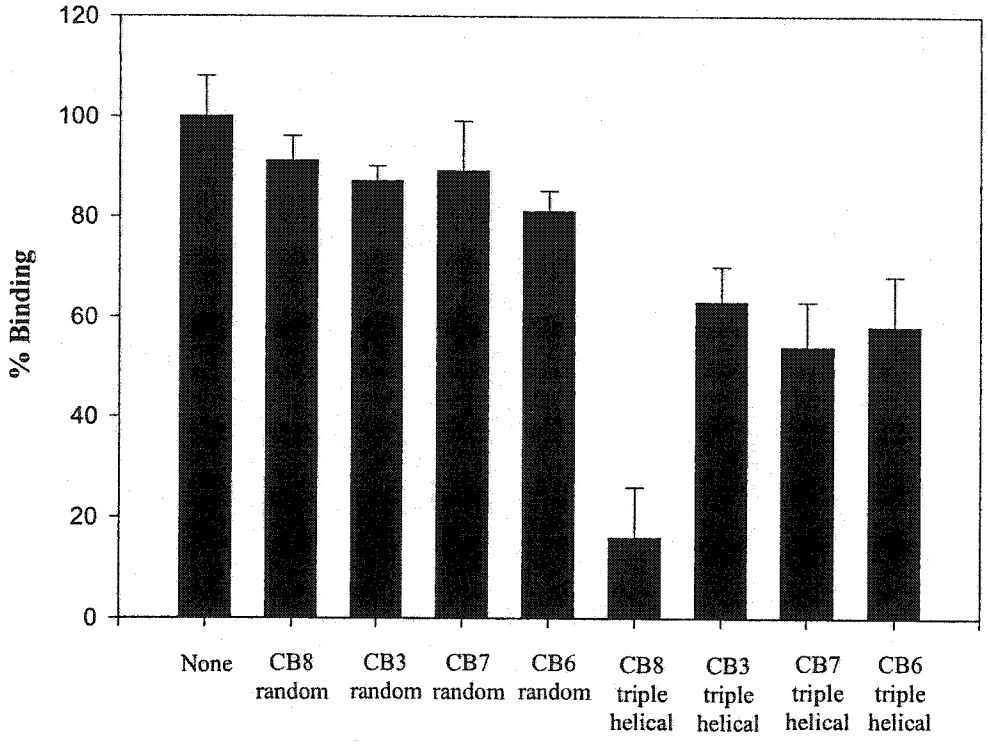


Figure 4.5b Inhibition of Hsp47 binding to collagen microtiter wells by CNBr fragments of type I collagen. Competition of ^{35}S-Hsp47 (0.6 μM) binding to collagen wells by CNBr peptides (12 μM) generated from the rest of the α1 chain of type I collagen. The triple helical conformation of the peptides was determined by CD and the random conformation was generated as described in Materials and Methods. Data shown are the mean (±S.E.) of triplicate measurements.

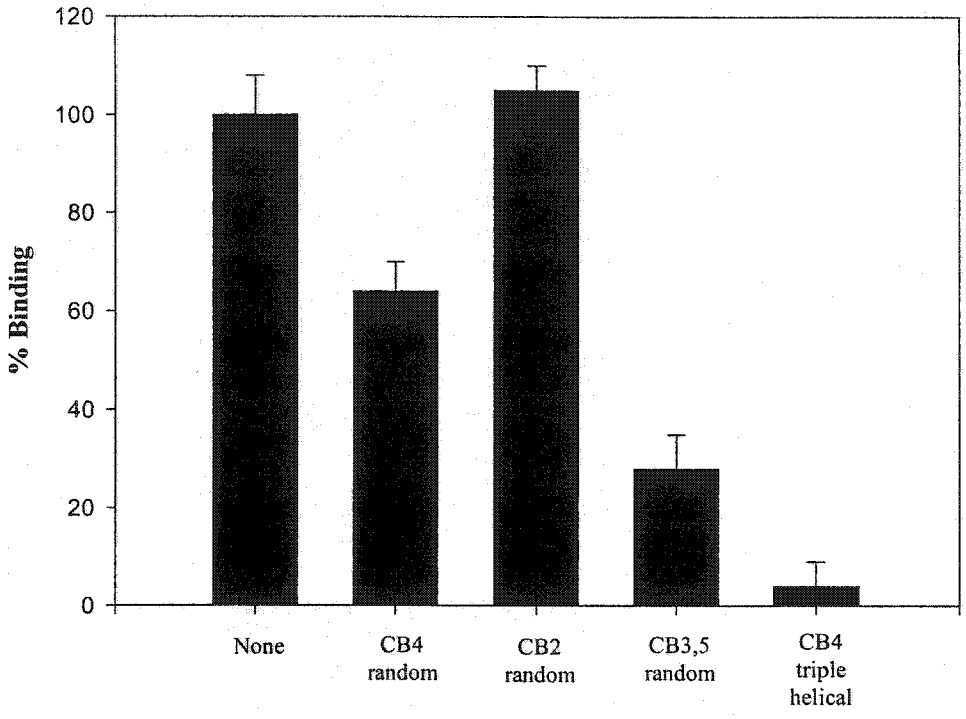


Figure 4.5c Inhibition of Hsp47 binding to collagen microtiter wells by CNBr fragments of type I collagen. Competition of ^{35}S-Hsp47 (0.6 μM) binding to collagen wells by CNBr peptides (12 μM) generated from the α2 chain of type I collagen. Peptides were in either a random and/or triple helical conformation as determined by CD. Data shown are the mean (±S.E.) of triplicate measurements.

CNBr-cleavage of the α1 chain of type I collagen results in seven peptides of varying lengths (Figure 4.1)(Rossi et al., 1996). The four large peptides, CB8, CB3, CB6 and CB7 all form triple-helical homotrimers in solution, whereas the smaller peptides, CB2, CB4 and CB5 adopt a random coil configuration at room temperature (Consonni et al., 2000; Rossi et al., 1996), possibly due to their short chain lengths. Each of the peptides was added at low μM concentrations to microtiter wells coated with type I collagen and their abilities to compete with collagen for ^{35}S-Hsp47 binding were examined. As shown in Figure 4.5, none of the CNBr peptides of the α1 chain had a significant effect on ^{35}S-Hsp47 binding to the collagen-coated wells when they were in the random coil conformation. However, in the triple-helical conformation, the peptides CB8, CB3, CB6 and CB7 were able to compete for ^{35}S-Hsp47 binding to the collagen-coated well. Among them, CB8 was found to exhibit the strongest binding to Hsp47 (Figure 4.5b).

CNBr-cleavage of the α2 chain of type I collagen yields three peptides of varying length, CB4, CB2 and CB3,5. Of these, only CB4 is able to form the triple-helical homotrimer in solution as reported earlier (Rossi et al., 1996) and as verified by our CD measurements (data not shown). Each of these peptides was assayed for its ability to bind ^{35}S-Hsp47 (Figure 4.5c). As with the shorter fragments of the α1 chain, Hsp47 binding to CB2 was insignificant. Maximum binding was seen with CB4 in the triple-helical conformation. This peptide showed reduced, but significant, binding to Hsp47 even in its random coil form. Moderately strong binding was observed with the CB3,5

peptide which, as reported earlier (Rossi et al., 1996), could not be obtained in the triple-helical conformation at room temperature.

The above results demonstrate that while Hsp47 binds to the CNBr peptides in solution when they adopt the triple-helix conformation, some peptides bind relatively more strongly than others. This was further tested by using the nitrocellulose transfer assay used earlier for the collagenase cleavage fragments. This assay singled out those peptides that possessed a relatively high binding affinity to Hsp47 so that some of the peptides identified in the radio-ligand competition assay were not picked up by this assay. As shown in Figure 4.6, labeled Hsp47 bound to the CB8 peptide of the α1 chain and CB4 peptide of the α2 chain. The bound Hsp47 could be released by the addition of 1 % gelatin to the binding buffer (data not shown). The results of this assay (Figure 4.6) combined with those obtained by the collagen competition assay (Figure 4.5) show that, although Hsp47 is capable of binding to several regions on the type I collagen molecule, the tightest binding occurs towards the N-terminus represented by the regions covered by CB8(α1) and CB4(α2), when these are in the triple-helical conformation.

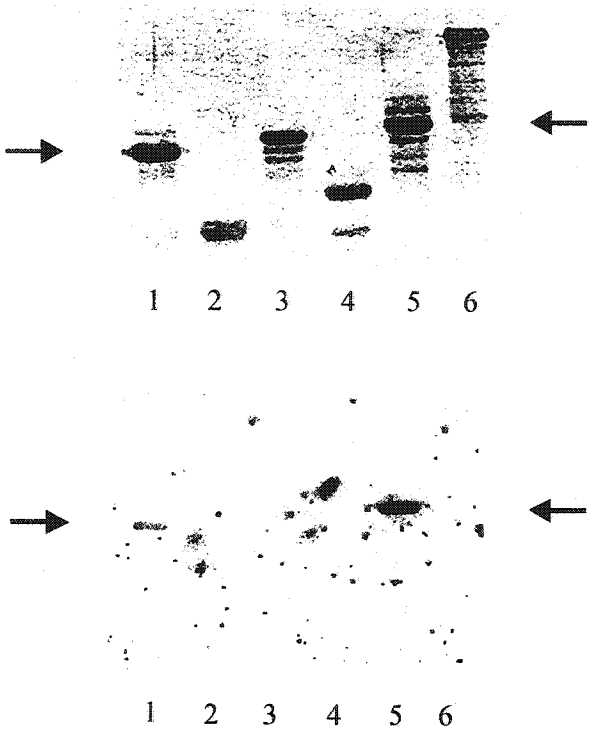


Figure 4.6 Binding of Hsp47 to representative CNBr fragments of type I collagen. CNBr peptides from the α1 (lanes 1-4) and α2 chains (lanes 5 & 6) of type I collagen were run in SDS-PAGE and were either stained with Coomassie Brilliant Blue (*top panel*) or transferred to nitrocellulose membrane and probed with ^{35}S-Hsp47 (*bottom panel*). *Lane 1,* CB8; *Lane 2,* CB3; *Lane 3,* CB7; *Lane 4,* CB6; *Lane 5,* CB4; *Lane 6,* CB3,5. The bottom panel shows an auto-radiograph of ^{35}S-Hsp47 binding to CB8(α1) and CB4(α2). Arrows indicate correspondence between the respective bands in the two panels.

4.5.5 Binding of Hsp47 to CNBr Peptides of Type II Collagen

CNBr cleavage of type II collagen results in the generation of six peptides of varying chain lengths (Figure 4.1)(Tenni et al., 2002). The CD spectral properties of CB6, CB12, CB11, CB8 and CB10 have been examined earlier (Tenni et al., 2002) and verified by us (data not shown). CB11, CB8 and CB10 are able to form triple-helical homotrimers in solution at room temperature. The random-coil CB6 and CB12 peptides did not show any significant binding to Hsp47 (Figure 4.7). Due to the limited supply of these peptides, experiments with CB11, CB8 and CB10 were performed only in their triple-helical and not random coil conformation. When added to the type I collagen-coated microtiter wells, all the triple-helical peptides competed for ^{35}S-Hsp47 binding with CB11 showing the strongest binding followed by CB10 and CB8 (Figure 4.7).

4.5.6 Binding of Hsp47 to Chemically Modified CB4 Peptide of α2(1)

To understand the possible involvement of lysine and hydroxylysine in Hsp47 binding, competition studies were conducted with chemically modified CB4 peptide of the α2 chain of type I collagen. The derivatized peptides were generated by N-methylation and N-acetylation of the primary amino group of lysine and hydroxylysine side chains under mild conditions as described previously (Tenni et al., 2002). Figure 4.8 shows that neither of these derivatives significantly affected CB4's ability to inhibit ^{35}S-Hsp47 binding to the wells. It may therefore be concluded that the absence of positive charges on lysine and hydroxylysine or the addition of a methyl group to these residues is not important for Hsp47 binding to collagen.

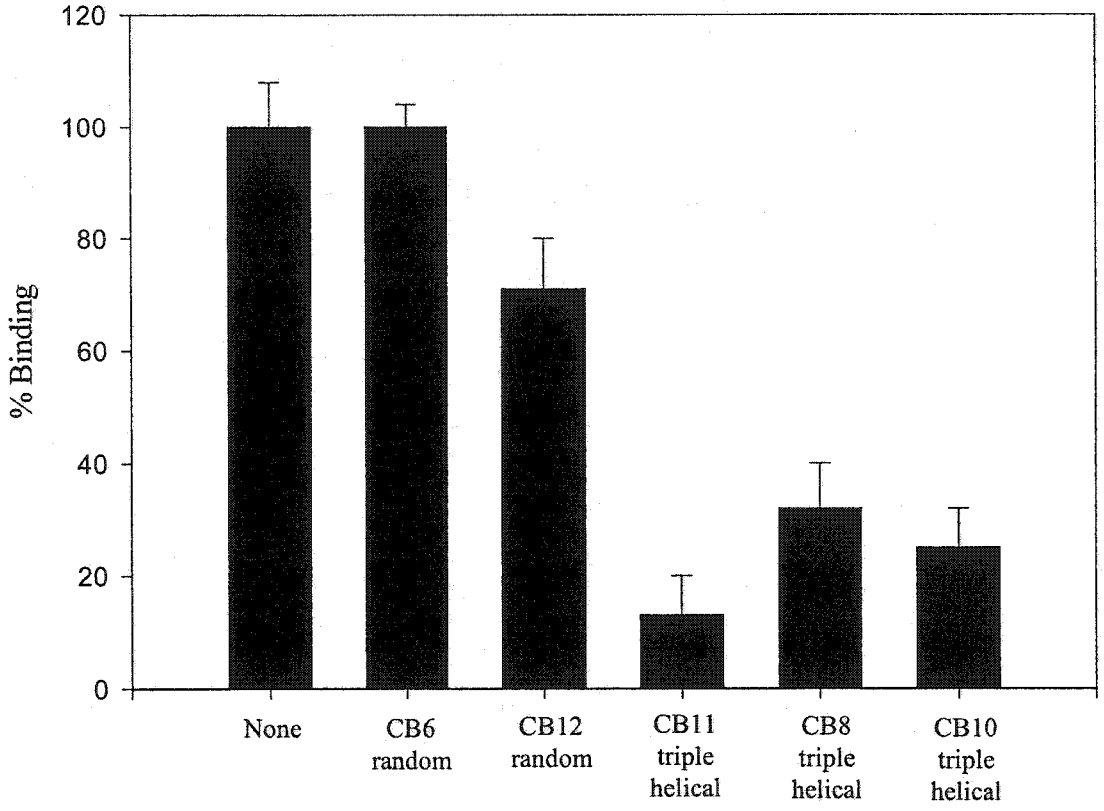


Figure 4.7 Inhibition of Hsp47 binding to collagen-coated microtiter wells by CNBr fragments of type II collagen. Competition of ^{35}S-Hsp47 (0.3 μM) binding to collagen wells by CNBr peptides (6 μM) generated from the α1 chain of type II collagen. Peptides used were in either a random and/or triple-helical configuration as determined by CD. Data shown are the mean (±S.E.) of triplicate measurements.

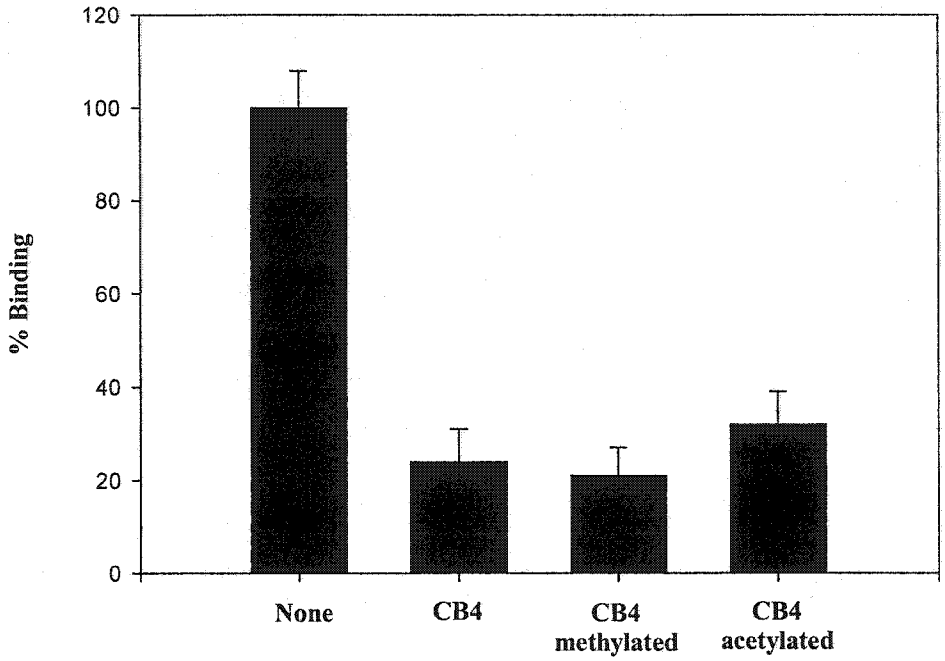


Figure 4.8 Inhibition of Hsp47 binding to collagen-coated microtiter wells by derivatized CB4 (α2) peptide of type I collagen. Competition of ^{35}S-Hsp47 (0.3 μM) binding to collagen wells by CB4 (α2) peptide modified by N-methylation and N-acetylation of its primary amino groups on lysine and hydroxylysine side chains. All peptides (3 μM) were in a triple helical conformation as determined by CD. Data shown are the mean (±S.E.) of triplicate measurements.

4.6 DISCUSSION

Hsp47 has been known to be intimately associated with one or more steps in the elaboration of nascent procollagen chains into the mature, functional collagen molecule. Knowledge of the structural details of the interaction of this chaperone protein with its substrate is, therefore, of importance from the basic and clinical points of view. In addition to its association with procollagen in the nascent, non-triple-helical form, Hsp47 also accompanies the triple-helical procollagen to the cis-Golgi compartment (Satoh et al., 1996). *In vitro*, Hsp47 binds to mature collagen type I-V with similar affinities ($\sim 10^{-7}$ M^{-1})(Natsume et al., 1994). In an earlier study, we showed that Hsp47 inhibits fibril formation among the triple-helical collagen monomer units *in vitro* (Thomson and Ananthanarayanan, 2000). It would be interesting to know the site(s) on collagen where Hsp47 binds and exerts its fibril inhibition effect. Several recent studies have sought to understand the intriguing collagen-binding behavior of Hsp47 and clarify the primary structural and conformational requirement for the binding. Using collagen model peptides *in vitro* (Koide et al., 1999) and an yeast two-hybrid system (Koide et al., 2000), Nagata and coworkers found that Hsp47 did not interact with the peptides in the single-stranded polyproline-II helix conformation and that the Gly-X-Y repeat sequence in the triple-helical, but not in the unfolded, conformation contained the required information for Hsp47 binding. With the significant exception of hydroxyproline (Hyp), residues in the Y position of the tripeptide repeat that enhance triple-helix stability were seen to also enhance Hsp47 binding. In particular, Arg and Pro residues were predominant in the peptides that showed enhanced Hsp47 binding. In contrast, Sauk et al. 2000, who

performed panning experiments with bacteriophage-peptides, found enriched binding of Hsp47 to sequences containing Trp, Leu, Val and Ala. In a recent study, Koide et al., found the X-Arg-Gly repeat forms the dominant Hsp47-binding site on collagen model peptides and in the homotrimeric type III collagen molecule (Koide et al., 2002).

The above structural studies using synthetic peptides have not been able to provide any clue about the location(s) in the triple-helical collagen molecule where Hsp47 is likely to bind. The recent suggestion by Koide et al. (2002) that Hsp47 would bind to X-Arg-Gly repeats in collagen, would imply an unrealistically large number of Arg-containing binding sites along the collagen molecule, since X-Arg-Gly occurs with a high frequency (13 %) in collagen (Yang et al., 1997). No indication of the relative affinities of these putative sites is obtainable from that study. In the present investigation, we have addressed this problem by examining discrete regions in the primary structure of collagen. This was made possible by the use of peptides generated by CNBr cleavage of the α chains of type I and type II collagens. These CNBr peptides have been well characterized for their primary structures and conformations in the laboratory of one of us (Consonni et al., 2000; Rossi et al., 1996; Tenni et al., 2002; Zanaboni et al., 2000).

To assess the binding of these peptides, we made use of the facile radio-labelling of the C-terminus of Hsp47 with ^{35}S during the purification of this protein using the Intein-based overexpression procedure (Evans and Ming-Qun, 2000) and developed a relatively simple binding assay (Figure 4.3). This competition assay is sensitive so that low μM amounts of the peptides were sufficient to compete with collagen bound to the microtiter wells. This is in contrast to other binding studies (Koide et al., 1999;

Macdonald and Bachinger, 2001) which necessitate the use of unrealistically high amounts of peptides. The competition assay used by us yielded a ranking of the peptides by their relative strengths of Hsp47 binding.

An alternate binding assay based on the visualization of ^{35}S-Hsp47 bound to peptides immobilized on nitrocellulose membrane was also used to examine binding of these as well as the collagenase cleavage fragments to Hsp47. This ligand blot assay was used as a verification of the results obtained from the competition assay on Hsp47 binding of the CNBr peptides. The ligand blot assay has been used in several other studies to identify ligand-binding sites in proteins (Kleinschmidt and Seiter, 1988; Kouklis et al., 1994). The results from this assay showed that both the α1 and α2 chains of collagen, as well as the N-terminal fragment A obtained by collagenase cleavage, bound Hsp47 (Figure 4.2). The absence of Hsp47 binding by the C-terminal fragment B would imply that the inhibition of collagen fibril formation by Hsp47 is mechanistically different from fibril inhibition by the peptides in fragment B between residues 776-796 of the α1(I) chain. These peptides were shown (at relatively high concentrations) to be capable of inhibiting fibril assembly and thought to represent regions important for the initial interactions during fibrillogenesis (Prockop and Fertala, 1998).

Examination of Hsp47 binding region(s) using the collagen competition assay showed that Hsp47 interacts weakly with $(Pro\text{-}Pro\text{-}Gly)_{10}$ and $(Pro\text{-}Hyp\text{-}Gly)_{10}$. This is in line with others' data (Macdonald and Bachinger, 2001). Also, as observed in the studies on synthetic peptides (Koide et al., 1999), the triple-helical conformation in the CNBr peptides was found to be more conducive for Hsp47 binding than the random coil

form (Figures 4.5 & 4.7). A range of affinities towards Hsp47 was found among the peptides binding Hsp47 in the triple-helical form. When comparing the inhibition of Hsp47 binding to the collagen wells by the peptides at a peptide: Hsp47 molar ratio of 20:1, by far the strongest inhibition (~96 %) was exhibited by the α2(I) CB4 peptide (Figure 4.5c). Very strong inhibition (~84-87 %) of Hsp47 binding was seen with α1(I) CB8 and α1(II) CB11 (Figures 4.5b & 4.7) and moderately strong inhibition (~68-75 %) was found with the α1(II) CB8 and CB10 peptides (Figure 4.7). Relatively weak inhibition was seen in the case of CB3, CB7 and CB6 peptides of the α1(I) chain (Figure 4.5b). Intriguingly, unlike the peptides derived from the α1(I) chain (Figure 4.5b), the peptides CB3,5 and CB4 of the α2(I) chain showed substantial inhibition even in their random coil forms (Figure 4.5c). The case of the CB3,5 peptide is particularly noteworthy since, in spite of its spanning nearly two-thirds of the α2(I) chain, it showed no significant triple-helical structure at room temperature (Rossi et al., 1996).

Taken together, the Hsp47 binding data on the CNBr peptides reinforce the idea that Hsp47 is capable of binding at several sites, possessing different affinities, along the collagen molecule and that some regions are able to bind the protein even in an apparently random coil form. Under similar experimental conditions, certain specific regions such as those represented by CB8 in the α1(I) chain, by CB4 in the α2(I) chain and by CB11 in the α1(II) chain form the strongest binding sites while they are in the triple-helical state. The former two peptides were also identified as strong binders in the nitrocellulose transfer assay (Figure 4.6). Interestingly, all the three peptides lie towards the N-terminus of the respective collagens. Because Arg residues are located throughout

the length of the collagen chain, it is unlikely that Arg alone represents the binding site for Hsp47.

In a recent study on the thermal stability of type I collagen it was suggested that collagen chaperone proteins such as Hsp47 may bind and stabilize regions in collagen which are relatively weak with respect to triple-helix stability (Leikina et al., 2002). These 'microunfolding regions' have been predicted to be important in collagen fibril formation (Kadler et al., 1988). Regions with relatively low triple-helix stability have earlier been identified in the α1(I) and α2(I) chains (Bachinger and Davis, 1991). These may well be candidates for micro-unfolding. However, it is unclear if Hsp47 would bind to these regions in light of the observed requirement for triple-helix for Hsp47 binding in this and earlier studies. Additional experiments using smaller fragments of the collagen chains with well-defined conformations are necessary to identify the optimal chain length, primary structure and tertiary structure required for Hsp47 binding. This will also help us understand the mechanism by which Hsp47 might prevent collagen fibril formation and design potent inhibitors that bind to Hsp47 *in vitro* and modify collagen biosynthesis under disease conditions such as fibrosis.

4.7 MATERIALS AND METHODS

4.7.1 Purification of Hsp47 and Radio-Labelling with ^{35}S-Cysteine

Hsp47 was purified as described previously using the IMPACT™ T7 system (New England Biolabs, Mississauga, Canada) followed by hydroxyapatite chromatography (Thomson and Ananthanarayanan, 2001). ^{35}S-cysteine labelling of

Hsp47 was performed as described elsewhere (Thomson and Ananthanarayanan 2001). It involved the incubation of 1.8 ml of purified Hsp47 (~150 μg/ml) with 5 μl of L-^{35}S-cysteine (1 μC); (NEN Life Science Products, Boston, MA) at 4°C for 8 hrs. The ^{35}S-Hsp47 was separated from free ^{35}S-cysteine using a PD-10 desalting column (Amersham Pharmacia Biotech, Quebec, Canada).

4.7.2 CNBr Peptides

The CNBr peptides used in the competition experiments and ligand blots were generated and characterized in acidic and neutral conditions as described elsewhere (Consonni et al., 2000; Rossi et al., 1996; Tenni et al., 2002; Zanaboni et al., 2000). The nomenclature of these peptides and their positions along the collagen molecule are shown in Figure 4.1. Each peptide was dissolved in a phosphate buffer (buffer A; 20 mM sodium phosphate, 50 mM NaCl, pH 7.4), its concentration verified by amino acid analysis (HSC/Pharmacia Biotechnology Service Center, Toronto, Canada), and the presence or absence of a triple helical conformation verified by CD spectroscopy. Molar concentrations are expressed in terms of the monomeric forms of the peptides. Random coil forms of the CNBr peptides were generated, immediately prior to experiments by incubation of the peptides at 50 °C for 3 min.

4.7.3 CD Spectroscopy

CD spectra used to determine the secondary structural characteristics of the peptides were obtained with a Jasco J-600 spectropolarimeter (JASCO Inc., Bethesda,

MD) using a 1 mm pathlength cell. Peptide concentrations were in the low μM range.

4.7.4 Collagenase Cleavage

To generate type I collagen fragments by digestion with vertebrate collagenase, 60 μg of acid-soluble calf skin collagen (Worthington Biochemical Corporation, Freehold, NJ) was incubated with 30 μl of a 38 μg/ml solution of the enzyme (MMP-1; Sigma Chemical Co., St. Louis, MO) to a final volume of 305 μl in 100 mM Tris, pH 7.4, 20 mM $CaCl_2$, 100 mM NaCl. The cleavage reaction was allowed to proceed at room temperature overnight. The reaction products were separated under reducing conditions in 10 % SDS-PAGE.

4.7.5 Ligand Blot Assay

Collagenase fragments and CNBr peptides were separated using 10 % and 12.5 % SDS-PAGE respectively, and were either stained with Coomassie Brilliant Blue or transferred to nitrocellulose at 100 V for 4 hrs. The membrane was blocked with 1 % BSA (Sigma), 1 % skim milk powder, 0.05 % Tween-20 (Sigma) in buffer A. The blot was rinsed for 5 min with the buffer and then incubated for 1 hr at room temperature in the same buffer containing 1 ml of ^{35}S-Hsp47. After binding, the excess ^{35}S-Hsp47 was decanted and the blot washed twice for 5 min in buffer A. The blot was then dried and exposed to an X-ray imaging film placed inside an intensifying screen (Kodak Laboratories, Rochester, NY).

4.7.6 Competition Assay

48-well plastic microtiter plates coated with rat tail collagen type I (BD Biosciences, Mississauga, Canada) were used to examine the interaction of Hsp47 with collagen. Individual wells were incubated in a constant final volume of 300 µl of buffer A containing increasing concentrations of ^{35}S-Hsp47 at room temperature (23±2°C) for 20 min. The binding solution was decanted and the wells washed twice with buffer A for 5 min. A 2 % SDS solution was added to the wells which were then left overnight with gentle agitation to dissociate the bound protein. The SDS solution was subsequently added to scintillation fluid for counting. Non-specific binding was determined by measuring ^{35}S-Hsp47 binding to the wells in the presence and absence of 0.1 % gelatin in buffer A. Specific binding was calculated by subtracting nonspecific binding from total binding at each concentration. All experiments were performed in triplicate using independent Hsp47 preparations. For competition experiments the radio-ligand assays were performed as above except that the peptides were added to the well buffer prior to the addition of a constant amount of ^{35}S-Hsp47. The nonspecific binding found in the presence of 0.1 % gelatin was subtracted from the binding values obtained. Normalized values were used to plot the histograms.

4.8 ACKNOWLEDGEMENTS

This work was supported by a grant to V.S. Ananthanarayanan from the Heart and Stroke Foundation of Ontario. C.A. Thomson was a recipient of a Research Traineeship from the Heart and Stroke Foundation of Canada.

5.9 REFERENCES

Bachinger, H. P., and Davis, J. M. (1991). Sequence specific thermal stability of the collagen triple helix. Int J Biol Macromol *13*, 152-6.

Consonni, R., Zetta, L., Longhi, R., Toma, L., Zanaboni, G., and Tenni, R. (2000). Conformational analysis and stability of collagen peptides by CD and by 1H- and 13C-NMR spectroscopies. Biopolymers *53*, 99-111.

Evans, T., and Ming-Qun, X. (2000). Intein-mediated protein ligation: Harnessing nature's escape artists. Biopolymers (Pep. Sci.) *51*, 355-362.

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

Gross, J., Harper, E., Harris, E. D., McCroskery, P. A., Highberger, J. H., Corbett, C., and Kang, A. H. (1974). Animal collagenases: specificity of action, and structures of the substrate cleavage site. Biochem Biophys Res Commun *61*, 605-12.

Hebert, C., Coletta, R. D., Norris, K., Nikitakis, N., Lopes, M., and Sauk, J. J. (2001). Non-natural CBP2 binding peptides and peptomers modulate carcinoma cell adhesion and invasion. J Cell Biochem *82*, 145-54.

Helenius, A., Marquardt, T., and Braakman, I. (1992). The endoplasmic reticulum as a protein-folding compartment. Trends in Cell Biology *2*, 227-231.

Hu, G., Gura, T., Sabsay, B., Sauk, J., Dixit, S. N., and Veis, A. (1995). Endoplasmic reticulum protein Hsp47 binds specifically to the N- terminal globular domain of the amino-propeptide of the procollagen I alpha 1 (I)-chain. J Cell Biochem *59*, 350-67.

Jain, N., Brickenden, A., Lorimer, I., Ball, E. H., and Sanwal, B. D. (1994). Interaction of procollagen I and other collagens with colligin. Biochem J *304*, 61-8.

Kadler, K. E., Hojima, Y., and Prockop, D. J. (1988). Assembly of type I collagen fibrils de novo. Between 37 and 41 degrees C the process is limited by micro-unfolding of monomers. J Biol Chem *263*, 10517-23.

Kleinschmidt, J. A., and Seiter, A. (1988). Identification of domains involved in nuclear uptake and histone binding of protein N1 of Xenopus laevis. Embo J *7*, 1605-14.

Koide, T., Asada, S., and Nagata, K. (1999). Substrate recognition of collagen-specific molecular chaperone HSP47. Structural requirements and binding regulation. J Biol Chem *274*, 34523-6.

Koide, T., Aso, A., Yorihuzi, T., and Nagata, K. (2000). Conformational requirements of collagenous peptides for recognition by the chaperone protein HSP47. J Biol Chem *275*, 27957-63.

Koide, T., Takahara, Y., Asada, S., and Nagata, K. (2002). Xaa-Arg-Gly triplets in the collagen triple helix are dominant binding sites for the molecular chaperone HSP47. J Biol Chem *277*, 6178-82.

Kojima, T., Miyaishi, O., Saga, S., Ishiguro, N., Tsutsui, Y., and Iwata, H. (1998). The retention of abnormal type I procollagen and correlated expression of HSP 47 in fibroblasts from a patient with lethal osteogenesis imperfecta. J Pathol *184*, 212-8.

Kouklis, P. D., Hutton, E., and Fuchs, E. (1994). Making a connection: direct binding between keratin intermediate filaments and desmosomal proteins. J Cell Biol *127*, 1049-60.

Leikina, E., Mertts, M. V., Kuznetsova, N., and Leikin, S. (2002). Type I collagen is thermally unstable at body temperature. Proc Natl Acad Sci U S A *99*, 1314-8.

Macdonald, J. R., and Bachinger, H. P. (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*, 25399-403.

Nagai, N., Hosokawa, M., Itohara, S., Adachi, E., Matsushita, T., Hosokawa, N., and Nagata, K. (2000). Embryonic lethality of molecular chaperone hsp47 knockout mice is associated with defects in collagen biosynthesis. J Cell Biol *150*, 1499-506.

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

Natsume, T., Koide, T., Yokota, S., Hirayoshi, K., and Nagata, K. (1994). Interactions between collagen-binding stress protein HSP47 and collagen. Analysis of kinetic parameters by surface plasmon resonance biosensor. J Biol Chem *269*, 31224-8.

Prockop, D. J., and Fertala, A. (1998). Inhibition of the self-assembly of collagen I into fibrils with synthetic peptides. Demonstration that assembly is driven by specific binding sites on the monomers. J Biol Chem *273*, 15598-604.

Rossi, A., Zuccarello, L. V., Zanaboni, G., Monzani, E., Dyne, K. M., Cetta, G., and Tenni, R. (1996). Type I collagen CNBr peptides: species and behavior in solution. Biochemistry *35*, 6048-57.

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

Sauk, J. J., Coletta, R. D., Norris, K., and Hebert, C. (2000). Binding motifs of CBP2 a potential cell surface target for carcinoma cells. J Cell Biochem *78*, 251-63.

Sauk, J. J., Smith, T., Norris, K., and Ferreira, L. (1994). Hsp47 and the translation-translocation machinery cooperate in the production of alpha 1(I) chains of type I procollagen. J Biol Chem *269*, 3941-6.

Smith, T., Ferreira, L. R., Hebert, C., Norris, K., and Sauk, J. J. (1995). Hsp47 and cyclophilin B traverse the endoplasmic reticulum with procollagen into pre-Golgi intermediate vesicles. A role for Hsp47 and cyclophilin B in the export of procollagen from the endoplasmic reticulum. J Biol Chem *270*, 18323-8.

Tenni, R., Viola, M., Welser, F., Sini, P., Giudici, C., Rossi, A., and Tira, M. E. (2002). Interaction of decorin with CNBr peptides from collagens I and II. Eur J Biochem *269*, 1428-37.

Thomson, C. A., and Ananthanarayanan, V. S. (2001). A method for expression and purification of soluble, active Hsp47, a collagen-specific molecular chaperone. Protein Expr Purif *23*, 8-13.

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

Yang, W., Chan, V. C., Kirkpatrick, A., Ramshaw, J. A., and Brodsky, B. (1997). Gly-Pro-Arg confers stability similar to Gly-Pro-Hyp in the collagen triple-helix of host-guest peptides. J Biol Chem *272*, 28837-40.

Zanaboni, G., Rossi, A., Onana, A. M., and Tenni, R. (2000). Stability and networks of hydrogen bonds of the collagen triple helical structure: influence of pH and chaotropic nature of three anions. Matrix Biol *19*, 511-20.

CHAPTER 5: CONCLUSIONS AND SUMMARY

5.1 TOWARDS UNDERSTANDING HSP47'S STRUCTURE AND FUNCTION

5.1.1 Purification of Hsp47

The aim of this thesis was the characterization of the structure of Hsp47 and of its interactions with collagen in an attempt to gain insight into the function of Hsp47 as a collagen-specific molecular chaperone. Initial efforts to purify Hsp47 were fraught with difficulties including problems with expression, solubility, purity, degradation and low yields. Many of these obstacles were overcome by use of the IMPACT^{TM} system, which not only allowed for the purification of soluble Hsp47 in mg quantities but also allowed for the facile labelling of the protein with ^{35}S-cysteine (Chapter 3). As a result, sufficient quantities of Hsp47 were purified for structural studies and the radio-labelled Hsp47 was used in binding studies.

However, it should be noted that use of the IMPACT^{TM} system did not resolve all of the difficulties in working with Hsp47 *in vitro*. In particular, the purified Hsp47 was still susceptible to degradation and formed insoluble aggregates over a relatively short period of time such that all of the experiments illustrated in this thesis were completed within 16 hrs of individual Hsp47 purifications. Attempts to stabilize or refold Hsp47 were unsuccessful so the question is raised as to why? An answer to this question is that Hsp47 is phosphorylated and glycosylated *in vivo* and that the lack of these modifications in recombinant Hsp47 could affect the protein's stability *in vitro*. This theory is not unreasonable but it is likely incorrect. This is because studies by Vaillancourt and Cates, (1991) indicated that Hsp47 purified from myoblasts, and thus phosphorylated and glycosylated, is only active for 48 hrs and is also inactivated by freezing and thawing. In

hindsight, the difficulties in stabilizing Hsp47 probably arise from its being a member of the serpin family. If homology modelling (Tasab et al., 2002) is correct, Hsp47 contains a reactive centre loop that acts as bait for serine proteases (Huntington et al., 2000). However, it does not contain all residues key for inactivation of the proteases and conceivably it is highly susceptible to being proteolyzed. Similarly, it is known that serpins exist in a metastable state when active and that subsequent to cleavage they quickly undergo structural rearrangement to a highly stable conformation (i.e. 'stressed to relaxed') (Bruch et al., 1988). The structural mobility of serpins may be responsible for difficulties in stabilizing Hsp47 *in vitro*. As well, polymers of serpins have been implicated in various disease states and crystallographic data show that they are the result of insertion of part of the reactive center loop of one molecule into a β-sheet of another (Dunstone et al., 2000). Aggregates of Hsp47 *in vitro* may be the result of this type of polymerization. Crystallographic data will be necessary to determine if Hsp47 does indeed mimic any known serpin structures and to what extent.

5.1.2 Structural Characterization of Hsp47

CD and fluorescence spectroscopy studies of Hsp47 revealed that it undergoes pH-induced structural changes. These changes in structure correlated with the pH-dependent binding and release of collagen by Hsp47. We therefore proposed that the pH-induced structural changes were responsible for binding and release of collagen by Hsp47, in the neutral ER and acidic Golgi, respectively. This mechanism of regulating binding and release of collagen by Hsp47 is a plausible one as other factors such as the

presence or absence of ATP have not been found to play a role (Nakai et al., 1992). Moreover, as mentioned previously, molecular modelling of Hsp47 mapped His residues to regions important for serpin conformational change (Dafforn et al., 2001).

During the purification of Hsp47, it was noted that Hsp47 existed as a monomer with a small proportion of trimer being present. Western analysis of low density microsomes isolated from fibroblasts, also revealed the presence of an Hsp47 trimer under non-reducing conditions. To address whether or not the ability to trimerize was necessary for Hsp47's activity, we performed our fibre assay under reducing conditions and showed that the ability of Hsp47 to prevent collagen fibrillogenesis was impaired (Figure 3.4c). This may have important *in vivo* implications as a trimeric form of Hsp47 has been identified during wound healing in an animal model (Wang et al., 2002).

Many other heat shock proteins have been shown to undergo conversion between monomeric and oligomeric forms (Miron et al., 1991; Wearsch and Nicchitta, 1996). It is believed that the ability to interconvert between these forms is associated with chaperone activity and the need of the cell to keep a constant supply these chaperone proteins for quick mobilization (Blond-Elguindi et al., 1993; Kokke et al., 1998). Our results agree with this idea and implicate Hsp47 as being active in its trimeric form. However, interpretation may not be that simple. This is because Hsp47 was shown to bind collagen in both its monomeric and trimeric forms (Dafforn et al., 2001) which raises the question if these are really 'inactive' and 'active' forms of the protein.

Hsp27 is a molecular chaperone and regulator of actin polymerization whose chaperone activity has been shown to be regulated by changes in oligomeric size (Miron

et al., 1991; Rogalla et al., 1999). Interestingly, S-thiolation was recently shown to disaggregate multimeric Hsp27 under conditions of oxidative stress (Eaton et al., 2002). Hsp27 contains a single cysteine, similar to Hsp47, and it would therefore be relevant to investigate in the future whether or not conditions such as oxidative stress affect the trimerization of Hsp47.

5.1.3 Function of Hsp47

The ultimate goal of this thesis was to determine a role for Hsp47 in collagen biosynthesis from *in vitro* studies. However, this proved a formidable task and many questions are still unanswered. My initial studies of Hsp47 were aimed at showing that Hsp47 could assist in collagen triple-helix formation or that Hsp47 played a role in preventing the thermal unfolding of collagen. These experiments were not fruitful and it wasn't until we investigated the role of Hsp47 in preventing collagen aggregation that an activity for Hsp47 was revealed. As outline in Chapter 2, Hsp47 effectively inhibited the fibrillogenesis of type I collagen. The ability of Hsp47 to inhibit collagen fibrillogenesis was dependent on the addition of Hsp47 during the lag phase of fibrillogenesis and this, not surprisingly, was pH-dependent.

To further our understanding of how Hsp47 might be inhibiting collagen fibrillogenesis, we initiated studies to map the binding of Hsp47 on collagen using CNBr peptides derived from type I and type II collagen (Chapter 4). Amino acids 781 to 794 of the α1(I) chain had been previously implicated as being important for initiating fiber formation (Prockop and Fertala, 1998). Our data shows that Hsp47 preferentially binds

to the opposite end of the collagen molecule and is therefore not likely exerting its fibre inhibition activity at this proposed site of fibre initiation. Although we were able to confirm that Hsp47 preferentially binds to collagen in its triple helical conformation, we also showed binding to peptides in the random coil form. According to the current literature, it is thought that Hsp47 binds to the triple-helical conformation of collagen and that Arg is required for Hsp47 binding with the presence of Hyp not having a major effect (Koide et al., 2002; Tasab et al., 2002). Our results do not refute these suggestions. However, they further show that binding of Hsp47 to collagen requires more than simply the presence of an Arg residue as Arg is found throughout the collagen molecule (Table 5.1) and we show binding by Hsp47 to collagen in discrete regions.

At this stage we may ask ourselves if by combining the results of the binding studies with that of the fibre inhibition by Hsp47, have we gained any insight into Hsp47's role as a molecular chaperone? We have shown that Hsp47 inhibits collagen fibrillogenesis and we have shown that it preferentially binds to triple-helical CNBr peptides derived from the N-terminus of the type I collagen molecule (e.g. α1 CB8, α2 CB4). What structural attributes do these peptides have in common that peptides derived from other regions of the molecule do not? Perhaps there is a preponderance of charged or hydrophobic amino acids in the Hsp47-binding peptides. This may not be the case since as mentioned in the introduction, the D period of collagen represents a repeating unit based on the distribution of hydrophobic and hydrophilic amino acids. This is further illustrated in Table 5.1 where the percentage of charged and hydrophobic amino acids is shown to be similar between all of the larger CNBr peptides of type I collagen.

In particular, one may compare the frequency of charged and hydrophobic amino acids in the pairs: α1 CB8 & α1 CB7 and α2 CB4 & α1 CB3, in which the first peptide if each pair shows significant binding by Hsp47 and the second does not.

Table 5.1

Compositional Analysis of Selected CNBr Peptides from Type I Collagen.

CNBr Peptide	% by Frequency of Charged Amino Acids (RKHYCDE)	% by Frequency of Hydrophobic Amino Acids (AILFWV)	% by Frequency of Arg
α1 CB8	16	18	5
α1 CB3	16	21	4
α1 CB7	17	18	5
α1 CB6	17	16	6
α2 CB4	15	22	5
α2 CB3,5	17	20	5

Analysis of the percentage of charged and hydrophobic amino acids along the collagen polypeptide chains is obviously an oversimplification of its primary structure. However, it does point to the difficulties that can be encountered when trying to map the binding of Hsp47 onto a molecule with such a repetitive sequence. As mentioned above, Arg has been implicated as being a key residue for Hsp47 binding to collagen (Koide et al., 2002; Tasab et al., 2002). X-ray crystallography studies of a collagen peptide show

that Arg residues can make direct contacts from one of their side-chain nitrogens to backbone carbonyl groups (Kramer et al., 2001). Koide et al., (2002) suggest that Hsp47 specifically recognizes and binds to collagen sequences containing an Arg residue in such a conformation. However, it is unclear whether or not all Arg residues in the collagen molecule exist in this conformation. Perhaps it is those Arg residues that make direct contacts from their side-chains to the peptide backbone and that are located in a stretch of hydrophobic amino acids that are specifically recognized by Hsp47.

Another approach to understanding the recognition elements in collagen by Hsp47 is to look at the process of fibrillogenesis and collagen conformation. It has been proposed that Hsp47 acts as a chaperone by binding to thermally unstable procollagen *in vivo* (Leikina et al., 2002). If it is the 'microunfolding' of collagen that gives it the 'flex' necessary to form fibres, it is possible that by masking the sites of 'microunfolding', Hsp47 may, in essence, prevent fibre formation. This concept has been investigated previously and negative results have been generated (Macdonald and Bachinger, 2001; Tasab et al., 2002). However, I believe this concept deserves a second look for a variety of reasons. Firstly, Vaillancourt and Cates (1991) showed that Hsp47 was capable of increasing the T_m of collagen by 1°C. They dismiss this result as being insignificant, but in lieu of the recent work of Leikina et al. (2002) it may be more significant than originally thought. This is because Leikina et al. (2002) show that collagen is unstable at body temperature and that a change in T_m of 1°C can have profound effects on the rate of collagen melting. Secondly, experiments based on selective cleavage of type I collagen with proteases as a means of probing the temporary relaxation of triple-helical

conformation, implicate the N-terminal three quarters of the molecule as being the primary site of microunfolding (Ryhanen et al., 1983). Coincidentally, it is this portion of the collagen molecule that we see as having the greatest degree of Hsp47 binding. Thirdly, Hsp47 is a heat shock protein that is upregulated under conditions that would exaggerate the unfolding of the collagen triple-helix. Taken together, I feel that there is sufficient reason to reinvestigate the potential of Hsp47 to bind to regions of 'microunfolding' and thus maintain the structural stability of procollagen *in vivo*.

5.2 FUTURE DIRECTIONS

The work presented in this thesis provides a strong starting point for future research on Hsp47. Our results show that Hsp47 preferentially binds to the N-terminal regions of the collagen molecule, however, the exact sequence of amino acids involved in binding is unknown. Future studies should therefore include attempts to narrow down the binding region of Hsp47 on collagen. In order to perform these studies, CNBr peptides shown to bind Hsp47 will have to be isolated as smaller fragments. To do this, selected sequences can be cloned into expression vectors and purified. Peptides purified in this manner will not contain Hyp residues. Conversely, peptide fragments can also be synthesized so that they will contain Hyp, however, considering the large regions of collagen that we are investigating this may prove a very expensive endeavour.

One of the major outcomes of this work was the identification of an activity assay for Hsp47. Recently, McMaster University opened a High Throughput Screening Facility and work is presently underway in our laboratory to identify small molecules that

would act as agonists or antagonists of Hsp47 activity. The impact of this work is potentially twofold. Firstly, the results of this thesis implicate Hsp47 as inhibiting the premature aggregation of procollagen *in vivo*. This proposed function is gaining support in the literature (Macdonald and Bachinger, 2001; Tasab et al., 2002), however it has yet to be proven *in vivo*. The ability to turn Hsp47 'on' or 'off', *in situ*, by the addition of a small molecule, may prove invaluable in determining the actual function of Hsp47. Similarly, it may further our understanding of collagen biosynthesis by allowing us to monitor its fate when Hsp47 is no longer able to function.

Secondly, antagonists of Hsp47 function may prove useful in design of therapeutics for pathological conditions resulting from collagen accumulation. Presently diseases such as fibrosis leave a patient with few medicinal options. Compounds inhibiting prolyl 4-hydroxylase activity are currently being identified but they have not yet been developed into antifibrotic drugs (As in Myllyharju and Kivirikko, 2001). Hsp47 is a good target for design of antifibrotic drugs because it has been shown to be necessary for collagen biosynthesis (Nagai et al., 2000). As well, the spatial and temporal expression of Hsp47 is highly correlated with that of its substrate, collagen, in health and disease (Nagata, 1998).

5.3 CLOSING REMARKS

Heat shock proteins perform a variety of functions in cells that can be generalized under two broad categories (Parsell and Lindquist, 1993). The first is their role as molecular chaperones that are required during the protein folding to prevent off-pathway

interactions and aggregation. The second is their role in the recognition of aberrant proteins and their subsequent turnover. Our results point to a role for Hsp47 in preventing the premature aggregation of procollagen through binding to the triple-helical N-terminal region of the collagen molecule. These *in vitro* studies characterizing the interaction of Hsp47 with collagen may potentially aid in our understanding of the complex collagen biosynthetic pathway and in the development of inhibitors against abnormal collagen accumulation in disease conditions.

5.4 REFERENCES

Blond-Elguindi, S., Fourie, A. M., Sambrook, J. F., and Gething, M. J. (1993). Peptide-dependent stimulation of the ATPase activity of the molecular chaperone BiP is the result of conversion of oligomers to active monomers. J Biol Chem *268*, 12730-5.

Bruch, M., Weiss, V., and Engel, J. (1988). Plasma serine proteinase inhibitors (serpins) exhibit major conformational changes and a large increase in conformational stability upon cleavage at their reactive sites. J Biol Chem *263*, 16626-30.

Dafforn, T. R., Della, M., and Miller, A. D. (2001). The molecular interactions of heat shock protein 47 (Hsp47) and their implications for collagen biosynthesis. J Biol Chem *276*, 49310-9.

Dunstone, M. A., Dai, W., Whisstock, J. C., Rossjohn, J., Pike, R. N., Feil, S. C., Le Bonniec, B. F., Parker, M. W., and Bottomley, S. P. (2000). Cleaved antitrypsin polymers at atomic resolution. Protein Sci *9*, 417-20.

Eaton, P., Fuller, W., and Shattock, M. J. (2002). S-thiolation of HSP27 regulates its multimeric aggregate size independently of phosphorylation. J Biol Chem *277*, 21189-96.

Huntington, J. A., Read, R. J., and Carrell, R. W. (2000). Structure of a serpin-protease complex shows inhibition by deformation. Nature *407*, 923-6.

Koide, T., Takahara, Y., Asada, S., and Nagata, K. (2002). Xaa-Arg-Gly triplets in the collagen triple helix are dominant binding sites for the molecular chaperone HSP47. J Biol Chem *277*, 6178-82.

Kokke, B. P., Leroux, M. R., Candido, E. P., Boelens, W. C., and de Jong, W. W. (1998). Caenorhabditis elegans small heat-shock proteins Hsp12.2 and Hsp12.3 form tetramers and have no chaperone-like activity. FEBS Lett *433*, 228-32.

Kramer, R. Z., Bella, J., Brodsky, B., and Berman, H. M. (2001). The crystal and molecular structure of a collagen-like peptide with a biologically relevant sequence. J Mol Biol *311*, 131-47.

Leikina, E., Mertts, M. V., Kuznetsova, N., and Leikin, S. (2002). Type I collagen is thermally unstable at body temperature. Proc Natl Acad Sci U S A *99*, 1314-8.

Macdonald, J. R., and Bachinger, H. P. (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*, 25399-403.

Miron, T., Vancompernolle, K., Vandekerckhove, J., Wilchek, M., and Geiger, B. (1991). A 25-kD inhibitor of actin polymerization is a low molecular mass heat shock protein. J Cell Biol *114*, 255-61.

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

Nagai, N., Hosokawa, M., Itohara, S., Adachi, E., Matsushita, T., Hosokawa, N., and Nagata, K. (2000). Embryonic lethality of molecular chaperone hsp47 knockout mice is associated with defects in collagen biosynthesis. J Cell Biol *150*, 1499-506.

Nagata, K. (1998). Expression and function of heat shock protein 47: a collagen-specific molecular chaperone in the endoplasmic reticulum. Matrix Biol *16*, 379-86.

Nakai, A., Satoh, M., Hirayoshi, K., and Nagata, K. (1992). Involvement of the stress protein HSP47 in procollagen processing in the endoplasmic reticulum. J Cell Biol *117*, 903-14.

Parsell, D. A., and Lindquist, S. (1993). The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. Annu Rev Genet *27*, 437-96.

Prockop, D. J., and Fertala, A. (1998). Inhibition of the self-assembly of collagen I into fibrils with synthetic peptides. Demonstration that assembly is driven by specific binding sites on the monomers. J Biol Chem *273*, 15598-604.

Rogalla, T., Ehrnsperger, M., Preville, X., Kotlyarov, A., Lutsch, G., Ducasse, C., Paul, C., Wieske, M., Arrigo, A. P., Buchner, J., and Gaestel, M. (1999). Regulation of Hsp27 oligomerization, chaperone function, and protective activity against oxidative stress/tumor necrosis factor alpha by phosphorylation. J Biol Chem *274*, 18947-56.

Ryhanen, L., Zaragoza, E. J., and Uitto, J. (1983). Conformational stability of type I collagen triple helix: evidence for temporary and local relaxation of the protein conformation using a proteolytic probe. Arch Biochem Biophys *223*, 562-71.

Tasab, M., Jenkinson, L., and Bulleid, N. J. (2002). Sequence-specific recognition of collagen triple helices by the collagen-specific molecular chaperone HSP47. J Biol Chem *277*, 35007-12.

Vaillancourt, J. P., and Cates, G. A. (1991). Purification and reconstitution of a collagen-binding heat-shock glycoprotein from L6 myoblasts. Biochem J *274*, 793-8.

Wang, J. F., Olson, M. E., Winkfein, R. J., Kulyk, W. M., Wright, J. B., and Hart, D. A. (2002). Molecular and cell biology of porcine HSP47 during wound healing: complete cDNA sequence and regulation of gene expression. Wound Repair Regen *10*, 230-40.

Wearsch, P. A., and Nicchitta, C. V. (1996). Endoplasmic reticulum chaperone GRP94 subunit assembly is regulated through a defined oligomerization domain. Biochemistry *35*, 16760-9.