Sequence-specific Recognition of Collagen Triple Helices by the Collagen-specific Molecular Chaperone HSP47*

Mohammed Tasab, Lynsey Jenkinson, and Neil J. Bulleid‡

From the School of Biological Sciences, 2.205 Stopford Building, University of Manchester, Manchester M13 9PT, United Kingdom

HSP47 is a molecular chaperone that plays an unknown role during the assembly and transport of procollagen. Our previous studies showed that, unlike most chaperones, HSP47 interacts with a correctly folded substrate. We suggested that HSP47 either stabilizes the correctly folded collagen helix from heat denaturation or prevents lateral aggregation prior to its transport from the endoplasmic reticulum. In this study we have addressed the role of triple helix stability in the binding of HSP47 to procollagen by expressing procollagen molecules with differing thermal stabilities and analyzing their ability to interact with HSP47 within the endoplasmic reticulum. Our results show that HSP47 interacts with thermostable procollagen molecules, suggesting that helix stabilization is not the primary function of HSP47 and that the interaction of HSP47 with procollagen depends upon the presence of a minimum of one Gly-X-Arg triplet within the triple helical domain. Interestingly, procollagen chains containing high proportions of stabilizing triplets formed triple helices and interacted with HSP47 even in the absence of proline hydroxylation, demonstrating that recognition does not depend upon this modification. Our results support the view that HSP47 functions early in the secretory pathway by preventing the lateral aggregation of procollagen chains.

The folding and assembly of procollagen molecules occurs within the lumen of the endoplasmic reticulum (ER) following translation and translocation of procollagen chains on membrane-bound ribosomes. Post-translational modifications take place as the chains enter the lumen of the ER, and the assembly process is known to involve interactions with multiple ER resident chaperones (1). Each procollagen chain contains a collagenous domain of repeating Gly-X-Y triplets that determines the characteristic triple helical structure of collagens. The contribution of specific Gly-X-Y triplets to the thermal stability of the triple helical structure depends upon the identity of the X and Y residues (2). The C-propeptide regions associate to form a trimeric procollagen molecule (3). This association ensures that the triple helix nucleation regions of the adjacent chains (4) are in close proximity, allowing for helix nucleation to occur followed by propagation of the triple helix toward the N terminus (5). As well as carrying out post-translational modifications, proteins such as protein disulfide isomerase and prolyl 4-hydroxylase assist in the folding process either by preventing nonspecific association of the individual chains prior to trimerization and triple helix formation (6) or by helping to retain unfolded molecules within the ER (7).

HSP47 is an ER resident glycoprotein that plays a crucial role during the folding, maturation, and secretion of procollagen (8). Evidence to support this role comes from the fact that HSP47 is up-regulated in collagen-producing cells (9–11) and binds specifically to collagenous peptides (12, 13). More convincingly, when the hsp47 gene was disrupted in mice, the resulting nullizygous animals were severely deficient in the mature processed form of collagen and did not survive past day 11.5 postcoitus (14). These results demonstrate that HSP47 is required to facilitate the secretion of stable and correctly processed collagen and that mice cannot develop without this specialized molecular chaperone. However, this work does not reveal the precise function of HSP47, whether HSP47 recognizes a specific sequence within the collagenous domain, or how binding of HSP47 to procollagen is regulated within the cell.

One approach that has been taken to address the question of substrate recognition has been to study the binding of collagen-like peptides to HSP47 in vitro (13). These studies demonstrate that HSP47 binds with highest affinity to peptides that are capable of forming thermostable triple helices. The binding of peptides to HSP47 can be abolished by lowering the pH, leading to the suggestion that the regulation of binding in vivo could be due to changes in pH during transport through the secretory pathway (15). These in vitro binding studies have also shown that the affinity of HSP47 for peptides containing arginine residues within the Y position of the Gly-X-Y collagen triplet is much higher than when this amino acid is absent (16). However a recent report on binding of HSP47 to (PGP)_{10} peptide demonstrated high affinity binding (15). Although there are conflicting results in the literature, these experiments do establish that HSP47 binds preferentially to procollagen peptides that adopt a triple helical conformation and contain the amino acid arginine.

Our previous studies demonstrated that HSP47 binds to procollagen molecules once they have formed triple helical molecules and dissociates upon thermal denaturation of the triple helical structure (17). The approach taken was to analyze the interaction of HSP47 within newly synthesized procollagen molecules within a functionally and morphologically intact ER (18). Having established that HSP47 interacts with correctly folded procollagen, we hypothesized that this interaction could lead to a stabilization of the procollagen triple helix, in particular regions of the helix with lower thermal stability. An alternative explanation for the binding of HSP47 to correctly folded procollagen is to prevent the lateral association of the chains occurring within the ER. It is known that once the procollagen...
molecule reaches the Golgi apparatus it forms higher order aggregates, which leads to distension of this organelle and could be a necessary intermediate prior to propeptide processing and formation of collagen fibrils (19). The formation of aggregates within the ER would prevent the transport of procollagen to the Golgi apparatus via vesicular transport; and hence the presence of HSP47 would be required to ensure efficient protein trafficking.

To investigate these hypotheses and to determine the subcellular specificity of HSP47 during procollagen folding and assembly, we determined its ability to interact with a variety of different procollagen chains with differing thermal stabilities in a well characterized semi-permeabilized cell system (20). Our results clearly demonstrate that HSP47 binds to triple helical molecules irrespective of their thermal stabilities and only to collagenous sequences containing at least one Gly-X-Tryp triplet. Such substrate recognition is consistent with a role in preventing the lateral association of procollagen chains within the ER lumen prior to their transport to the Golgi apparatus.

MATERIALS AND METHODS

Construction of Host Plasmid and Other Constructs—Construction of plasmid pC1(III)A1 has been described previously (21, 22). The host plasmid was made by constructing two intermediate plasmids. The first plasmid, pNEMG, was made by amplification of the N-terminal fragment of pC1(III)A1 using primers T3 and MB1 and cloned into the NotI and Smal sites into pGEMEasy (Promega, Southampton, UK). The second plasmid, pCSK, was made by conversion of the C-terminal fragment of pC1(III)A1 using primers MB2 and MB3 and subcloned into plBluescript SK+ (Stratagene, UK) via the Smal and SalI sites. The N-terminal NotI-Smal fragment was excised from pNEMG and ligated into pCSK plasmid restricted with NotI and Smal. The resulting host plasmid contains a Smal site that allows the insertion required DNA sequences. Other constructs were made by cloning “guest” DNA fragments, produced from annealed primer sequences, into the Smal site of the host plasmid. The first set of primers were ligated together to produce the guest insert fragment that was subsequently blunt-end ligated into the Smal site of the host vector. GXRGXP21 insert was made from oligonucleotides MBA, MBB, MBC, and MBD. The GPX2 insert made from MTP1, MTP2, MTP3, and MTP4. The GXR, GPX, GP2 insert was made with MTP5, MTP6, MTP3, and MTP4. Primer sequences available on request from the author.

Oligonucleotide Phosphorylation and Ligation—Complementary oligonucleotide primers were phosphorylated using T4 polynucleotide kinase (Roche Molecular Biochemicals) at 37 °C for 1 h. Annealing was carried out in annealing buffer (1 mM Tris-HCl, pH 7.5, 100 μM EDTA, 15 mM NaCl) by heating to 95 °C and slow cooling to 25 °C. Ligations were carried out using T4 DNA ligase (Promega).

Transcription in vitro—Transcription reactions were carried out as described previously (23). All recombinant plasmids were linearized with KpnI and transcribed using T3 RNA polymerase (Promega). Reactions were incubated at 37 °C for 4 h. Following purification over RNase columns (Qiagen, Dorking, UK), RNA was resuspended in 100 μl of RNase-free water containing 1 mM dithiothreitol and 40 units of RNasin (Promega).

Translation in Vitro—RNA was translated using a rabbit reticulocyte lysate (FlexiLysate, Promega) for 90 min at 30 °C. The translation reactions contained 16.5 μl of reticulocyte lysate, 0.25 μl of amino acids (minus methionine), 0.4 μl of 100 mM KCl, 0.25 μl of ascorbic acid (5 mg/ml), 15 μl of T[35S]methylion (Amersham Biosciences), 0.5 μl of transcribed RNA, and 5 μl (~2 × 1010) of semipermeabilized HT-1080 cells (SP cells) prepared as described (20). After translation, N-ethylmaleimide was added to a final concentration of 20 mM (in vitro) or to cross-linking with diathyrosuccinimidyl propionate (DSP); SP cells were solubilized by centrifugation in a microfuge at 10,000 × g for 5 min, and the pellet was resuspended in an appropriate buffer for subsequent enzymic digestion or gel electrophoresis.

Proteolytic Digestion—Isolated SP cells were resuspended in 20 μl of CTT buffer (1% (v/v) Triton X-100, 50 mM Tris-HCl, pH 7.4, containing 0.15 mM NaCl, 10 mM EDTA). Samples were then digested with a combination of chymotrypsin (250 μg/ml) and trypsin (100 μg/ml) (Sigma) for 5 min at room temperature. The reactions were stopped by the addition of soybean trypsin inhibitor (Sigma) to a final concentration of 500 μg/ml. Pepsin digests were carried out in the presence of 0.1 mM HCl. Acidified samples were incubated with pepsin (100 μg/ml) for 2 h at 30 °C or 16 h at 4 °C. The reactions were stopped by neutralization with Tris base (100 mM) and boiling in SDS-PAGE loading buffer.

Chemical Cross-linking—Translation products in SP cells were resuspended in 50 μl of KH1 (10 mM KOAc, 2 mM MgOAc, 20 mM HEPES, pH 7.2). Cross-linking was carried out using 1 mM DSP (stock 50 mM in Me2SO) for 10 min at 20 °C followed by a further 10-min incubation after the addition of 100 mM glycine to quench the DSP reaction. Cross-linked products were immunoprecipitated using the appropriate antibody.

Immunoprecipitation—Detailed protocols have been described previously (6). Briefly SP cells following cross-linking were isolated and solubilized in denaturing buffer (50 mM Tris-HCl, pH 7.4, 1% SDS, and 1% Nonidet P-40). Insoluble material was removed by centrifugation at 13,000 × g for 10 min. Supernatants were diluted in immunoprecipitation buffer (50 mM Tris-HCl, 0.15 mM NaCl, 2 mM EDTA, 1% (v/v) Triton X-100) to 1 ml. Samples were preincubated with 50 μl of protein A-Sepharose (10% (w/v) in phosphate-buffered saline) (Zymed Laboratories Inc., San Francisco, CA) for 30 min at 4 °C to remove protein A-binding components. Immunoprecipitation of cross-linked products was carried out using the appropriate antibody and 50 μl of protein A-Sepharose overnight at 4 °C. Antibodies to Myc peptide were from Calbiochem (Nottingham, UK) and antibodies against HSP47 were from Stressgen Biotechnologies Corp. Immunoprecipitated complexes were retrieved by centrifugation at 13,000 × g for 30 s and washed three times in immunoprecipitation buffer.

SDS-Polyacrylamide Gel Electrophoresis—Samples were resuspended in SDS-PAGE loading buffer (0.0625 M Tris-HCl, pH 6.8, SDS (2% w/v), glycerol (10% v/v), and bromphenol blue) in the presence of 50 mM dithiothreitol and boiled for 5 min. After electrophoresis gels were processed for autoradiography and exposed to Kodak X-Omat AR film or images quantified by phosphorimaging analysis.

RESULTS

We have previously suggested that one of the main functions of HSP47 may be to bind to the less stable regions of the collagen triple helix to prevent local unfolding especially during conditions such as heat shock. Given that the melting temperature of native fibrillar collagens is only a few degrees above body temperature, it seems natural to propose that stabilization of regions of the triple helical domain with limited thermal stability may be an important role for HSP47. One consequence of this theory is that we would expect HSP47 to bind specifically to regions of the triple helix with reduced thermal stability and to have low affinity for a triple helix that is composed of large numbers of stable Gly-X-Y triplelets. The basis for this specificity of binding could be the amino acid sequence or a particular conformation adopted by these regions of the triple helix. To test this theory we designed procollagen molecules with potentially thermal stable triple helices. To facilitate this construction we first prepared a host procollagen molecule that contains a substantial deletion within the triple helical domain of the proα1(III) chain preserving only six Gly-X-Y triplets from the N-terminal end and the last four triplets at the C-terminal end. This yields a very short collagenous domain mimicking that of collagen-like peptides but retaining the C- and N-propeptides and the telopeptides (Fig. 1). The cDNA construct was engineered with a strategically placed Smal site within the DNA sequence encoding the short triple helical domain so that we would be able to place DNA fragments into the Smal site. In this way we initially produced two larger constructs, one a 32 amino acid mixture of highly stable triplets including GXRGXP21 (5) and one containing just GPX triplets (6). The GPR triplets were chosen because of its highly stabilizing effect on the triple helix (24) and because it is unaffected by the in vitro hydroxylation process. Thus we could compare binding of HSP47 to two helices that we would predict would be highly thermal stable, one of which has a high proportion of hydroxyproline residues and the other having a mixture of hydroxyproline and arginine residues within the Y-
position of the Gly-X-Y triplet. Several methionine residues were also placed in the X position of both helices to allow radiolabeling and detection of the triple helical domain.

We initially analyzed the synthesis of the host, the GXR₆GXP₁₅, and the GXP₂₂ procollagen constructs to determine whether they had become hydroxylated and assembled to form triple helical molecules within our SP cell translation system. Transcripts encoding the constructs were translated in the presence of SP cells for 90 min, and the products were isolated and separated by SDS-PAGE under both reducing and nonreducing conditions (Fig. 2, A and B). Under reducing conditions the products migrated according to the expected sizes for the host, GXR₆GXP₁₅, and GXP₂₂ procollagen molecules (Fig. 2, A, lane 1, and B, lanes 1 and 2). Under nonreducing conditions the host and GXP₂₂ procollagen translation products migrated as higher molecular weight species corresponding to an interchain disulfide-bonded trimer (Fig. 2, A, lane 2, and B, lane 4). The GXR₆GXP₁₅ translation product migrated either with a higher molecular weight, although the polypeptide migrated as a smear rather than a sharp band (panel B, lane 3), or as an intrachain disulfide bonded species. When translations were carried out in the presence of an inhibitor of prolyl 4-hydroxylase (α,α’-dipryridyl), translation products migrated with a faster mobility indicating that all chains become modified by proline hydroxylation (data not shown). These results demonstrate that all of the procollagen chains are hydroxylated and can assemble into trimeric molecules within SP cells. However, the GXR₆GXP₁₅ molecule either adopted a disulfide-bonded trimer that did not migrate as a single species, probably because of the retention of some secondary structure even in the presence of SDS, or as a non-interchain disulfide-bonded species.

To assess further the ability of GXP₂₂ and GXR₆GXP₁₅ to form fully folded triple helical molecules, we carried out translations in the presence and absence of α,α’-dipryridyl and treated the translation products with proteases prior to separation by SDS-PAGE (Fig. 2, C and D). Following proteolytic digestion of the GXR₆GXP₁₅ translation product, a protease resistant fragment of 15 kDa remained (Fig. 2C, lane 3) indicative of the formation of a compact triple helical domain. We have shown in previous studies that translation products of procollagen molecules are completely susceptible to proteolytic degradation when hydroxylation is inhibited with α,α’-dipryridyl, as is the case with a truncated version of type III procollagen (Fig. 2C, lanes 1 and 2). In contrast, the GXR₆GXP₁₅ translation products, formed in the presence of α,α’-dipryridyl, were resistant to proteolysis, producing a band consistent with that expected for an unhydroxylated triple helical domain (panel C, lane 4).

Thermal denaturation studies revealed that the $T_m$ of GXR₆GXP₁₅ triple helical product was higher than 55 °C (data not shown). Following proteolytic digestion of the GXP₂₂ translation products synthesized in the absence of added α,α’-dipryridyl, protease resistant fragments of 15 and 30 kDa, along with higher molecular mass species, were observed (panel D, lane 1). When α,α’-dipryridyl was included in the translations, only the 15- and 30-kDa protease-resistant products were observed (panel D, lane 2). The 15-kDa protease-resistant fragment is of the expected molecular mass for the triple helical domain, and it is likely that the higher molecular mass species represent multimers of this fragment that are particularly resistant to denaturation even after boiling in the presence of reducing agent and SDS. These fragments would be particularly stable after proline hydroxylation. Thus, the presence of high stability triplets such as GPR and GXP appears to confer high thermal stability upon the triple helical domain. The presence of con-

![Fig. 1. Diagrammatic representation of recombinant procollagen chains.](image)

![Fig. 2. Folding of recombinant procollagen molecules.](image)
sequent GXP triplets is particularly stabilizing, giving rise to triple helical domains that are resistant to denaturation even in the presence of SDS. The GXR,GXP₁₅ and GXP₂₂ chains are also able to assemble into triple helical molecules even in the absence of proline hydroxylation. It is also highly likely that the host procollagen molecule was also able to form a triple helix, although we were unable to determine this as there are no methionine residues in this molecule and therefore the triple helical domain could not be radiolabeled.

Does HSP47 Bind to Thermally Stable Procollagen Molecules?—To test whether HSP47 binds only to regions of instability within the procollagen triple helix, we determined whether it could bind to the very stable triple helical domains described above. Transcripts encoding the host procollagen molecule and the GXR,GXP₁₅ and GXP₂₂ chains were translated in the presence of SP cells, and potential interactions were stabilized using the chemical cross-linking agent DSP. Cross-linked products were immunoprecipitated with antibodies raised against HSP47 (data not shown). We have demonstrated above that the host procollagen and GXP₂₂ are hydroxylated and that GXP₂₂ forms a triple helix. These results suggest that binding of HSP47 to procollagen triple helical domains within the ER requires more than simply a triple helical domain and that the binding maybe reliant upon a specific amino acid sequence. This negative result also underlines the fact that HSP47 interacts with the triple helical domain as the C- and N-propeptides from these molecules are identical to the GXR,GXP₁₅ construct.

Recognition of Gly-X-R Triplet by HSP47—Comparisons of the amino acid sequences of the triple helical domain within the host procollagen chain, the GXR,GXP₁₅, and the GXP₂₂ reveal that the major differences between the chains lie with the presence of arginine in the Y position. To ascertain whether Gly-X-Arg promotes the interaction of HSP47 with procollagen within a cellular context, we replaced the proline in the Y position of a Gly-Pro-Pro triplet in the triple helical domain in GXP₂₂ with an arginine residue (Fig. 1). The new construct, GXR,GXP₂₁, was translated in the presence and absence of α,α′-dipyridyl and analyzed to assess folding and association with HSP47. The GXR,GXP₂₁ construct was capable of forming a disulfide-bonded trimer as demonstrated when the translation products were separated under nonreducing conditions (Fig. 4A). Separation of the protease-digested products under reducing condition produced a pattern similar to that seen with the GXP₂₂ construct (Fig. 4B). Even in the presence of α,α′-dipyridyl, the synthesized translation product formed a protease-resistant fragment that was also resistant to complete denaturation with SDS. The protease-resistant fragments migrate as clearer distinct bands when separated under nonreducing conditions confirming that the GXR,GXP₂₁ construct forms a correctly aligned triple helix that is interchain disulfide bonded and that the absence of hydroxylation does not prevent this triple helix from forming (Fig. 4C). Having determined that GXR,GXP₂₁ forms triple helical molecules irrespective of hydroxylation, we carried out cross-linking experiments to establish whether these molecules were also substrates for HSP47. The hydroxylated products associate with HSP47 (Fig. 5, lane 1). Under conditions where hydroxylation was prevented we also observed an interaction with HSP47 (lane 2).
The formation of the triple helix conformation must be considered to be one of the contributing factors required to allow HSP47 interaction. Studies of collagen model peptides (13) and in vitro binding studies (26) also confirm the binding of HSP47 to the triple helix. The fact that HSP47 bound equally well to the unhydroxylated and hydroxylated triple helix in our studies indicates that the 4-hydroxyl group does not play a role in the recognition process. It is generally accepted that HSP47-procollagen association is not limited to the ER but also occurs in the intermediate compartment and perhaps in the cis-Golgi (19, 27, 28). If hydroxylation were to exert a negative influence on HSP47 association, then very little association would occur even in the ER because hydroxylation tends to occur rapidly following translation. Thus the observation indicating that HSP47 has reduced affinity for hydroxylated peptides in vitro (12) is not supported by in vivo data.

Using triple helical domains consisting of a (Gly-X-Pro)n repeat sequence, we have shown for the first time that this type of sequence is not a substrate for HSP47 in vitro. The failure of the (GXP)n-type triple helix to interact with HSP47 under hydroxylation or nonhydroxylation conditions shows that (GXP)n triple helix either does not possess the recognition signal for HSP47 binding or forms an aggregate that cannot be recognized. In support of the latter possibility we have observed that upon treatment with proteases the collagen domains form SDS-stable aggregates. This ability is considerably reduced if hydroxylation is prevented using a,a′-dipyridyl, indicating that the 4-hydroxyl group exacerbates the aggregation effect. However, the fact that HSP47 did not interact even with the unhydroxylated triple helix shows that a recognition factor for HSP47 binding is absent. We showed that a single Gly-X-Arg triplet is sufficient to allow HSP47 to recognize the triple helix of GXR,GXP31. The GXR triplet therefore serves as the HSP47-binding site on the triple helix as well as contributing toward the thermal stability and flexibility of the procollagen molecule. A recent in vitro study using collagen host-guest peptides confirms this finding and also shows that the Arg residue must be located in the Y position of a GXY triplet for binding of HSP47 to occur. Furthermore, the binding is specific to the side chain of arginine (16).

Studies of the small naturally occurring collagens that form the capsule wall of nematocytes support our observations on the enhanced ability of (GXP)n triple helical sequences to aggregate. The (GXP)n-type triple helices of these Hydra minicollagens exhibit lateral aggregation to form superstructures (29). This observation along with the inhibition of fibrillogenesis by HSP47 (30) is possibly an important clue to understanding the function of HSP47. If uncontrolled aggregation were to occur in the environment of the ER, it would interfere with the efficient expression of other proteins. The HSP47 molecule might therefore function in limiting or controlling the aggregation and perhaps even in targeting the procollagen aggregates toward a particular transport mechanism. Current evidence points to transport of procollagen cargo occurring in the form of large-scale aggregate structures by a mechanism involving progressive maturation of Golgi cisternae (19, 28, 31). The dissociation mechanism is likely to exploit the exquisite pH sensitivity of HSP47 (32) that leads to dissociation of HSP47 from procollagen below pH 6.3 probably in the more acidic environment of the cis-Golgi (27). Evidence to support the theory that HSP47 modulates collagen aggregation is also available from studies on pN-collagen type III. This collagen has a marked tendency to aggregate even under high salt conditions. However, aggregation appears to be suppressed in the presence of HSP47 (33).

In summary we propose that HSP47 combines with procoll-
lumen molecules to control and limit lateral aggregation in the early secretory pathway. The trimeric form of HSP47 observed in other studies (15, 30) may provide a clue as to how the limitation actually works. The formation of a “ring” of HSP47 around a small bundle (for example a trimer) of procollagen triple helices is one mechanism that may prevent premature large-scale lateral aggregation. Subsequent pH induced conformational changes in HSP47 in the cis-Golgi would then provide the mechanism for HSP47 dissociation and formation of pro-collagen bundles as a precursor to the formation of collagen fibrils during the later stages of the secretory pathway.

Acknowledgments—We thank Dr. Margaret Batten for assistance in construction of the “host” plasmid.

REFERENCES

Sequence-specific Recognition of Collagen Triple Helices by the Collagen-specific Molecular Chaperone HSP47
Mohammed Tasab, Lynsey Jenkinson and Neil J. Bulleid

doi: 10.1074/jbc.M202782200 originally published online July 11, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M202782200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 33 references, 21 of which can be accessed free at http://www.jbc.org/content/277/38/35007.full.html#ref-list-1