Conformational Requirements of Collagenous Peptides for Recognition by the Chaperone Protein HSP47*

Received for publication, April 10, 2000, and in revised form, June 16, 2000
Published, JBC Papers in Press, June 20, 2000, DOI 10.1074/jbc.M003026200

Takaki Koide†‡§, Akiko Aso†‡, Tetuya Yorihizu†‡§, and Kazuhiro Nagata†‡§**

From the †Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST) and the ‡Department of Molecular and Cellular Biology, Institute for Frontier Medical Sciences, Kyoto University, Sakyo-ku, Kyoto 606-8397, Japan

The collagen binding chaperone HSP47 interacts with procollagen in the endoplasmic reticulum and plays a crucial role in the biosynthesis of collagen. We recently demonstrated that typical collagen model peptides, (Pro-Pro-Gly)n, possess sufficient structural information for interaction with HSP47 (Koide, T., Asada, S., and Nagata, K. (1999) J. Biol. Chem. 274, 34523–34526). Here, we show that binding of (Gly-Pro-Pro)n peptides to HSP47 can be detected using the two-hybrid system in yeast if a trimerizing domain is fused to the C termini of the peptides. Some peptides interacted with HSP47 at a lowered assay temperature at 24 °C but not at 30 °C, indicating the importance of conformational change of the substrate peptides. To analyze the spectrum of HSP47 substrate sequences, we performed two-hybrid screening of collagen-like peptides in designed random peptide libraries using HSP47 as a bait. In selected peptides, the enrichment ratio calculated for each amino acid residue correlated strongly with the contribution of the residue to triple-helix stability independently determined using synthetic collagen model peptides. Taken together, our results suggest that HSP47 preferentially recognizes collagenous Gly-X-Y repeats in triple-helical conformation. We also demonstrated that screening of combinatorial peptide libraries is a powerful strategy to determine conformational requirements as well as the elucidation of binding motifs in primary structure.

HSP47 is an endoplasmic reticulum (ER)1 resident stress protein, which is thought to function as a collagen-specific molecular chaperone. This protein associates with procollagen during its folding and/or post-translational modification in the ER (1–3). Recent studies have revealed that HSP47 plays a critical role in collagen biosynthesis; hsp47 null mice show abnormal collagen synthesis and die before E11.5.2 Because HSP47 binds in vitro to various types of collagen (at least types I-V; Ref. 4), as well as to collagen-like proteins such as C1q,3 it appears that HSP47 may function through specific binding to the helix-forming portions of procollagen. However, the function of HSP47 at the molecular level remains unclear. In a previous report, we showed that typical collagen model peptides ((Pro-Pro-Gly)n, n ≥ 7) were recognized by HSP47 in an in vitro binding assay (5). The strength of this interaction increased with increasing length of the model peptides, and the interaction was negatively regulated by prolyl 4-hydroxylation at the second Pro residues of the triplets.

In immature procollagen α-chains (prepro α-chains), Pro is the most common amino acid residue in the X and Y positions of Gly-X-Y triplets, but a variety of residues other than Pro are also found in both positions (6). The longest uninterrupted series of (Gly-Pro-Pro)n repeats in a preprotocollagen is the five-repeat series (n = 5) found in the human, bovine, chick, rat, and mouse α1(I) chain. And the longest (Pro-Pro-Gly)n repeats are also the five repeats in the human α2(I) chain. The initial aims of the present study were to elucidate amino acid preferences in HSP47 substrate peptides and to try to identify a putative consensus sequence for HSP47 binding using the yeast two-hybrid system to screen collagenous sequences.

We developed a method of detecting specific interaction between HSP47 and collagen model peptides using the yeast two-hybrid system. Using this method, we analyzed the effects of model (Gly-Pro-Pro)n substrate length and assay temperature on interaction with HSP47. We also investigated the effect of amino acid replacement in the (Gly-Pro-Pro)n peptides as a function of assay temperature. Finally, we constructed two peptide libraries containing diverse collagen-like sequences and screened these libraries for HSP47 binding sequences using the yeast two-hybrid system. The data obtained in this study highlighted the conformational basis of HSP47 substrate preference rather than identifying a consensus sequence for HSP47 binding.

**EXPERIMENTAL PROCEDURES**

General—Restriction enzymes, Taq polymerase (Ex Taq), DNA ligase, and other DNA-modifying enzymes were purchased from Takara Shuzo Co. Ltd. (Kusatsu, Japan). DNA fragments were synthesized by Hokkaido System Science (Sapporo, Japan). Chemicals were purchased from either Nacalai Tesque (Kyoto, Japan) or Wako Pure Chemical Co. (Osaka, Japan). DNA sequencing was carried out on an ABI PRISM 377A sequencer (Perkin-Elmer).

Plasmid Constructs—The identities and orientations of the inserts of all plasmid constructs, except for those in random peptide libraries,

---

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

† Present address: Dept. of Biological Science and Technology, Faculty of Engineering, The University of Tokushima, Tokushima 770-8506, Japan.

‡ These authors contributed equally to this work.

** To whom correspondence should be addressed. Tel.: 81-75-751-3848; Fax: 81-75-751-4645; E-mail: nagata@frontier.kyoto-u.ac.jp.

† The abbreviations used are: ER, endoplasmic reticulum; PCR, polymerase chain reaction; kb, kilobase; AD, activation domain; GST, glutathione S-transferase; PAGE, polyacrylamide gel electrophoresis; BD, DNA binding domain; LB medium, Luria-Bertani medium; Hyp, 4-hydroxyproline.

‡ N. Nagai, M. Hosokawa, S. Itohara, T. Matsushita, Y. Ninomiya, Y. Sado, E. Adachi, N. Hosokawa, and K. Nagata, submitted for publication.

§ T. Koide, unpublished observation.
To construct pYL94 encoding the GAL4 DNA binding domain fused to the C1q-like domain of the K100 gene (GenBank accession number AB044560) was amplified by PCR with the primers YOR68–6 (5′-TTA GGA TCC CAT CTT-3′) and YOR68–7R (5′-GGG GAA TTC TCA GTC AGC ATA AAT AAT AAA-3′). The resulting 0.5-kb fragment was digested with BamHI and EcoRI and subcloned into the BamHI and EcoRI sites of pAS2−1 (CLONTECH) to yield pYL94.

To construct pYL97 encoding the GAL4-activation domain (AD) of the HSP47, the mouse hsp47 gene (7) was amplified by PCR with the primers Eco-pYL68 (5′-GGG GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT-3′) and pYL68-Xho (5′-TTC ACC GGG CCA GCC GGT GCA CCC GGA CCT-3′) and subcloned into the BamHI and EcoRI sites of pACT2 (CLONTECH) to yield pYL97.

**TABLE I**

| Entry | Construct | Synthetic DNA sequence (5′→3′) |
|-------|-----------|---------------------------------|
| 1     | (GPP)5    | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 2     | (GPP)6    | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 3     | (GPP)7    | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 4     | (GPP)8    | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 5     | (GPP)9    | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 6     | (GPP)-K100-C1qD | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 7     | (GPP/-K100-C1qD | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 8     | (GPP)-K100-C1qD | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 9     | (GPP)-K100-C1qD | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 10    | GAP1      | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 11    | GAP2      | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 12    | GAP3      | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 13    | GAP2      | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 14    | X librarya | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |
| 15    | Y librarya | G GAA ACA GCT ATG ACC ATG GAT CCC GGC CCA CCT |

**a** The following mixtures were used at the x, y, and z positions: x = G (25%), C (25%), A (25%), T (25%); y = G (19%), C (19%), A (31%), T (31%); z = G (39%), C (39%), A (0%), T (22%) (8).

**Note:** The substrate recognition by HSP47 was confirmed by sequencing.

To generate pYL94 encoding the GAL4 DNA binding domain fused to HSP47, the mouse *hsp47* gene (7) was amplified by PCR with the 5′-primer YOR1D (5′-ACC GGA TCC CCG CCG AGG TGA AGA AAC C-3′) and the 3′-primer YOR2M (5′-TTG GAA TTC TCA CAA CTC ATC-3′). The resulting 1.2-kb fragment was subcloned into the BamHI and EcoRI sites of pAS2−1 (CLONTECH) to yield pYL94.

To construct pYL97 encoding GAL4AD-(Gly-Pro-Pro)\textsubscript{n} (n = 5–9), synthetic single-stranded DNA (Table I, entries 1–5) was used as a template for PCR with the primers RI-P1 (5′-5′ AAT TCA CTG GCC GTC GTT TTA C-3′) and Xho-R1 (5′-5′ GGA AAC AGC TAT GAC GGA CCG CCT GGG CCA CCT GGA CCT-3′). The resulting fragments were inserted into the BamHI and EcoRI sites of pAS2−1 (CLONTECH) to yield pYL97.

Plasmids encoding GAL4AD-(Gly-Pro-Pro)\textsubscript{n}, K100 C1q-like domain were constructed as follows. A 0.5-kb DNA fragment was digested with *Bam*HI and *Eco*RI and subcloned into the *Bam*HI and *Eco*RI sites of pAS2−1 (CLONTECH) to yield pYL97. To construct plasmids encoding GALAD-(Gly-Pro-Pro)\textsubscript{n} (n = 5–9), synthetic single-stranded DNA (Table I, entries 1–5) was used as a template for PCR with the primers RI-P1 (5′-5′ AAT TCA CTG GCC GTC GTT TTA C-3′) and Xho-R1 (5′-5′ GGA AAC AGC TAT GAC GGA CCG CCT GGG CCA CCT GGA CCT-3′). The resulting fragments were inserted into the *Bam*HI and *Eco*RI sites of pAS2−1 (CLONTECH) to yield pYL97. To construct plasmids encoding GALAD-(Gly-Pro-Pro)\textsubscript{n}, K100 C1q-like domain were constructed as follows. A 0.5-kb DNA fragment was digested with *Bam*HI and *Eco*RI and subcloned into the *Bam*HI and *Eco*RI sites of pAS2−1 (CLONTECH) to yield pYL97. To construct plasmids encoding GALAD-(Gly-Pro-Pro)\textsubscript{n}, K100 C1q-like domain were constructed as follows. A 0.5-kb DNA fragment was digested with *Bam*HI and *Eco*RI and subcloned into the *Bam*HI and *Eco*RI sites of pAS2−1 (CLONTECH) to yield pYL97.
and subcloned into the BamHI and EcoRI sites of pGAD-C1q. Plasmids encoding GAP1, GAP2, GPA1, and GPA2 peptides (see Fig. 3) were constructed by the same procedure using the corresponding synthetic inserts (Table I, entries 10–13). To generate pYL103 for bacterial expression of a GST fusion protein of the K100 C1q-like domain, pYL97 was digested with BamHI and EcoRI, and the resulting 0.5 kb fragment was subcloned into the BamHI and EcoRI sites of pGEX-3X (Amersham Pharmacia Biotech) to yield pYL103.

The inserts for the X and Y libraries (Table I, entries 14 and 15) were chemically synthesized as single-stranded DNA fragments. To minimize bias in the amino acid distribution, the optimized mixtures of nucleotide derivatives developed by Cwirla et al. (8) were used at the degenerate positions. The libraries were constructed by a procedure similar to the construction of the plasmids encoding the GAL4AD-(Gly-Pro-Pro)n-K100 C1q-like domain.

**Yeast Culture and Transformation**—The yeast strain CG-1945 (MATa, ura3–52, his3–200, ade2–101, lys2–801, trp1–901, leu2–3, 112, gal4–52, gal80–538, cyh2, lys2::GAL4AD, GAL1-TATA, HIS3, URA3::GAL1::CYC1::TATA, LacZ) was used for two-hybrid screening and was grown in either yeast-peptone-dextrose or synthetic dropout medium containing appropriate mixtures of amino acids and nucleosides. Introduction of plasmids into yeast cells was performed using YEASTMAKER Yeast Transformation System (CLONTECH) according to the manufacturer's instructions.

**Yeast Two-hybrid Binding Assays**—Two-hybrid binding assays to detect interaction between HSP47 in the binding hybrid and peptides in the activation hybrids were performed according to the instructions for the MATCHMAKER two-hybrid kit (CLONTECH). A human lamin C fragment in pAS2–1 (pLAM5–1, CLONTECH) was used as a negative control.

**Library Screening**—Two-hybrid screening of random collagenous peptide libraries was performed at a culture temperature of 30 °C according to the instructions for the MATCHMAKER two-hybrid kit (CLONTECH), using HSP47 (pYL94) as a bait. Initial screening was carried out using histidine auxotrophic assays on agar plates, and the resulting positive clones were further screened using β-galactosidase assays on nitrocellulose membranes.

**Western Blotting of Yeast Cell Lysates**—Yeast cells harvested from 5 ml of liquid culture were frozen in liquid nitrogen and thawed in 100 μl of 10 mM Tris-HCl (pH 7.5) containing 1 mM phenylmethanesulfonyl fluoride and 0.2% Nonidet P-40. The cells were lysed by vortexing with glass beads (Sigma) at 4 °C for 30 min. After centrifugation, the supernatant was collected, and the protein concentration was determined using a protein assay kit (Bio-Rad). Lysate proteins (10 μg) were separated by SDS-PAGE, and transferred to nitrocellulose membranes.

To detect the activation hybrid proteins, anti-GAL4 monoclonal antibody (CLONTECH) was used as a primary antibody, and immuno-reactive bands were visualized using an ECL kit (Amersham Pharmacia Biotech).

**Expression, Purification, and Chemical Cross-linking of Recombinant K100 C1q-like Domain**—Escherichia coli JM109 cells carrying pYL103 were grown in 5 ml of LB medium containing 50 μg/ml ampicillin, 30 μg/ml kanamycin, and 50 μg/ml chloramphenicol at 37 °C. The expression of the recombinant protein was induced by adding isopropyl-β-D-thiogalactopyranoside (0.5 mM), and the culture was continued for 1 h. Cell lysates were prepared as described elsewhere (9). GST-K100 C1q-like domain fusion protein in the lysate was adsorbed onto 0.3 ml of glutathione-Sepharose 4B (Amersham Pharmacia Biotech), and the beads were washed four times with 10 ml of 50 mM Hepes-Na (pH 7.5), 3.7 mM EDTA, 0.4 mM NaCl. The fusion protein on the beads was cleaved by overnight treatment with 2 units of Factor Xa (Amersham Pharmacia Biotech), and the resulting 0.5 kb-fragment was subcloned into the BamHI and EcoRI sites of pGEX-3X (Amersham Pharmacia Biotech).

**Substrate Recognition by HSP47**—To detect the specific binding of recombinant GST-HSP47 fusion protein to immobilized synthetic (Pro-Pro-Pro)n peptides when n was not less than 7 and showed that prolyl 4-hydroxylation (which converts Pro residues to 4-hydroxyprolyl (Hyp) residues) at the second Pro residue of each triplet has a negative effect on the interaction. As yeast cytosol is unlikely to contain prolyl 4-hydroxylase activity, a similar binding assay was performed using the yeast two-hybrid system. A plasmid encoding mouse HSP47 fused to the GAL4 DNA binding domain (GAL4BD) was transfected into yeast cells harboring HIS3 and LacZ genes in GAL4-driven reporter constructs. Synthetic DNA fragments containing various numbers of Gly-Pro-Pro repeats were fused to the gene encoding the GAL4 activation domain (GAL4AD), and these plasmids were also introduced into the yeast cells. The binding of the hybrid proteins expressed in the yeast cells was assayed by growth on histidine-depleted agar plates and by β-galactosidase assay on nitrocellulose membranes. In this system, no interaction was detected between HSP47 and peptides containing 5–10 repeats of (Gly-Pro-Pro)n (Fig. 1A), in contrast to the results of the in vitro solid-phase binding assay using synthetic (Pro-Pro-Gly)n peptides (5).

In parallel with the experiments above, one of the authors (T. Yorihuzi) cloned a gene encoding a novel C1q-like protein, tentatively named K100 protein, from an E17 mouse embryo cDNA library by two-hybrid screening using HSP47 as a bait. The open reading frame of the gene encodes an N-terminal signal sequence and 17 repeats of Gly-X-Y followed by a C-terminal globular domain. The globular domain is equivalent to the so-called C1q module (10, 11) and shows 28 and 34% amino acid identity with the corresponding domains of mouse C1q-B.
chain and type X collagen, respectively. Interaction between HSP47 and K100 protein was detected in the two-hybrid system only when the C1q-like domain was present at the C terminus of the Gly-X-Y repeat sequence of K100, although the C1q-like domain itself did not interact with HSP47 (data not shown). We therefore fused the C1q-like domain to the C terminus of the (Gly-Pro-Pro)ₙ peptides in the activation hybrids and tested for binding to HSP47. Specific interaction was detected when the number of triplet repeats fused to the C1q-like domain was greater than 7 (Fig. IA). This result demonstrates that the interaction between HSP47 and (Gly-Pro-Pro)ₙ peptides can be detected in the yeast two-hybrid system. In the presence of the K100 C1q-like domain, the strength of the interaction depended on the length of the peptides in a manner consistent with our previous observations based on the in vitro binding assay (5).

Next, we wished to clarify the role of the K100 C1q-like domain in HSP47-peptide interactions in the two-hybrid system. To determine whether the K100 C1q-like domain simply stabilized the activation hybrids by preventing enzymatic degradation in the cells, we examined the cellular protein levels of the activation hybrids by Western blotting using anti-GAL4AD antibody. The protein levels of the activation hybrids remained relatively constant, regardless of the presence or absence of the K100 C1q-like domain (Fig. 1B). The C1q-like domain of type X collagen is reported to form a very stable trimer that remains intact even in SDS-containing buffers (12–14). We looked for the formation of such a stable trimer by the K100 C1q-like domain using unboiled SDS-PAGE samples of the yeast lysate but did not detect any bands corresponding to the expected molecular mass of a trimer (Fig. 1B).

We further investigated the possible oligomer-forming properties of the K100 C1q-like domain using a more sensitive method. The recombinant K100 C1q-like domain was expressed in E. coli cells, purified, and cross-linked in solution using glutaraldehyde, and the products of the cross-linking reaction were analyzed by SDS-PAGE. The major cross-linked products migrated at positions corresponding to dimeric and monomeric proteins. Molecular sizes are shown in kDa. The monomer, dimer, and trimer of the recombinant K100 C1q-like domain are indicated by asterisks. B, untreated and cross-linked recombinant K100 C1q-like domain was analyzed by SDS-PAGE on 12% gels followed by Western blotting using rabbit antiserum raised against recombinant K100 C1q-like domain.

Effect of Assay Temperature and Amino Acid Replacement on the Interaction of (Gly-Pro-Pro)ₙ Peptides with HSP47 in the Yeast Two-hybrid System—The thermal stability of collage-

4 T. Yorihuzi and K. Nagata, manuscript in preparation.
and that any five (Gly-Pro-Pro)\textsubscript{n} peptides of different lengths. Interaction in the yeast screen was estimated to be 6.05 residues are conserved in all types of collagen. If all of the residues at every third position were also fixed, because the Gly peptide libraries were fixed to 24-mers (encoding 8 triplets). The Gly selection. Because (Gly-Pro-Pro)\textsubscript{8} combined with a C-terminal domain nature of the amino acid residues present in the degenerate (Gly-X-Pro)\textsubscript{n} peptides (23). Thus, residues at the X and Y positions in the Gly-X-Y repeats appear to contribute differently to either the triple helix stability or the formation of the HSP47 binding surface. Because GAP1 interacted with HSP47 at 24 °C (Fig. 3B), most of the peptides in the X library appear to have failed to form triple-helical structures under the screening conditions at 30 °C.

To exclude any possible experimental bias that might have affected the selection, we compared the amounts of the activation hybrids expressed in the HSP47-selected yeast clones with those in randomly chosen negative clones. Protein levels estimated by immunoblotting using anti-GAL4AD antibody were not significantly different between the clones examined, regardless of the HSP47 binding activity of the peptides (Fig. 5). We therefore conclude that the result of two-hybrid selection may be attributed directly to the HSP47 binding property of the substrate peptides.

All of the peptide sequences deduced from the DNA sequences of the selected clones are shown in Fig. 6. We could not identify any characteristic order in the primary amino acid sequences of the individual peptides. Most of the amino acid residues were almost uniformly distributed throughout the randomized Y positions of HSP47 binding peptides. Arginine residues, which have been reported to possess a triple-helix stabilizing effect comparable to that of Hyp residues (24), were most strongly enriched (up to 6-fold) in the HSP47 binding peptides relative to basal values taken from 200 randomly chosen clones. Proline residues were also highly enriched (about 5-fold). On the other hand, Asp, Phe, Gly, Asn, and Trp residues were apparently excluded (less than 20% of the basal frequency, Fig. 7).

Brodsky and co-workers (6, 17, 24–26) have studied the contribution of various amino acid residues to the stability of triple-helical structures using sets of designed collagen models such as acetyl-(Gly-Pro-Hyp)\textsubscript{4}-Gly-X-Y-(Gly-Pro-Hyp)\textsubscript{4}-Gly-Gly-amide. Surprisingly, the enrichment ratios for amino acid residues at the Y position of our model peptides correlated strongly with the thermal stabilities of corresponding collagen model peptides reported in the literature (Fig. 8). This correlation clearly demonstrates the importance of triple-helical structure in substrate recognition by HSP47. Selection by HSP47 binding from the Y library appears to be attributable to the

Fig. 4. Design of collagenous peptide libraries. Plasmids in the X and Y libraries were designed to express activation hybrid proteins containing random mixtures of all possible amino acid residues in the X\textsubscript{n} (blue) and Y\textsubscript{n} (red) positions.
Concluding Remarks—In this study, we have focused on the molecular determinants of HSP47 substrate recognition. In this study, we established a yeast two-hybrid assay system to study the interaction between HSP47 and substrate peptides. Using this assay, we have obtained the following evidence that HSP47 preferentially recognizes triple-helical peptides as binding substrates: 1) a trimerizing domain, such as the K100 C1q-like domain, fused to the peptides was required for detectable interaction (Figs. 1 and 2); 2) some peptides were bound by HSP47 at 24 °C but not at 30 °C, suggesting that the melting temperature of the triple helix may be an important factor in binding (Fig. 3); and 3) amino acid residues, which stabilize the triple helix, such as Arg and Pro, were enriched in the group of peptides selected from the Y library and vice versa (Figs. 6–8).

Although the real function of HSP47 is still ambiguous, the studies on the substrate recognition lead us to some working hypotheses. In the previous paper, we showed that HSP47 prefers Pro residues at the Y positions rather than Hyp in the context of the (Pro-Y-Gly)n sequence (5). Combined with the result shown in this paper, it seems that HSP47 retains triple helical procollagen molecules having less prolyl 4-hydroxylated portions in the ER. Prolyl 4-hydroxylase is also reported to bind to a less prolyl 4-hydroxylated form of single chain procollagen (27). These mechanisms may be expected to ensure the quality of procollagen molecules to be secreted, although it is not known whether HSP47 can exert a quality control mechanism alone or in cooperation with prolyl 4-hydroxylase. It is not clear whether HSP47 exclusively binds to triple helical portions containing Gly-Pro-Pro triplets or other triplets such as Gly-Pro-Arg can be accommodated to the HSP47-binding sites, because all peptides designed for two hybrid screening contain known HSP47 binding sequences such as Gly-Pro-Pro-Gly and Gly-Pro-Pro at both ends of the randomized sites (Fig. 4). If the sequences other than (Gly-Pro-Pro)n can form HSP47-binding sites, our finding that HSP47 prefers triple helical substrates suggests another role of HSP47; HSP47 might facilitate procollagen folding in the ER by stabilizing partially folded triple helical intermediates that would be otherwise unstable at body temperature. Further biochemical and physicochemical studies using individual model peptides would clarify the molecular function of HSP47 as collagen-specific chaperone.

Although we initially undertook two-hybrid screening of the peptide libraries with a view to identifying HSP47 binding motifs in the primary structure, no such motifs were apparent from our results. Instead, we have obtained information re-
garding the preferred secondary structure of HSP47 substrates (Figs. 6–8). It is of note that this is the first case, to our knowledge, in which biological selection from combinatorial libraries has provided information about intermolecular conformation. We believe that the similar two-hybrid method will be a powerful tool to elucidate the substrate conformation of other collagen-binding proteins.

Acknowledgment—We are grateful to Dr. Konrad Beck for valuable discussion.

REFERENCES
1. Nakai, A., Satoh, M., Hirayoshi, K., and Nagata, K. (1992) J. Cell Biol. 117, 903–914
2. Satoh, M., Hirayoshi, K., Yokota, S., Hosokawa, N., and Nagata, K. (1996) J. Cell Biol. 133, 469–483
3. Nagata, K. (1996) Trends Biochem. Sci. 21, 23–26
4. Natsume, T., Kide, T., Yokota, S., Hirayoshi, K., and Nagata, K. (1994) J. Biol. Chem. 269, 31224–31228
5. Kido, T., Asada, S., and Nagata, K. (1999) J. Biol. Chem. 274, 34523–34526
6. Ramshaw, J. A. M., Shah, N. K., and Brodsky, B. (1998) J. Struct. Biol. 122, 86–91
7. Takechi, H., Hirayoshi, K., Nakai, A., Kudo, H., Saga, S., and Nagata, K. (1992) Eur. J. Biochem. 206, 323–329
8. Cwirla, S. E., Peters, E. A., Barrett, R. W., and Dower, W. J. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 6378–6382
9. Asada, S., Kido, T., Yasui, H., and Nagata, K. (1999) Cell Struct. Funct. 24, 187–196
10. Bork, P., and Koonin, E. V. (1996) Curr. Opin. Struct. Biol. 6, 366–376
11. Hegyi, H., and Bork, P. (1997) J. Protein Chem. 16, 545–551
12. Frischholz, S., Beier, F., Girkontaite, I., Wagner, K., Poschl, E., Turnay, J., Mayer, U., and von der Mark, K. (1998) J. Biol. Chem. 273, 4547–4555
13. Marks, D. S., Gregory, C. A., Wallis, G. A., Brass, A., Kadler, K. E., and Boot-Handford, R. P. (1999) J. Biol. Chem. 274, 3632–3641
14. Zhang, Y., and Chen, Q. (1999) J. Biol. Chem. 274, 22409–22413
15. Maru, Y., Afar, D. E., Witte, O. N., and Shibuya, M. (1996) J. Biol. Chem. 271, 15353–15357
16. Fields, G. B., and Prockop, D. J. (1996) Biopolymers 40, 345–357
17. Shah, N. K., Ramshaw, J. A. M., Kirkpatrick, A., Shah, C., and Brodsky, B. (1996) Biochemistry 35, 10262–10268
18. Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991) Nature 353, 726–730
19. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kuzmierski, W. M., and Knapp, R. J. (1991) Nature 354, 82–84
20. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnoffsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) Cell 72, 767–778
21. Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J., and Brent, R. (1996) Nature 386, 548–550
22. Bella, J., Eaton, M., Brodsky, B., and Berman H. M. (1994) Science 266, 75–81
23. Feng, Y., Melacini, G., and Goodman, M. (1997) Biochemistry 36, 8716–8724
24. Yang, W., Chan, V. C., Kirkpatrick, A., Ramshaw, J. A. M., and Brodsky, B. (1997) J. Biol. Chem. 272, 28837–28840
25. Chan, V. C., Ramshaw, J. A. M., Kirkpatrick, A., Beck, K., and Brodsky, B. (1997) J. Biol. Chem. 272, 31441–31446
26. Ackerman, M. S., Bhate, M., Shenoy, N., Beck, K., Ramshaw, J. A. M., and Brodsky, B. (1999) J. Biol. Chem. 274, 7668–7673
27. Walsmsley, A. R., Batten, M. R., Lad, U., and Bulleid, N. J. (1999) J. Biol. Chem. 274, 14884–14892
