Bioorganic & Medicinal Chemistry 18 (2010) 3767–3775

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

A structure–activity relationship study elucidating the mechanism of sequence-specific collagen recognition by the chaperone HSP47

Yoshimi Nishikawa [a], Yoshifumi Takahara [b], Shinichi Asada [c], Akira Shigenaga [d], Akira Otaka [d], Kouki Kitagawa [c], Takaki Koide [a,*]

[a] Department of Chemistry and Biochemistry, School of Advanced Science and Engineering, Waseda University, Tokyo 169-8555, Japan
[b] Faculty of Engineering, The University of Tokushima, Tokushima 770-8505, Japan
[c] Faculty of Pharmaceutical Sciences, Niigata University of Pharmacy and Applied Life Sciences, Niigata 956-8603, Japan
[d] Institute of Health Biosciences and Graduate School of Pharmaceutical Sciences, The University of Tokushima, Tokushima 770-8505, Japan

ARTICLE INFO

Article history:
Received 25 March 2010
Accepted 17 April 2010
Available online 21 April 2010

Keywords:
Heat-shock protein 47 (HSP47)
Chaperone
Collagen
Peptide

ABSTRACT

Heat-shock protein 47 (HSP47) is a chaperone that facilitates the proper folding of procollagen. Our previous studies showed that the high-affinity HSP47-binding motif in the collagen triple helix is Xaa-(Thr/Pro)-Gly-Xaa-Arg-Gly. In this study, we further investigated structural requirements for the HSP47-binding motif, using synthetic triple-helical collagen-model peptides with systematic amino acid substitutions at either the Thr/Pro (=Yaa^{-3}) or the Arg (=Yaa^{0}) position. Results obtained from in vitro binding assays indicated that HSP47 detects the side-chain structure of Arg at the Yaa^{0}-position, while the Yaa^{-3} amino acid serves as the secondary recognition site that affects affinity to HSP47.

1. Introduction

Collagen is the most abundant protein in mammals, making up about 1/4 to 1/3 of their total protein content. Its triple helix is a unique tertiary structure in which three left-handed polyproline type II helices are supercoiled. The close packing of the three chains requires Gly at every third residue, generating an (Xaa-Yaa-Gly) repeating sequence pattern. The Xaa and Yaa positions are often occupied by Pro and 4(*R*)-hydroxyproline (4(*R*)Hyp) residues, respectively.[1,2] Fibrosis is a progressive disease caused by excessive accumulation of collagen fibers. The expanding formation of collagen deposits lead to disruption of normal tissue structure and can cause organ dysfunction. The synthetic pathway of procollagen is known to be facilitated by HSP47, an endoplasmic reticulum-resident molecular chaperone.[3,4] Its chaperone function is essential for the synthesis of properly folded triple-helical procollagen molecules.[5,6]

Expression of HSP47 is reportedly up-regulated during the progression of various fibrotic lesions, and its role in collagen biosynthesis is also important in the progression of fibrosis. HSP47 is hence regarded as a candidate for the targets of antifibrotic drugs.[7,8] Recently, Niitsu and co-workers showed that liposomal administration of small interfering RNA (siRNA) against HSP47 was effective in treating experimental liver cirrhosis.[9] In addition, Ananthanarayanan and co-workers reported the discovery of small molecular compounds that inhibit HSP47–collagen interactions in vitro,[10] though the efficacy of the compounds is still debated.[11] Our research group has been focusing on molecular recognition between collagen and HSP47 (HSP47 binds to both procollagen and mature collagen), taking advantages of systems that involve chemically synthesized collagen-model peptides. To date, we have found that: (a) HSP47 recognizes collagen only when the polypeptide chains assume the triple-helical conformation;[12,13] (b) Arg residue at a Yaa-position of collagenous (Xaa-Yaa-Gly)-repeats is important for high-affinity binding;[14,15] (c) amino acid residues at Xaa positions do not greatly affect binding;[14] and (d) the amino acid (named Yaa^{-3}) at the Yaa position in the N-terminal contiguous tripeptide unit to the Arg (at the Yaa^{0}-position) is also directly recognized by HSP47.[13,16]

In this paper, we further investigated the HSP47–collagen recognition mechanism, focusing especially on key amino acid residues at the Yaa^{0}- and Yaa^{-3}-positions by examining interactions between synthetic triple-helical collagen-model peptides and recombinant HSP47.

2. Results and discussion

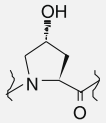
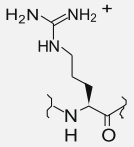
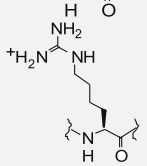
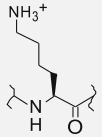
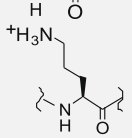
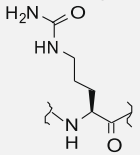
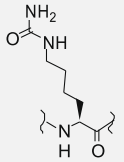
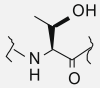
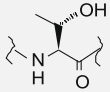
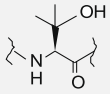
2.1. Design, synthesis and characterization of collagen-model peptides

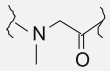
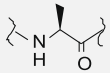
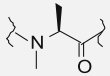
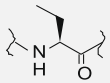
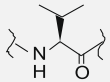
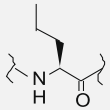
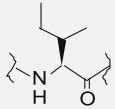
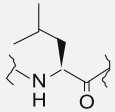
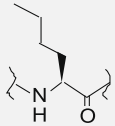
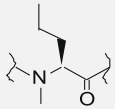
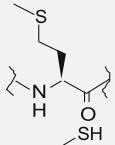
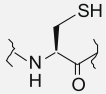
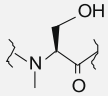
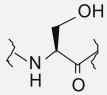
Collagen-model peptides used in this study are listed in Table 1. Here we used open-chain 'host–guest' peptides that spontaneously

* Corresponding author. Tel./fax: +81 3 5286 2569.
E-mail address: koi@waseda.jp (T. Koide).

0968-0896/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2010.04.054

Table 1
Structures and characterization of the collagen-model peptides

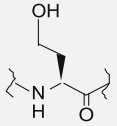
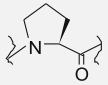
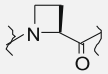
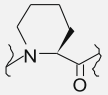
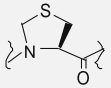
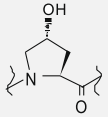
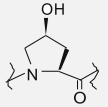
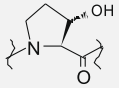
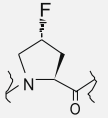
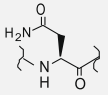
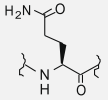
H-**Zaa**-(Pro-4(*R*)Hyp-Gly)**n**-Pro-**Yaa**$^{-3}$-Gly-Pro-**Yaa**^{0}-Gly-(Pro-4(*R*)Hyp-Gly)$_4$-NH_2											
Compound	Zaa	*n*	Yaa^{-3}		Yaa^{0}	Formula	MS [M+H]$^+$		CD max		T_m (°C)
							Calculated	Found	nm	$[\Theta]_{mrw}$ (deg cm^2/dmol)	
1	Cys	2		4(*R*)Hyp		Arg	$C_{100}H_{149}N_{29}O_{32}S_1$	2302.54	2302.42[b]	225.4	3643	44.5
2	Cys	2		4(*R*)Hyp		Har	$C_{101}H_{151}N_{29}O_{32}S_1$	2316.56	2316.91[b]	224.4	4457	42.0
3	Cys	2		4(*R*)Hyp		Lys[a]	$C_{100}H_{148}N_{26}O_{33}S_1$	2274.05	2276.81[b]	225.2	4004	35.0
4	Cys	2		4(*R*)Hyp		Orn[a]	$C_{99}H_{146}N_{26}O_{33}S_1$	2261.48	2259.36[b]	225.4	3585	32.5
5	Cys	2		4(*R*)Hyp		Cit	$C_{100}H_{148}N_{28}O_{33}S_1$	2303.52	2304.71[b]	225.4	3632	46.0
6	Cys	2		4(*R*)Hyp		Hci	$C_{101}H_{150}N_{28}O_{33}S_1$	2316.07	2315.84[b]	224.2	4458	40.5
7	Cys	2		4(*R*)Hyp	4(*R*)Hyp		$C_{99}H_{144}N_{26}O_{33}S_1$	2259.47	2259.24[b]	224.4	3515	46.5
8	None	3		Thr		Arg	$C_{108}H_{161}N_{31}O_{35}$	2453.18	2452.73[c]	224.1	4258	45.5[d]
9	None	3		*allo*-Thr		Arg	$C_{108}H_{161}N_{31}O_{35}$	2453.18	2453.40[c]	223.7	4052	45.5
10	None	3		3-methyl Thr		Arg	$C_{109}H_{163}N_{31}O_{35}$	2467.20	2467.40[c]	224.6	3784	43.5

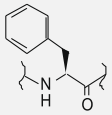
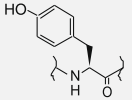
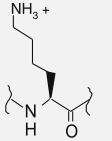
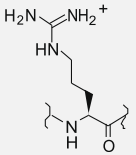
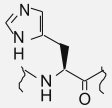
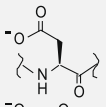
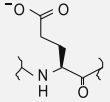
11	None	3		NMG	Arg	$C_{107}H_{159}N_{31}O_{34}$	2423.17	2423.39[c]	224.6	3925	42.0
12	None	3		Ala	Arg	$C_{107}H_{159}N_{31}O_{34}$	2423.17	2423.98[c]	224.6	4511	47.0[d]
13	None	3		NMA	Arg	$C_{108}H_{161}N_{31}O_{34}$	2437.19	2437.38[c]	224.4	4848	43.7
14	None	3		Abu	Arg	$C_{108}H_{161}N_{31}O_{34}$	2437.19	2437.36[c]	224.5	3956	48.5
15	None	3		Val	Arg	$C_{109}H_{163}N_{31}O_{34}$	2451.21	2450.98[c]	224.6	4167	47.8[d]
16	None	3		Nva	Arg	$C_{109}H_{163}N_{31}O_{34}$	2451.21	2451.37[c]	224.3	3986	48.2
17	None	3		Ile	Arg	$C_{110}H_{165}N_{31}O_{34}$	2465.22	2464.73[c]	224.8	4274	47.8[d]
18	None	3		Leu	Arg	$C_{110}H_{165}N_{31}O_{34}$	2465.22	2465.98[c]	224.6	4124	39.5[d]
19	None	3		Nle	Arg	$C_{110}H_{165}N_{31}O_{34}$	2465.22	2466.10[c]	224.2	4097	46.0
20	None	3		NMNva	Arg	$C_{110}H_{165}N_{31}O_{34}$	2465.22	2465.44[c]	225.5	3998	34.4
21	None	3		Met	Arg	$C_{109}H_{163}N_{31}O_{34}S_1$	2483.18	2482.73[c]	224.8	4204	49.5[d]
22	None	3		Cys	Arg	$C_{107}H_{159}N_{31}O_{34}S_1$	2455.15	2455.40[c]	223.6	4129	43.5
23	None	3		NMS	Arg	$C_{108}H_{161}N_{31}O_{35}$	2453.18	2453.36[c]	224.2	4092	38.7
24	None	3		Ser	Arg	$C_{107}H_{159}N_{31}O_{35}$	2439.17	2438.98[c]	223.8	4121	42.5[d]

(*continued on next page*)

Table 1 (*continued*)

H-**Zaa**-(Pro-4(*R*)Hyp-Gly)**n**-Pro-**Yaa**$^{-3}$-Gly-Pro-**Yaa**^{0}-Gly-(Pro-4(*R*)Hyp-Gly)$_4$-NH_2

Compound	Zaa	*n*	Yaa^{-3}	Yaa^{0}	Formula	MS [M+H]$^+$		CD max		T_m (°C)
						Calculated	Found	nm	$[\Theta]_{mrw}$ (deg cm^2/dmol)	
25	None	3	 Hse	Arg	$C_{108}H_{161}N_{31}O_{35}$	2453.18	2453.47[c]	223.8	3994	45.5
26	None	3	 Pro	Arg	$C_{109}H_{161}N_{31}O_{34}$	2449.19	2449.23[c]	225.2	3958	51.3[d]
27	None	3	 Aze	Arg	$C_{108}H_{159}N_{31}O_{34}$	2435.17	2435.40[c]	225.3	3391	41.5
28	None	3	 Pic	Arg	$C_{110}H_{163}N_{31}O_{34}$	2463.21	2463.29[c]	224.8	4805	43.5
29	None	3	 Thp	Arg	$C_{108}H_{159}N_{31}O_{34}S_1$	2467.15	2467.22[c]	224.9	4077	50.2
30	None	2	 4(*R*)Hyp	Arg	$C_{109}H_{161}N_{31}O_{34}$	2199.39	2199.78[b]	224.8	4231	45.2[d]
31	None	3	 4(*S*)Hyp	Arg	$C_{109}H_{161}N_{31}O_{35}$	2465.18	2465.41[c]	224.5	4135	49.4
32	None	3	 3(*R*)Hyp	Arg	$C_{109}H_{161}N_{31}O_{35}$	2465.18	2465.44[c]	224.8	4018	45.5
33	None	3	 4(*R*)Flp	Arg	$C_{109}H_{160}N_{31}O_{34}F_1$	2467.18	2467.37[c]	225.2	3685	55.5
34	None	3	 Asn	Arg	$C_{108}H_{160}N_{32}O_{35}$	2466.18	2466.48[c]	224.6	4112	34.3[d]
35	None	3	 Gln	Arg	$C_{109}H_{162}N_{32}O_{35}$	2480.20	2480.98[c]	224.0	4001	47.4[d]

36	None	3		Phe	Arg	$C_{113}H_{163}N_{31}O_{34}$	2499.21	2498.98[c]	222.4	4919	32.9[d]
37	None	3		Tyr	Arg	$C_{113}H_{163}N_{31}O_{35}$	2515.20	2515.23[c]	224.8	6568	33.0[d]
38	None	3		Lys	Arg	$C_{110}H_{166}N_{32}O_{34}$	2480.23	2479.73[c]	224.6	4025	43.5[d]
39	None	3		Arg[e]	Arg	$C_{110}H_{166}N_{34}O_{35}$	2524.23	2525.11[b]	224.6	3727	37.5[d]
40	None	3		His	Arg	$C_{110}H_{161}N_{33}O_{34}$	2489.20	2489.48[c]	223.6	4416	39.0[d]
41	None	3		Asp	Arg	$C_{108}H_{159}N_{31}O_{36}$	2467.16	2466.48[c]	223.8	4146	34.5[d]
42	None	3		Glu	Arg	$C_{109}H_{161}N_{31}O_{36}$	2481.18	2481.48[c]	224.8	4246	44.0[d]

Abbreviations: Cit, citrulline; Har, homoarginine; Orn, ornithine; Hci, homocitrulline; 4(*R*)Flp, 4(*R*)-fluoroproline; Pic, pipecolic acid; NMA, *N*-methylalanine; NMS, *N*-methylserine; Thp, thioproline; Abu, 2-aminobutyric acid; NMNva, *N*-methylnorvaline; Aze, azetidine-2-carboxylic acid; 3(*R*)Hyp, 3(*R*)-hydroxyproline; 4(*R*)Hyp, 4(*R*)-hydroxyproline; 4(*S*)Hyp, 4(*S*)-hydroxyproline; Nva, norvaline; NMG, *N*-methylglycine; Hse, homoserine; Nle, norleucine.

[a] Peptide with a free C-terminal carboxyl group.

[b] MALDI-MS.

[c] ESI-MS.

[d] Ref. 16.

[e] H-(Pro-4(*R*)Hyp-Gly)$_2$-Pro-Glu-Gly-Pro-Arg^{-3}-Gly-Pro-Arg^{0}-Gly-(Pro-4(*R*)Hyp-Gly)$_4$-NH_2.

form homotrimeric collagen helices in aqueous solutions. In the design, the 'guest' Xaa-Yaa^{-3}-Gly-Xaa-Yaa^{0}-Gly sequences with various Yaa^{0}- or Yaa^{-3}-amino acid substitutions were flanked by several repeats of the helix-stabilizing 'host' triplet. We chose (Pro-4(*R*)Hyp-Gly)-repeats as the 'host sequence,'[17] because it has a high helix-stabilizing effect with negligible HSP47-binding activity.[18] Since amino acid residues at Xaa-positions do not largely contribute to the binding,[14] the Xaa residues were fixed to Pro, the most frequently appearing residue at this position. The C-terminal ends of the peptides were amidated, to further stabilize the triple helices by avoiding charge repulsions between carboxylate anions.[19] The N-termini of peptides **8**–**42** were left free to be immobilized to NHS-activated Sepharose-beads through their N^{α}-amine functionality. In the case of the immobilization of a set of Y^{0}-substituted peptides including those containing Lys and Orn residues (**1**–**7**), we instead utilized the thiol group of N-terminal-added Cys residues to ensure the orientation of immobilized peptides to prevent coupling with the side-chain-amine groups.

The peptides were synthesized by the *N*-9-fluorenylmethoxycarbonyl (Fmoc)-based solid-phase method, and purified by reversed phase (RP)-HPLC. The identities of the purified peptides were confirmed by mass spectrometry; all the measured masses agreed with the theoretical values (Table 1).

As triple-helical conformation is essential to the interaction with HSP47,[12,13] the conformational states of the peptides were analyzed by means of circular dichroism (CD) spectrometry prior to the binding assays. Solutions of all peptides showed positive CD signals around 225 nm at 25 °C (Table 1). The wavelengths at the maximal signals are very similar to one another, and the mean residue molecular ellipticity ($[\Theta]_{mrw}$) values at the maximal wavelengths around 4000 deg cm^2/dmol are consistent with those of typical collagen triple helix,[17,20] indicating that these peptides assumed triple-helical conformations, and their overall helical structures were very similar to one another (Table 1).

2.2. Assay systems for evaluating HSP47-binding activity

HSP47-binding activities of the peptides were estimated by means of two assay systems. One is a bead-pull-down assay using affinity beads on which peptides are immobilized. This assay allowed us to detect direct binding between HSP47 and peptides. The other is a competition assay on a surface plasmon resonance (SPR) biosensor to evaluate the relative HSP47-binding affinities of the peptides. In this assay, a peptide of interest was used as an inhibitor for binding between HSP47 and an R/R/R-peptide, a previously reported HSP47-binder, immobilized on a sensor chip.[21] Here, we used the IC$_{50}$s of trimeric peptides as indexes that represent their relative HSP47-binding affinities.

2.3. HSP47 recognizes side-chain structure of Arg residue at the Yaa^{0}-position

First, we investigated HSP47 recognition of the Yaa-position (Yaa^{0}), using collagen-model peptides that had Arg surrogates (Table 1, **1**–**6**). We examined homoarginine (Har **2**), lysine (Lys **3**), ornithine (Orn **4**), citrulline (Cit **5**), and homocitrulline (Hci **6**) residues at the Yaa^{0}-position. A 4(*R*)Hyp containing peptide (**7**) was used as a negative control.[18] In the pull-down assay, significant interaction with HSP47 was observed when Arg (**1**) or Har (**2**) was incorporated as shown in Figure 1. The other peptides (**3**–**7**) did not exhibit detectable HSP47 binding. Relative HSP47-binding affinities of the Har-containing peptide **2** (IC$_{50}$ = 14 μM), determined by a peptide competition assay using SPR-based technique, was 3.4 times lower than that of the Arg-containing peptide **1** (IC$_{50}$ = 4.1 μM; Fig. 1B and Tables 1 and 2). When Lys **3** or Orn **4** was incorporated at the Yaa^{-3}-position, significant binding of HSP47 was not detected (IC$_{50}$s >100 μM). This indicated that the guanidium moiety of the side-chain is essential to the binding, and the positive charge of the Yaa^{0}-residue by itself does not account for the specific binding to HSP47. It also suggests the importance of the length of the side-chain. This finding is consistent with the previous finding that HSP47 binding would depend mainly on hydrophobic interaction, rather than ionic interaction.[22] The X-ray crystal structure of a collagen-model peptide with the guest

Table 2
IC$_{50}$ values for Yaa^{0}-substituted collagen-model peptides

H-Cys-(Pro-4(*R*)Hyp-Gly)$_3$-Pro-**Yaa**^{0}-Gly-(Pro-4(*R*)Hyp-Gly)$_4$-NH$_2$		
Compound	Yaa^{0}	IC$_{50}$[a] (μM)
1	Arg	4.1 ± 0.22
2	Har	14 ± 0.25
3	Lys[b]	>100
4	Orn[b]	>100
5	Cit	>100
6	Hci	>100
7	4(*R*)Hyp	>100

Abbreviations: Cit, citrulline; Har, homoarginine; Orn, ornithine; Hci, homocitrulline; 4(*R*)Hyp, 4(*R*)-hydroxyproline.

[a] IC$_{50}$ values represent the means ± SD obtained from three independent determinations. Molar concentrations of peptide trimers are used.

[b] Peptide with a free C-terminal carboxyl group.

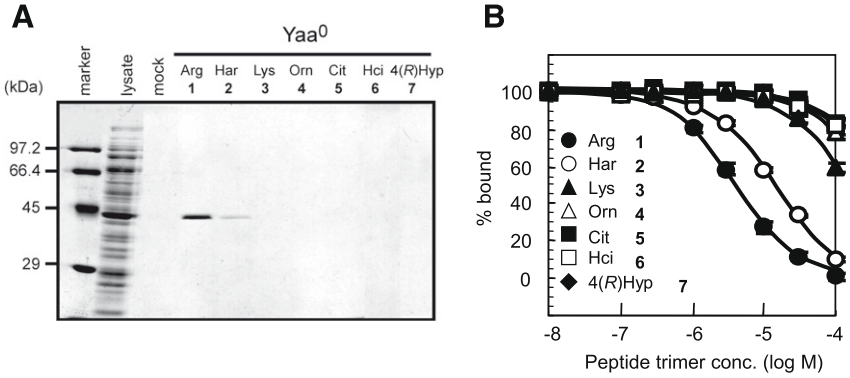


Figure 1. Interaction of HSP47 with Yaa^{0}-substituted collagen-model peptides. (A) Binding of HSP47 to immobilized H-Cys-(Pro-4(*R*)Hyp-Gly)$_3$-Pro-Yaa^{0}-Gly-(Pro-4(*R*)Hyp-Gly)$_4$-NH$_2$. Bacterial lysates containing recombinant HSP47 were mixed with the peptide-immobilized beads. After washing the beads, proteins retained on the beads were separated by 12% SDS–PAGE and visualized by CBB staining. Molecular sizes are shown in kilodaltons. (B) Competitive binding assays on Biacore. HSP47-binding was measured as described in Section 4. Values are means ± SD (n = 3).

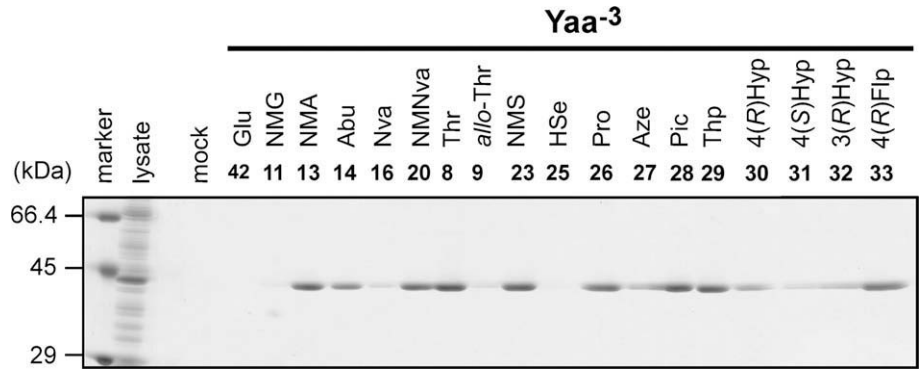


Figure 2. Interaction of HSP47 with Yaa^{-3}-substituted collagen-model peptides. Binding of HSP47 to H-(Pro-4(*R*)Hyp-Gly)$_3$-Pro-Yaa^{-3}-Gly-Pro-Arg^0-Gly-(Pro-4(*R*)Hyp-Gly)$_4$-NH_2 analyzed in a manner similar to that in Figure 1A.

sequence corresponding to human α1(III) 785–796 (T3–785), showed that the Arg side-chain at the Yaa position radiating from the triple helix makes direct contact with backbone carbonyl groups through its N^{ε} or N^{η}.[23] We suggest that the hydrophobic ring structure, formed by the Arg side-chain, is a key structure for specific recognition by HSP47.

2.4. Structure–activity relationships of the Yaa^{-3}-residue

Our previous study involving peptides with the replacement of the Yaa^{-3} residues with natural amino acids showed that HSP47-binding affinity was largely affected by the amino acid at the Yaa^{-3} position. Especially, Thr and Pro at the Yaa^{-3} position exhibited high affinity for HSP47 binding.[16] In the current study, we sought to identify structural elements of the Yaa^{-3} amino acids required for high-affinity binding to HSP47, by using triple-helical peptides containing non-natural amino acids at this position. Based on previous findings, we mainly focused on Thr analogs and Pro analogs. In this series of peptides, the Yaa^0 residue was fixed to Arg.

The HSP47-binding affinities of peptides with non-natural amino acid substitutions at Yaa^{-3} positions were analyzed by a similar pull-down assay. Here, we used the Glu-containing peptide **42** as a negative control.[16] As shown in Figure 2, amounts of bead-bound HSP47 varied according to the amino acid incorporated, indicating that HSP47 recognized the side-chain structure of the Yaa^{-3} residue. The IC_{50} values were further estimated from inhibition curves obtained from the SPR competition assay (Table 3). No relationships between the binding affinities and conformational parameters, such as CD maxima or melting temperatures (T_m) of the triple helices, were found. The best structure for HSP47 binding had Thr (**8**, IC_{50} = 0.58 μM) at the Yaa^{-3} position, followed by 4(*R*)-fluoroproline (4(*R*)Flp) (**33**, IC_{50} = 0.64 μM), pipecolic acid (Pic) (**28**, IC_{50} = 0.69 μM), *N*-methyl Ala (NMA) (**13**, IC_{50} = 0.94 μM) and Pro (**26**, IC_{50} = 0.94 μM). Hydrophobic five- or six-membered ring structures apparently contribute to the high-affinity binding because Flp **33**, Pic **28**, Pro **26,** and Thp **29** showed relatively small IC_{50} values. Introduction of a four-membered ring (Aze **27**) decreased affinity to a certain degree, presumably because of local distortion of the triple helical structure.[24] The introduction of hydroxyl groups to the ring (3(*R*)Hyp **32**, 4(*R*)Hyp **30**, and 4(*S*)Hyp **31**) decreased affinity.

In cases using amino acids with linear side-chains, existence of three or more atoms led to substantially decreased affinity (Nva **16**, Ile **17**, Hse **25**, Nle **19**, Leu **18**, Asn **34**, Met **21**, Lys **38**, Gln **35**, Glu **42**, Asp **41**, and Arg **39**). This result suggests that the depth of the binding pocket for Yaa^{-3} could accommodate linear side-chains shorter than three atoms in length. Since peptides with aromatic or imidazole rings attached to the β-carbon (Phe **36**, Tyr **37** and His **40**) showed very weak affinities, this binding pocket would not suit these rings, either.

Focusing on the β-substituents of the Thr residue at Yaa^{-3}, the configuration of the Thr hydroxyl group was shown to be critical for high-affinity binding, because peptide **9,** possessing an *allo*-Thr residue, showed marked decrease in binding affinity (IC_{50} = 11 μM). Trisubstitution of the β-carbon was also disfavored (3-methyl Thr **10**, IC_{50} = 9.1 μM). The β-methyl group of the Thr residue was found to be more important than the hydroxyl group because Abu **14** showed smaller IC_{50} (2.4 μM) values than Ser **24** (IC_{50} = 4.0 μM). This hypothesis is also supported by the finding that the hydroxyl group (Ser **24**) could be changed to a thiol group (Cys **22**, IC_{50} = 4.2 μM) without affecting the HSP47-binding affinity.

The *N*-methyl group increases binding affinity. *N*-methyl Ala (NMA)-substituted peptide **13** (IC_{50} = 0.94 μM) showed higher binding affinity than Ala **12** (IC_{50} = 9.3 μM). Similarly, peptides containing *N*-methyl norvaline (NMNVa) **20** (IC_{50} = 2.5 μM) or *N*-methyl Ser (NMS) **23** (IC_{50} = 1.1 μM) showed about fourfold higher binding affinities than peptides containing Nva **16** (IC_{50} = 9.8 μM) or Ser **24** (IC_{50} = 4.0 μM), respectively. This result is consistent with the finding that peptides with Pro, and its analogs with five- and six-membered rings, have high affinities for HSP47. At this stage it is unclear whether HSP47 directly recognizes the *N*-alkyl group or the *N*-methyl group stabilizes local conformation of the favored structure.

3. Conclusion

In this study, we found the structural elements of the collagen triple helix required for recognition by the chaperone HSP47 by focusing on the essential Arg-like structure at the Yaa^0 position and the secondary position at Yaa^{-3} that regulate relative affinity to HSP47. It has been suggested that HSP47 stabilizes the triple-helical conformation of the procollagen intermediate, which is otherwise unstable at body temperature;[25] this implies that HSP47 binds to relatively unstable portions of the procollagen triple helix. However, we found no significant relationships between triple-helical stability of collagen sequences and their HSP47-binding affinities.

Although we used only triple-helical peptides as assessed by CD spectrometry, the local conformation (or helical twists) of the guest sequence could differ according to the amino acids incorporated.[23,26] At this stage, we could not eliminate the factor of local conformation. It is possible that differences in local structure could be clarified by X-ray-crystallographic analyses of the peptides in the future. Nevertheless, the amino acid preferences at Yaa^0 and Yaa^{-3} positions obtained from systematic amino acid replacements would give clues

Table 3
IC_{50} values of Yaa^{-3}-substituted collagen-model peptides

H-(Pro-4(*R*)Hyp-Gly)$_3$-Pro-**Yaa**$^{-3}$-Gly-Pro-**Arg**^{0}-Gly-(Pro-4(*R*)Hyp-Gly)$_4$-NH_2

Compound	Yaa^{-3}	IC_{50}[a] (μM)
8	Thr	0.58 ± 0.13[a]
33	4(*R*)Flp	0.64 ± 0.052
28	Pic	0.69 ± 0.065
13	NMA	0.94 ± 0.066
26	Pro	0.94 ± 0.095[b]
23	NMS	1.1 ± 0.11
29	Thp	1.4 ± 0.063
14	Abu	2.4 ± 0.21
20	NMNva	2.5 ± 0.15
24	Ser	4.0 ± 0.73[b]
22	Cys	4.2 ± 0.44
27	Aze	4.2 ± 0.26
32	3(*R*)Hyp	4.5 ± 0.21
30	4(*R*)Hyp	4.6 ± 0.81[b]
15	Val	8.4 ± 0.64[b]
10	3-methyl Thr	9.1 ± 0.38
12	Ala	9.3 ± 1.1[b]
31	4(*S*)Hyp	9.5 ± 0.59
16	Nva	9.8 ± 0.79
9	*allo*-Thr	11 ± 0.89
11	NMG	18 ± 0.12
17	Ile	19 ± 7.6[b]
25	Hse	23 ± 1.5
19	Nle	35 ± 3.5
18	Leu	37 ± 0.84[b]
34	Asn	41 ± 5.6[b]
21	Met	43 ± 3.3[b]
40	His	64 ± 12[b]
36	Phe	79 ± 13[b]
37	Tyr	95 ± 9.8[b]
38	Lys	>100[b]
35	Gln	>100[b]
42	Glu	>100[b]
41	Asp	>100[b]
39	Arg[c]	>100[b]

Abbreviations: 4(*R*)Flp, 4(*R*)-fluoroproline; Pic, pipecolic acid; NMA, *N*-methylalanine; NMS, *N*-methylseline; Thp, thioproline; Abu, 2-aminobutyric acid; NMNva, *N*-methylnorvaline; Aze, azetidine-2-carboxylic acid; 3(*R*)Hyp, 3(*R*)-hydroxyproline; 4(*R*)Hyp, 4(*R*)-hydroxyproline; 4(*S*)Hyp, 4(*S*)-hydroxyproline; Nva, norvaline; NMG, *N*-methylglycine; Hse, homoserine; Nle, norleucine.

[a] IC_{50} values represent the means ± SD obtained from three (or six for **8**) independent assays. Molar concentrations of peptide trimers are used.

[b] Ref. 16.

[c] H-(Pro-4(*R*)Hyp-Gly)$_2$-Pro-Glu-Gly-Pro-Arg^{-3}-Gly-Pro-Arg^{0}-Gly-(Pro-4(*R*)Hyp-Gly)$_4$-NH_2.

for rational design or selection of effective HSP47 inhibitors as potential antifibrotic drugs.

4. Experimental procedures

4.1. General

Fmoc-protected amino acids and resins were purchased from Watanabe Chemical Industries., Ltd (Hiroshima, Japan) and Novabiochem (Darmstadt, Germany). All other chemicals were purchased from Nacalai Tesque (Kyoto, Japan), Sigma (St. Louis, MO) and Dojindo (Kumamoto, Japan). For analytical RP-HPLC, a Cosmosil 5C_{18} AR-II column (4.6 × 250 mm, Nacalai Tesque) was used at a flow rate of 1 mL/min. For preparative RP-HPLC, a Cosmosil 5C_{18} AR-II column (20 × 250 mm, Nacalai Tesque) was used at a flow rate of 3.5–7 mL/min.

4.2. Synthesis of collagen model peptides

Peptide chains were constructed manually based on the standard Fmoc-based strategy on Rink-amide resins (Novabiochem) or H-Gly-ClTrt resin. *tert*-Butyl, *tert*-butoxycarbonyl, 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf), and triphenylmethyl were used as side-chain-protecting groups for hydroxyl, amino, guanidino, and thiol groups, respectively. Peptide cleavage/deprotection was performed by treatment with trifluoroacetic acid (TFA)/*m*-cresol/thioanisole/water/1,2-ethanedithiol (82.5:5:5:5:2.5, v/v) at room temperature for 1 h. For Arg(Pbf)-containing peptides, the treatment was prolonged to 3–4 h. All peptides were purified by RP-HPLC with a linear gradient of acetonitrile in water, both containing 0.05% TFA, and identified with electron-spray-ionization mass spectrometry (ESI-MS) or matrix-assisted laser desorption–ionization time-of-flight mass spectrometry (MALDI-TOF-MS).

Prior to use, aqueous solutions of the peptides were kept at 4 °C for at least two days to allow the formation of triple helices. Stock solutions of 10 mg/mL Cys-containing peptides were prepared in degassed 0.05% aqueous TFA to prevent formation of disulfide-bridged dimers by oxidation, and diluted with appropriate buffers before use.

4.3. CD spectroscopy

CD spectra of refolded peptides (1 mg/mL in 50 mM bis-Tris–HCl [pH 7.4], 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid [EDTA]) were recorded over a range of 190–260 nm on a JASCO J-820 spectropolarimeter using a 0.5 mm path-length quartz cell at 25 °C. Temperature was controlled using a Peltier thermal controller. Each measurement was repeated three times, and the final result was averaged of the three independent scans. The CD spectra were expressed as mean residual ellipticity, $[\Theta]_{mrw}$. Thermal transition curves of the triple helices were obtained by recording the ellipticity at 225 nm at a heating rate of 0.3 °C/min. The T_m values were evaluated from the midpoint of the transition.

4.4. Protein preparation

Recombinant mouse HSP47 was produced and purified essentially as described previously.[18] Briefly, *Escherichia coli* (strain JM109) that harbored expression plasmids for mouse HSP47 was cultured overnight in Luria-Bertani medium, diluted 50-fold, and continued to cultivate at 32 °C. Protein synthesis was induced by raising the temperature to 40.5 °C, for 1 h. Cells were then harvested by centrifugation and lysed in lysis buffer (50 mM Tris–HCl [pH 8.0], 150 mM NaCl, 5 mM EDTA, 30% [v/v] glycerol) containing 1 mg/mL lysozyme (Wako, Osaka, Japan), followed by the addition of protease inhibitors and Nonidet P-40 (final concentration 0.2%). The resulting suspension was sonicated; cell debris was then removed by centrifugation. The lysates were mixed with gelatin-Sepharose and rotated at 4 °C for 1 h. After washing the beads with high-ionic-strength buffer (50 mM bis–Tris–HCl [pH 7.4] containing 0.4 M NaCl and 1 mM EDTA), retained HSP47 was eluted with 50 mM bis–Tris–HCl (pH 5.4) containing 0.4 M NaCl and 1 mM EDTA. The HSP47-containing fractions were neutralized to pH 7.4 with 0.5 M aqueous NaOH, and filtered using 0.2-μm PTFE filters (Millipore, MA). Protein concentrations were determined using a protein assay kit (BioRad, Hercules, CA) with bovine serum albumin as the standard. The purified HSP47 obtained was used within a day.

4.5. Solid-phase binding assay

Peptide immobilization onto NHS-activated Sepharose (GE Healthcare, Milwaukee, WI) and solid-phase binding assays were performed as reported earlier.[14] The set of Cys-containing peptides was immobilized onto thiopropyl-Sepharose 6B (GE Healthcare). *E. coli* lysate containing recombinant mouse HSP47 (70 μL) was mixed with 130 μL of binding buffer (10 mM HEPES·Na [pH 7.5], 3.7 mM EDTA, 0.4 M NaCl, 0.005% Tween 20) and 25-μL bed of affin-

ity beads. Binding was carried out at 4 °C for 1 h. Beads were washed with high-ionic-strength buffer; proteins retained on beads were then eluted by adding 25 μL of 2 × Laemmli sample buffer (50 mM Tris–HCl [pH 6.7], 2% SDS, 10% glycerol, 0.002% bromophenol blue), separated by 12% SDS–PAGE, and visualized by staining with Coomassie Brilliant Blue R-250 (CBB).

4.6. SPR

Competitive binding assays were carried out using an SPR biosensor (Biacore X; GE Healthcare) essentially as described previously.[16] Briefly, R/R/R peptides[21] were immobilized on the surface of a CM5 sensor chip using the standard amine coupling method at pH 5.6 as described by the manufacturer. Resonance was measured in 50 mM bis–Tris–HCl (pH 7.4) containing 0.4 M NaCl, 1 mM EDTA and 0.005% Tween 20 as a running buffer at 25 °C at a flow rate of 40 μl/min. Before each measurement, HSP47 (final concentration: 25 or 50 μg/mL) was incubated with various concentrations of the competitor peptides for 10 min at 25 °C. Measurements were carried out three times. Obtained data were analyzed using Origin ver.7.5 data analysis software (OriginLab, Northampton, MA); IC_{50} values were estimated from the midpoint of the inhibition curves.

Acknowledgments

This work was supported by a Grant-in Aid for Scientific Research (B) No. 19390032 from the Japan Society for the Promotion of Science, and a Waseda University Grant for Special Research Projects (No. 2008B-098).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2010.04.054.

References and notes

1. Beck, K.; Brodsky, B. *J. Struct. Biol.* **1998**, *122*, 17.
2. Okuyama, K.; Hongo, C.; Fukushima, R.; Wu, G.; Narita, H.; Noguchi, K.; Tanaka, Y.; Nishino, N. *Biopolymers* **2004**, *76*, 367.
3. Nagata, K. *Semin. Cell Dev. Biol.* **2003**, *14*, 275.
4. Koide, T.; Nagata, K. *Top. Curr. Chem.* **2005**, *247*, 85.
5. Nagai, N.; Hosokawa, M.; Itohara, S.; Adachi, E.; Matsushita, T.; Hosokawa, N.; Nagata, K. *J. Cell Biol.* **2000**, *150*, 1499.
6. Ishida, Y.; Kubota, H.; Yamamoto, A.; Kitamura, A.; Bächinger, H. P.; Nagata, K. *Mol. Biol. Cell* **2006**, *17*, 2345.
7. Sauk, J. J.; Nikitakis, N.; Siavash, H. *Front. Biosci.* **2005**, *10*, 107.
8. Taguchi, T.; Razzaque, M. *Trends Mol. Med.* **2007**, *13*, 45.
9. Sato, Y.; Murase, K.; Kato, J.; Kobune, M.; Sato, T.; Kawano, Y.; Takimoto, R.; Takada, K.; Miyanishi, K.; Matsunaga, T.; Takayama, T.; Niitsu, Y. *Nat. Biotechnol.* **2008**, *26*, 431.
10. Thomson, C. A.; Atkinson, H. M.; Ananthanarayanan, V. S. *J. Med. Chem.* **2005**, *48*, 1680.
11. Okano-Kosugi, H.; Matsushita, O.; Asada, S.; Herr, A. B.; Kitagawa, K.; Koide, T. *Anal. Biochem.* **2009**, *394*, 125.
12. Koide, T.; Aso, A.; Yorihuzi, T.; Nagata, K. *J. Biol. Chem.* **2000**, *275*, 27957.
13. Koide, T.; Asada, S.; Takahara, Y.; Nishikawa, Y.; Nagata, K.; Kitagawa, K. *J. Biol. Chem.* **2006**, *281*, 3432.
14. Koide, T.; Takahara, Y.; Asada, S.; Nagata, K. *J. Biol. Chem.* **2002**, *277*, 6178.
15. Tasab, M.; Jenkinson, L.; Bulleid, N. *J. Biol. Chem.* **2002**, *277*, 35007.
16. Koide, T.; Nishikawa, Y.; Asada, S.; Yamazaki, C. M.; Takahara, Y.; Homma, D. L.; Otaka, A.; Ohtani, K.; Wakamiya, N.; Nagata, K.; Kitagawa, K. *J. Biol. Chem.* **2006**, *281*, 11177.
17. Shah, N. K.; Ramshaw, J. A. M.; Kirkpatrick, A.; Shah, C.; Brodsky, B. *Biochemistry* **1996**, *35*, 10262.
18. Koide, T.; Asada, S.; Nagata, K. *J. Biol. Chem.* **1999**, *274*, 34523.
19. Venugopal, M. G.; Ramshaw, J. A.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33*, 7948.
20. Chan, V. C.; Ramshaw, J. A. M.; Kirkpatrick, A.; Beck, K.; Brodsky, B. *J. Biol. Chem.* **1997**, *272*, 31441.
21. R/R/R-peptide is a disulfide-tethered trimer with the core sequence of [βAla-(Gly-Pro-4(*R*)Hyp)$_2$-Gly-Pro-Arg-(Gly-Pro-4(*R*)Hyp)$_2$-Gly-NH$_2$]$_3$ for detail see: Koide, T.; Nishikawa, Y.; Takahara, Y. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 125.
22. Saga, S.; Nagata, K.; Chen, W. T.; Yamada, K. M. *J. Cell Biol.* **1987**, *105*, 517.
23. Kramer, R. Z.; Bella, J.; Brodsky, B.; Berman, H. M. *J. Mol. Biol.* **2001**, *311*, 131. T3-785: H-(Pro-4(*R*)Hyp-Gly)$_3$-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-Pro-Hyp-Gly-(Pro-4(*R*)Hyp-Gly)$_3$-OH.
24. Zagari, A.; Nemethy, G.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 967.
25. Leikina, E.; Mertts, M. V.; Kuznetsova, N.; Leikin, S. *Proc. Natl. Acad. Sci., U.S.A.* **2002**, *99*, 1314.
26. Okuyama, K.; Wu, G.; Jiravanichanun, N.; Hongo, C.; Noguchi, K. *Biopolymers* **2006**, *84*, 421.