Chaperonin GroEL from *Escherichia coli* binds to the non-native states of many unrelated proteins, and GroEL-recognizable structural features have been argued. As model substrate proteins of GroEL, we used seven artificial proteins (138–141 residues), each of which has a unique but randomly chosen amino acid sequence and no propensity to fold into a certain structure. Two of them were water-soluble, and the rest were soluble in 3 M urea. The soluble ones interacted with GroEL in a manner similar to that of a natural substrate; they stimulated the ATPase cycle of GroEL and GroEL/GroES and inhibited GroEL-assisted folding of other protein. All seven artificial proteins were able to bind to GroEL. The results suggest that the secondary structure as well as the specific sequence motif of the substrate proteins are not necessary to be recognized by GroEL.

Molecular chaperones bind to the non-native states of many unrelated proteins, thereby preventing aggregation or reverting folding reactions (1–7). Chaperonin GroEL from *Escherichia coli* is the best understood molecular chaperone (1–13). GroEL is composed of 14 identical 57-kDa subunits arranged in two heptameric rings stacked back to back (14, 15). Each ring defines a central cavity ~45 Å in diameter. The cavity surface of GroEL ring is lined with hydrophobic residues that can bind non-native polypeptides that expose regions of hydrophobic surface (16–24). GroEL protects these non-native substrate proteins from irreversible aggregation and improves the final yield of productive folding in a co-chaperonin GroES- and ATP-dependent manner.

A wide range of proteins, unrelated in their primary sequences and native tertiary structures, can bind to GroEL (25–27). Because GroEL cannot interact with proteins in the native state, there should be specific local structural features that are recognized by GroEL and are commonly found in many proteins in the non-native, but not native, state. However, it remains yet unclear what kinds of structural features of these proteins are recognized by GroEL. Studies using a variety of proteins as a model of substrate protein have been carried out, but they have led to several different proposals on the GroEL-recognizable structural features, ranging from a fully unfolded or native-like conformation (28–36). Some proteins (short peptides) containing α-helices were reported to bind to GroEL (37–40), although it was shown that an all or mainly β-sheet protein (Fab fragment, pepsin, and green fluorescent protein) could bind to GroEL (16, 41, 42). On the other hand, binding of GroEL to unfolded proteins without an ordered secondary structure, apocytochrome c (33) and α-casein (21), was demonstrated. In the crystal structure of the isolated apical domain of truncated GroEL with a 17-residue NH2-terminal tag, the region identified previously as a polypeptide binding site of GroEL indeed held a peptide stretch of the tag (residues –1 to –7) region with a nearly fully extended configuration probably as a mimic of the natural substrate protein (17).

Generally, natural proteins have an inherent propensity to form secondary (and tertiary) structures. This may cause ambiguity in the search for a GroEL-recognizable structure because it is not easy to exclude the possibility that the folding state of the substrate protein changes upon binding to GroEL. Even such a possibility exists that a portion of the bound protein which is not involved directly in the binding is allowed to fold some secondary structure. Recently Yomo and his colleagues (43) reported that the plasmid library generates numerous artificial proteins, each of which has a unique and randomly chosen amino acid sequence. These proteins are free from evolutionary constraints and, except for very rare probability, have no propensity to fold into a certain structure (44). Taking advantage of this, we used seven different artificial proteins with 138–141 amino acids as model substrate proteins of GroEL and found that all of them are able to bind to GroEL. Thus the secondary structure as well as specific sequence motif of the substrate proteins are not necessary for recognition by GroEL.

**EXPERIMENTAL PROCEDURES**

**Materials**—Bovine apo-αLA (type III, Ca2+-depleted), α-casein (from bovine milk), and rhodanese (type II, from bovine liver) were purchased from Sigma. Alkaline phosphatase-conjugated anti-T7 tag antibody was purchased from Novagen.

**Mutation**—pEL237E carrying a gene of the GroEL mutant (L237E), which is unable to bind polypeptide (18), was generated by site-directed mutagenesis using an oligonucleotide GCAACAGCTTCTTCAACCGG-

**Protein Expression and Purification**—The purification of GroEL from *E. coli* JM109 cells bearing the expression plasmid pKY206 (45) and GroES was described in Ref. 26. The GroEL fractions of a butyl-Toyopearl column (Toyo) were concentrated with ultrafiltration (Amicon) with a 100-kDa cutoff filter (YM100) and were diluted with a buffer containing 20% methanol. This concentration-dilution procedure was repeated three times. GroEL fractions were dialyzed overnight in 25 mM Tris–HCl, pH 7.5, and stored at ~20 °C after the addition of glycerol (final concentration, 25%). GroES fractions were dialyzed overnight in 20 mM Na2S, pH 7.5, and stored at ~80 °C. The mutant GroEL (L237E) was purified from *E. coli* BL21(DE3) cells bearing the pEL237E by the same procedures except that the ammonium sulfate concentration in the lysis buffer and the column buffer was 15% saturation. Proteins

---

* This work was supported by research fellowships from the Japan Society for the Promotion of Science for Young Scientists. 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.

§ To whom correspondence should be addressed. Tel.: 81-45-924-226-8503 and the Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

1 The abbreviations used are: LA, lactalbumin; rLA, α-lactalbumin from bovine whose four disulfide bonds are reduced; RP, artificial protein with random amino acid sequences; ANS, 1-anilino-8-naphtalene-sulfonate.
with random sequences, RP3-42 and RP3-45, were expressed as soluble proteins in E. coli and purified as follows. E. coli K9989 expressing the RPs. The consensus sequence of NH₂-terminal region (11 residues) is a T7 epitope tag sequence and linker sequence (6 residues). The amino acid sequence at the COOH-terminal region of multiple cloning sites (22 residues) was one of three common sequences arising from three different reading frames.

Fig. 1. Amino acid sequences of RPs. The consensus sequence of NH₂-terminal region (11 residues) is a T7 epitope tag sequence and linker sequence (6 residues). The amino acid sequence at the COOH-terminal region of multiple cloning sites (22 residues) was one of three common sequences arising from three different reading frames.

![fig1](Image)

Fig. 2. CD spectra of substrate proteins. Trace 1, 3.9 µM RP3-45; trace 2, 5.4 µM RP3-42; trace 3, 4.5 µM α-casein; trace 4, 6.8 µM holo-LA in 25 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 100 mM KCl, 1 mM dithiothreitol. The consensus sequence of NH₂-terminal region (11 residues) is a T7 epitope tag sequence and linker sequence (6 residues). The amino acid sequence at the COOH-terminal region of multiple cloning sites (22 residues) was one of three common sequences arising from three different reading frames.

![fig2](Image)

Fig. 3. Stimulation of GroEL ATPase by soluble RPs. The influence of substrate proteins and GroES on the ATPase activity of GroEL was examined. The added substrate proteins (α-casein, rLA, RP3-42, and RP3-45) are indicated above each set of columns. The ATPase activity was estimated from the initial rate of the time course.

![fig3](Image)

Jasco J-720 spectropolarimeter at room temperature. A quartz cell with a path length of 0.1 cm was used for far (195–250 nm) ultraviolet regions. The mean residual ellipticity, [θ], was calculated from the equation [θ] = θ(10 μl/cm), where θ is the observed ellipticity in millidegrees, l the path length in cm, c the molar concentration of the protein, and n is the number of amino acid residues in the sequence. Fluorescence spectra of tryptophan by excitation light at 285 nm were measured with a fluorometer (FP-777, Jasco) at 25 °C. Binding of 1-anilino-8-naphthalenesulfonate (ANS) to proteins was assessed by fluorescence (excitation 380 nm, emission 480 nm) at 25 °C. The solution contained 18 µM proteins and 50 µM ANS in buffer B (25 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 100 mM KCl, 1 mM dithiothreitol), and emission spectra (400–700 nm) were measured.

ATPase Assay—ATPase activities were assayed by measuring the amount of produced inorganic phosphate with malachite green (48). The reaction was started by the addition of ATP (final concentration, 2 mM) to the assay solution containing buffer B, 0.5 µM GroEL, and, when indicated, 1 µM RP3-42, 1 µM RP3-45, 1 µM α-casein, 1 µM rLA, and 1 µM GroES. The assay solution was preincubated for 10 min at 25 °C prior to the addition of ATP. The reactions were terminated by the addition of perchloric acid after incubations at 25 °C for 0, 5, 10, 15, 20, and 40 min. The reaction was linear as a function of time for 10 min. Thus the ATPase activity was obtained from this term. The solution was treated with a malachite green reagent, and the absorbance at 630 nm was measured.

Detection of Binding of Random Sequence Proteins to GroEL with Western Blotting—1 µl of random sequence proteins (1 mg/ml) with or without 3 µM urea was diluted to 400 µl of buffer B containing 40 µg of GroEL. Then the solution was concentrated with an ultrafiltration centrifuge device (Ultracentrifuge-MC, 100-kDa cutoff filter, Millipore) to approximately 20 µl. The concentrated solution was electrophoresed on 6% polyacrylamide gel without SDS and blotted onto a polyvinylidene difluoride membrane. Random sequence proteins were detected using alkaline phosphatase-conjugated anti T7-tag antibody. Blots were developed with nitro blue tetratoluidine and 5-bromo-4-chloro-3-indolyl phosphate (49).
FIG. 4. Binding of soluble RPs probed by intrinsic fluorescence. Panel A, Trp fluorescence emission spectra of soluble RPs. Trace 1, 0.5 μM RP3-42; trace 2, 0.5 μM RP3-45; trace 3, 0.5 μM RP3-45; trace 4, 0.5 μM RP3-45 in the presence of 1 μM GroEL. Panel B, titration of soluble RPs with GroEL. GroEL was added either to RP3-42 (closed circles) or RP3-45 (closed triangles) to the indicated concentrations. The concentrations of RP3-42 and RP3-45 were 0.56 and 0.61 μM, respectively. The fluorescence emission at 343.5 nm (RP3-42) and 342 nm (RP3-45) was corrected for background fluorescence and for change in volume. The fluorescence intensity was normalized. The solid lines represent a simulation to the data using estimated Kd values and binding stoichiometries.

RESULTS AND DISCUSSION

Random Sequence Proteins—An expression plasmid library encoding artificial genes for proteins with 138–141 amino acid residues was made as described previously (43). The nucleotide sequences were nearly random, although randomness was restricted by three factors: the appearance of stop codons was avoided, a 7T epitope tag sequence (11 residues) and linker sequence (6 residues) were introduced at the NH2-terminal region for detection of the proteins, and multiple cloning restriction sites were put into the COOH-terminal region for easy genetic handling. The amino acid sequence at the region of multiple cloning sites (22 residues) of each expressed protein was one of three common sequences arising from three different reading frames. The remaining ~90-amino acid sequence contained 20 kinds of amino acids and was nearly random. We transformed E. coli by the plasmid library, chose seven colonies expressing proteins, and deduced the amino acid sequences from nucleotide sequences (Fig. 1). Five transformants produced proteins as inclusion bodies that we solubilized in 3 M urea, designated as RP3-04, -06, -34, -54, and -64 (insoluble RPs).2 The remaining two, RP3-42 and RP3-45, were expressed as soluble proteins and purified without exposure to urea (soluble RPs). Structures of soluble RPs have been studied previously, and it was suggested that they are rather compact like the molten globule but did not contain a marked secondary structure (44). Indeed, CD spectra of soluble RPs (Fig. 2) were very similar to the spectrum of α-casein, a model unfolded protein with no secondary structure. Reference proteins, native holo-LA, and rLA in a lesser degree, showed an α-helix signal, a trough around 222 nm. RP3-42 and -45 did not have a trough at 222 nm or at 214 nm, a β-sheet signal. Another indication of disordered structures of soluble RPs was a deep trough at 190 nm (50). Exposure of the hydrophobic surface to the bulk water phase was assessed by binding of ANS. The relative fluorescence of ANS at 466 nm at the same protein concentration (18 μM) was 100% (apo-LA, a typical molten globule), 57% (RP3-45), 36% (holo-LA), and 18% (RP3-42). In 6 M guanidine HCl, all of the proteins showed no fluorescence (<5%). Taking into account the results from small angle x-ray scattering (44), it was indicated that although soluble RPs lack the secondary structure, they have a compact shape with some hydrophobic surface (RP3-45–RP3-42).

Soluble RPs Stimulated GroEL ATPase Cycle—As reported (34, 51–53), when substrate proteins such as α-casein and rLA were present, the catalytic cycle of GroEL and GroEL/GroES was accelerated, and the rate of ATP hydrolysis was enhanced.

2 In the original paper (43), RP3-04 and RP3-54 were classified as soluble. However, in this report, these were classified as insoluble because of purification from inclusion bodies.

FIG. 5. Detection of binding of random sequence proteins to GroEL with Western blotting. Panel A, Western blotting of the simple mixture of soluble RPs with GroEL (left) or peptide binding-incompetent GroEL mutant (L237E) (right). The arrowhead indicates the position of GroEL. Panel B, Western blotting of the mixtures resulted from dilution of RPs in 3 M urea into a solution of GroEL (upper) or GroEL mutant (L237E) (lower). The arrowheads indicate the positions of GroEL. In the panel B, control means GroEL only (upper) or RP3-04 diluted into wild-type GroEL solution (lower). (Fig. 3). Similarly, RP3-42 and RP3-45 activated the ATPase activity of GroEL and GroEL/GroES. One of the known features of GroEL ATPase activity is the suppression by GroES, and this feature was retained for the natural and soluble RPs. GroEL/GroES-dependent folding of rhodanese from the guanidine HCl-denatured state was inhibited when RP3-42 was present. As the amount of RP3-42 increased, the yield of recovered rhodanese activity decreased, reaching 30% of the original activity in the presence of a 27 molar excess of RP3-42 over rhodanese (data not shown). Similar inhibition was observed for rLA, a well characterized model substrate protein of GroEL. Thus, the soluble RPs compete for the same binding site on
GroEL with natural substrate protein. Based on these results we can conclude that GroEL recognizes random sequence proteins as natural substrate proteins.

**Binding of Soluble RPs Probed by Fluorescence—Because GroEL has no trypsinoid residue, the change of the intrinsic fluorescence emission spectrum of the substrate protein induced by binding to GroEL reflects solely the change of the local environment around trypsinoid residues of the substrate proteins.** The binding of RP3-42 and -45 to GroEL caused an increase of fluorescent intensities of about 1.6-fold and blue shifts of the emission maximum from 350.5 to 341.5 nm (RP3-42) and from 347 to 342 nm (RP3-45) (Fig. 4A). Such an increase of fluorescence intensity and blue shift of the maximum emission wavelength were also observed when α-casein bound to GroEL (not shown). These changes are attributable to the increased hydrophobicity around trypsinoid residues which result from interaction with hydrophobic residues of polypeptide binding sites of GroEL. We titrated the increase of fluorescence of RPs with varying amounts of GroEL (Fig. 4B). The dissociation constants (Kd) were estimated by fitting the data to a simple titration equation, and the values of 17 nm (RP3-42) and 145 nm (RP3-45) were obtained. Interestingly, the binding affinity is not parallel to the extent of hydrophobic surface probed by ANS. The real reason for this apparent discrepancy is not known, but the acidic nature of RP3-45 (pI = 4.9) relative to RP3-42 (pI = 5.5) could attenuate the binding to GroEL, which is also acidic (pI = 4.5). The binding stoichiometry was calculated to be 1.75 and 2.05 mol/mol GroEL, respectively. Binding stoichiometry over the value 2 has been reported in Ref. 54. When the mixture of GroEL and dansyl-labeled RP3-42 was subjected to gel filtration high performance liquid chromatography, fluorescence was eluted at the position of GroEL, and the prior incubation of the mixture with ATP resulted in a loss of 65% of the fluorescence (data not shown). This indicates that ATP released GroEL-bound RP3-42 from GroEL like a natural substrate protein.

**Direct Detection of Binding of Random Sequence Proteins to GroEL—The mixture of soluble RPs and GroEL was analyzed with native polyacrylamide gel electrophoresis (Fig. 5A). Both RPs were detected at the position of GroEL. The possibility that GroEL retained the T7 tag and/or multicloning sequences of RPs, rather than randomized sequence, was excluded because we confirmed that natural proteins (green fluorescent protein) with a T7 tag at the NH2 terminus and multicloning sequence at the COOH terminus did not bind to GroEL (data not shown). A GroEL mutant (L327E), in which Leu-237 was replaced by a glutamic acid residue, was shown previously to be unable to bind a natural substrate protein (18). This mutant GroEL also failed to bind soluble RPs (Fig. 5A). This ensured that soluble RPs bind to the normal polypeptide binding site of GroEL. Whether GroEL could bind insoluble RPs was examined by dilution of an RP solution in 3 M urea into a GroEL solution followed by native polyacrylamide gel electrophoresis (Fig. 5B). All of the RPs tested bound to GroEL, again failing to bind to mutant GroEL (L327E).

Taken together, the results in this report show that secondary structure is not prerequisite for the substrate protein to bind GroEL. The same conclusion can be drawn from the studies using α-casein as a substrate protein and crystal structure of minichaperone in which a peptide with a roughly extended conformation occupied the polypeptide binding site of GroEL (17). It should be added, however, that this does not exclude the possibility that some peptides that interact weakly with GroEL adopt a conformation of α-helix when binding to GroEL (37–40). It is our current view that the polypeptide binding site of GroEL may be flexible enough to accept an α-helix form of peptides that expose hydrophobic residues but that an extended form of peptides with hydrophobic residues should be a more favorable substrate.

**REFERENCES**

1. Beissinger, M., and Buchner, J. (1998) **Biochim. Biophys. Acta** 1401, 571–580.
2. Bukau, B., and Horwich, A. L. (1998) **Cell** 92, 351–366.
3. Gething, M. J., and Sambrook, J. (1992) **Nature** 356, 33–45.
4. Hartl, F. U. (1996) **Nature** 381, 571–580.
5. Clark, A. C., and Buchner, J. (1998) **Biochim. Biophys. Acta** 1411, 10–13.
6. Hartl, F. U. (1998) **Trends Biochem. Sci.** 23, 68–73.
7. Coyle, J. E., Jaeger, J., Gross, M., Robinson, C. V., and Radford, S. E. (1997) **Fold. Design** 2, R93–R104.
8. Ellis, R. J., and Hartl, F. U. (1996) **FEBS Lett.** 390, 10–13.
9. Fenton, W. A., and Horwich, A. L. (1997) **Protein Sci.** 6, 743–760.
10. Horwich, A. (1998) **Curr. Opin. Struct. Biol.** 8, 93–100.
11. Ranson, N. A., White, H. E., and Saibil, H. R. (1998) **Biochem. J.** 333, 233–242.
12. Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A., and Horwich, A. L. (1998) **Annu. Rev. Biochem.** 67, 581–608.
13. Braig, K., Otowinski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1994) **Nature** 371, 1080–1085.
14. Xu, Z., Horwich A. L., and Sigler P. B. (1997) **Proc. Natl. Acad. Sci. U. S. A.** 94, 3571–3575.
15. Fenton, W. A., Kashi, Y., Furtak, K., and Horwich, A. L. (1994) **Nature** 371, 214–219.
16. Gervasoni, P., Staudenmann, W., James, P., and Plückthun, A. (1998) **Biochemistry** 37, 11660–11669.
17. Izbakli, L. S., Oten, D. E., and Fersht, A. R. (1995) **Biochemistry** 34, 14581–14587.
18. Lin, Z., Schwartz, F. P., and Eisenstein, E. (1995) **J. Biol. Chem.** 270, 1011–1014.
19. Vinicka, A., Gervasoni, P., Zauzig, F., Ziegler, U., Lindner, P., Goverschütz, P., Pluckthun, A., and Semenza, G. (1998) **Biophys. J.** 75, 3256–3263.
20. Zahn, R., Axmann, S. E., Rucknagel, K. P., Jaeger, E., Laminet, A. A., and Plückthun, A. (1994) **J. Mol. Biol.** 240, 150–164.
21. Zahn, R., and Plückthun, A. (1994) **J. Mol. Biol.** 242, 165–174.
22. Ewalt, R. L., Hendrick, J. P., Hury, W. A., and Hartl, F. U. (1997) **Cell** 90, 491–500.
23. Sakikawa, C., Taguchi, H., Makino, Y., and Yoshida, M. (1999) **J. Biol. Chem.** 274, 21215–21256.
24. Viitanen, P. V., Gatenby, A. A., and Lerner, G. H. (1992) **Protein Sci.** 1, 363–369.
25. Chen, Y. H., Yang, J. T., and Martinez, H. M. (1972) **J. Biol. Chem.** 247, 267, 12468–12474.
26. Makino, Y., Amada, K., Taguchi, H., and Yoshida, M. (1997) **Biochim. Biophys. Acta** 1352, 38–42.
27. Mendoza, J. A., Butler, M. C., and Horwich, P. M. (1999) **J. Biol. Chem.** 274, 24648–24654.
28. Oku, T., Iikura, T., Kido, K., and Kojima, H. (1994) **Nature** 367, 419–426.
29. Braig, K. (1998) **FEBS Lett.** 428, 147–151.
30. Schneider, J., Fritsch, E. F., and Maniatis, T. (1989) **Annu. Rev. Biochem.** 58, R93–R104.
31. Sugiyama, K., Saitoh, S., and Fujita, T. (1991) **Cell** 67, 159–165.
32. Chen, Y. H., Yang, J. T., and Martinez, H. M. (1972) **J. Biol. Chem.** 247, 267, 12468–12474.
33. Jackson, G. S., Staniforth, R. A., Halsall, D. J., Atkinson, T., Halbrook, J. J., Clarke, A. R., and Burton, S. G. (1998) **Biochemistry** 37, 2554–2563.
34. Rye, H. S., Roseman, A. M., Chen, S., Furtak, K., Fenton, W. A., Saibil, H. R., and Horwich, A. L. (1999) **Cell** 97, 325–338.
35. Yифrach, O., and Horovitz A. (1996) **J. Mol. Biol.** 256, 356–361.
36. Clark, A. C., and Buchner, J. (1997) **Biochemistry** 36, 5893–5901.
GroEL Binds Artificial Proteins with Random Sequences
Katsuhiko Aoki, Fumihiro Motojima, Hideki Taguchi, Tetsuya Yomo and Masasuke Yoshida

J. Biol. Chem. 2000, 275:13755-13758.
doi: 10.1074/jbc.275.18.13755

Access the most updated version of this article at http://www.jbc.org/content/275/18/13755

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 53 references, 12 of which can be accessed free at http://www.jbc.org/content/275/18/13755.full.html#ref-list-1