Cell, Vol. 99, 757–768, December 23, 1999, Copyright ©1999 by Cell Press

The Crystal Structure of a GroEL/Peptide Complex: Plasticity as a Basis for Substrate Diversity

Lingling Chen† and Paul B. Sigler*†‡
*Howard Hughes Medical Institute
†Department of Molecular Biophysics and Biochemistry
Yale University
New Haven, Connecticut 06511

Summary

The chaperonin GroEL is a double toriodal assembly that with its cochaperonin GroES facilitates protein folding with an ATP-dependent mechanism. Nonnative conformations of diverse protein substrates bind to the apical domains surrounding the opening of the double toroid's central cavity. Using phage display, we have selected peptides with high affinity for the isolated apical domain. We have determined the crystal structures of the complexes formed by the most strongly bound peptide with the isolated apical domain, and with GroEL. The peptide interacts with the groove between paired α helices in a manner similar to that of the GroES mobile loop. Our structural analysis, combined with other results, suggests that various modes of molecular plasticity are responsible for tight promiscuous binding of nonnative substrates and their release into the shielded *cis* assembly.

Introduction

GroEL, one of the earliest and certainly the best studied of the molecular chaperonins (Ellis, 1996; Hartl, 1996; Fenton and Horwich, 1997), is an essential protein that assists the folding of a variety of proteins in *E. coli*. GroEL consists of 14 subunits of 58 kDa each that are organized into two rings stacked back-to-back, sharing seven-fold rotational symmetry (Braig et al., 1994; Boisvert et al., 1996) (Figure 1A, left panel). The two stacked rings create a central channel that is split into two functionally separate cavities at the ring interface by the protrusion of the poorly ordered N and C termini of each subunit into the channel. Each of the subunits comprises three structural domains linked by flexible hinges: an apical domain that forms the opening of the channel; an equatorial domain, located at the base of the ring that maintains all of the inter- and most of the intraring interactions; and an intermediate domain that bridges the apical and the equatorial domains. GroEL-mediated folding also requires the cochaperonin GroES, a dome-like molecule made of seven 10 kDa identical subunits (Hunt et al., 1996; Mande et al., 1996). In the GroEL/GroES complex (Figure 1A, right panel), GroES caps one end of GroEL, converting the GroEL central channel to a sealed chamber (the *cis* ring or *cis* assembly), where the protein can fold in isolation (Chen et al., 1994; Xu et al., 1997). Based upon a wealth of biochemical and structural data, the following mechanism of GroEL-assisted protein folding has been proposed (for review, Sigler et al., 1998; Xu and Sigler, 1998). Initially, GroEL captures the misfolded protein at the opening of one of the central cavities (Chen et al., 1994; Thiyagarajan et al., 1996) mainly through interactions between the presumably accessible hydrophobic residues of the nonnative polypeptide and the exposed hydrophobic residues on the channel surface of the apical domains (Fenton et al., 1994; Hlodan et al., 1995; Itzhaki et al., 1995; Lin et al., 1995), thus preventing the substrate from aggregating with other misfolded or partially folded protein molecules. Next, cooperative binding of seven ATP molecules and subsequent binding of GroES to the same GroEL ring (the *cis* ring) displaces (and possibly unfolds) the substrate from the cavity lining and initiates folding of the isolated substrate within the enclosed GroEL-GroES chamber, or the so called "Anfinsen" cage (Weissman et al., 1995, 1996; Mayhew et al., 1996; Shtilerman et al., 1999). During this process, all seven ATPs are hydrolyzed to ADPs, thereby weakening the *cis* assembly. Substrate binds to the unoccupied opposite or *trans* ring, where the binding of ATP drives the formation of a new folding chamber with the concordant disassembly of the weakened original *cis* assembly, thereby releasing the folded peptide. The now empty unliganded ring can initiate another half cycle of this "two-stroke system."

One of the most intriguing aspects of the GroEL system is its surprisingly diverse range of substrates. About 30% of *E. coli* proteins will be misfolded in a GroEL-deficient cell, and about 50% of soluble *E. coli* proteins have been shown to interact with GroEL in their nonnative states (Viitanen et al., 1992; Horwich et al., 1993). GroEL specifically binds to substrate polypeptides in nonnative conformations (Martin et al., 1991), which presumably have more accessible hydrophobic surfaces with which to interact with exposed nonpolar residues in the central cavity of GroEL (Fenton et al., 1994). Despite the wide range of biochemical and biophysical techniques that have been applied to study the substrate–GroEL interaction, no sequence specificity has been found in substrates, and the stereochemistry of the binding interaction has only recently been addressed in stereochemical terms (Chatellier et al., 1999; Kobayashi et al., 1999; Tanaka and Fersht, 1999).

The apical domain of GroEL has been implicated in substrate binding. In cryo-EM studies, electron density for the substrate is found trapped in the opening of the central cavity of GroEL formed by the apical domains (Chen et al., 1994). Small-angle neutron scattering and structure-based mutational studies also provide evidence that apical domains are involved in substrate interactions (Fenton et al., 1994; Thiyagarajan et al., 1996). Furthermore, the apical domains in the crystal structures of GroEL (Braig et al., 1994, 1995; Boisvert et al., 1996) have rather high temperature factors, indicative of either local flexibility and/or en bloc movement of the apical domain as reflected in a normal mode analysis (Ma and Karplus, 1998). This mobility has been ascribed to the

‡To whom correspondence should be addressed (e-mail: sigler@csb.yale.edu).

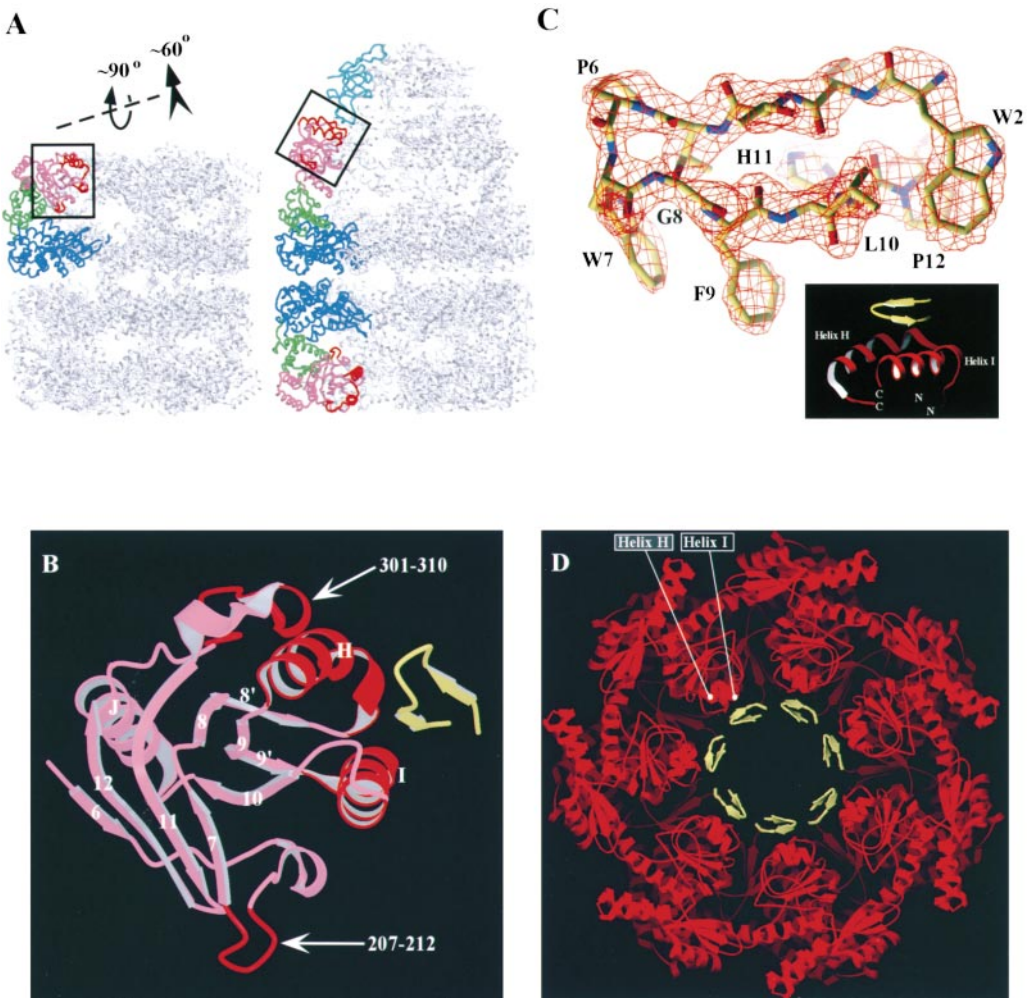

Figure 1. Structure of the Peptide/Apical Domain Complex

(A) Structures of unliganded GroEL (Braig et al., 1995) and GroEL/GroES/$(ADP)_7$ (Xu et al., 1997). Color codes for the subunits are pink for the apical domain, green for the intermediate domain, blue for the equatorial domain, and cyan for the GroES subunit. Subunits immediately in front of those highlighted are removed for clarity. The four flexible regions (see Structural Plasticity) are highlighted in red. The isolated fragment of the apical domain used in this study is enclosed in the rectangular box. Direction and magnitude of the apical domain movement within *cis* ring upon the formation of GroEL/GroES assembly are shown.

(B) The structure of the strongly binding peptide (SBP)/apical domain complex. The view of the apical domain is the same as in the left of (A). SBP is in yellow. Consistent with the secondary structure nomenclature in the tetradecameric GroEL (Boisvert et al., 1996), α helices are labeled H to K, and β strands are numbered as 6 to 12. Two additional β strands found in SBP/apical domain complex are denoted as 8′ and 9′.

(C) A difference Fo-Fc omit electron density map at 2.1 Å in the vicinity of SBP bound to the apical domain. The map, contoured at 2.5 σ, was generated using Fo, the observed structure factor of the crystalline complex; and Fc and ϕc, the structure amplitude and phases, respectively, calculated from a model in which the bound SBP was removed from the refinement. SBP is illustrated as a guide. Insert shows the position of bound SBP in relation to the binding groove formed by helices H and I.

(D) Top view of the model of the SBP/GroEL complex. To generate this model, Cα atoms of the apical domain in the SBP/apical domain complex are superimposed to the apical domains of the unliganded GroEL structure (Braig et al., 1995). This model is identical to the SBP/GroEL crystal structure within the limits of partial refinement and resolution of the latter. SBPs are colored in yellow; GroEL is in red. Helices H and I of one subunit are labeled. For clarity, only one ring of GroEL is shown.

(A) and (B) were produced with MOLSCRIPT and RASTER 3D (Kraulis, 1991; Merritt and Bacon, 1997); (C) and (D) were generated in SETOR (Evans, 1993).

structural adaptability of this domain that would be required for binding a wide range of substrates (Braig et al., 1994).

To investigate the nature of the interactions between GroEL and its substrate, the structural basis for GroEL's promiscuous substrate binding, and the mechanism of substrate entrapment, we have determined the high-resolution crystal structure of a complex formed by the apical domain of GroEL with a peptide selected by phage display technique (Scott and Smith, 1990) from a random 12-mer peptide library. For comparison, we have determined the structure of the unliganded apical domain, or the "apo" form. In addition, we solved and partially refined the complex formed by the intact GroEL and the peptide and found the peptide bound in the same manner as to the isolated apical domain. A pair

of parallel α helices in the apical domain, helix H and helix I (15 residues each), and the groove between them constitute the peptide-binding site (Figure 1B). The structure of this binding site in each of the two representations within the asymmetric unit of the apo domain is different, indicating that the open binding site is flexible. Moreover, the open binding sites are also different in structure than those seen in the crystal structures of the unliganded apical domain (Braig et al., 1994; Boisvert et al., 1996; Zahn et al., 1996). Most importantly, we have found that the binding site assumes a different and unique structure when bound to a particular peptide. This was seen in the structures of the apical domain/peptide complex studied here as well as that of the complex formed by the mobile loops of GroES with the apical domains in GroEL/GroES/$(ADP)_7$ assembly (Xu et al., 1997), and that of a complex formed in the crystal structure of Buckle et al. (1997) between the helices H and I of the apical domain and a seven-residue N-terminal extension of a neighboring molecule in the lattice. The binding site's conformation is distinct in each of these three peptide-bound crystal structures, suggesting that the flexible binding site adjusts to a different complementary conformation upon binding different oligopeptide sequences.

Results and Discussion

Biopanning Selection for High-Affinity Peptides

We selected peptides with phage display methods (Scott and Smith, 1990) that bind to the apical domain of GroEL with higher than average affinity. Table 1 lists the sequences of 41 12-residue peptides present in clones randomly taken from the enriched library left after four rounds of selection. Interestingly, five sequences are represented more than once in this pool of sequences, and one of these repeats six times, suggesting that the selection was convergent. The selected peptides (counting the repeats as one sequence) show an enrichment of tryptophan and a slight bias for leucine, relative to that found in the original peptide library (data not shown). Although some sequences have a short stretch of hydrophobic residues, particularly in the five repetitive sequences, a consensus motif is not immediately obvious. Despite attempts to saturate the resin with the His_6-tagged apical domain, ten peptides may have been selected due to their interactions with exposed Ni^{2+}–NTA complexes on the resin, as there are either at least two consecutive histidine residues or three or more nonconsecutive histidine residues in the sequence. These sequences were therefore excluded in the above analysis.

Binding Affinity Measured by Fluorescence Anisotropy

The affinity of seven peptides for the GroEL apical domain was determined by fluorescence anisotropy. One peptide (Type 1, represented in 2 of the 41 clones), called the "strongly binding peptide" (SBP), displayed a significantly higher affinity for the apical domain than the other peptides tested. The dissociation constant (K_D) was 2.0 μM at 20°C in 150 mM NaCl (Figure 2A) and dependent slightly on temperature (1.4 μM and 6.0 μM at 4°C and 37°C, respectively). The K_D varied with the concentration of salt in a manner implying a significant contribution from hydrophobic interactions. The affinity of SBP for the apical domain was enhanced about 2-fold at higher salt concentration (500 mM) and reduced to almost half at lower salt concentration (10 mM) (data not shown). The affinity of six additional peptides was tested, including the other four sequences selected more than once by phage display, a peptide corresponding to the GroES mobile loop, and a peptide corresponding to the near N-terminal fragment that interacts with the helix H/helix I region of a neighboring molecule in the crystal structure of Buckle et al. (1997). None of these peptides bound with a K_D less than 200 μM (the upper limit of our fluorescence anisotropy assay). Three peptides (peptide 3 and 5 on Table 1, and the GroES mobile loop) showed evidence of binding, but the inflection point of the assay could not be reached. However, these peptides also showed a similar affinity to hen egg white lysozyme. Only the SBP showed a remarkable preference (>150-fold) for the apical domain over hen egg white lysozyme.

Table 1. Sequences of Randomly Selected Peptides after Four Rounds of Biopanning Selections

Type	Sequences
1	SWMTTPWGFLHP
	SWMTTPWGFLHP
2	FHYEIWIPPHRG
	FHYEIWIPPHRG
	FHYEIWIPPHRG
	FHYEIWIPPHRG
	FHYEIWIPPHRG
	FHYEIWIPPHRG
3	SSPWWLVSFTST
	SSPWWLVSFTST
	SSPWWLVSFTST
4	SHSLIWRIPLLH
	SHSLIWRIPLLH
5	IYVPWYYAENLP
	IYVPWYYAENLP
6	YNYSWNGVVFVP
7	AQSTPLMKPQKS
8	DQTTLQRFLGSH
9	QTIKPPITVHPS
10	QYNHILGYLPFQ
11	IMDPQNSKVTVA
12	LPIQNAKRSMVS
13	IMSPWDESFWNY
14	ASESYVLFPGTR
15	SNWHGPLSYQLM
16	ALPLQDTAATLS
17	QEIYLTPRGPQQ
18	IDRTQMWRQSDL
19	INRDHPLHAGQP
20	HQTPQSLARWSL
21	HSLRAIQLITGM
22	LPSHHHHRVPAA
23	IPTYHHHHPSLR
24	QMTHHHTHRPPI
25	DLHSHHHGHMNH
26	SMHHHHRPASPT
27	WIGDAKSSLHHA
28	HNHPHTTSHVSM
29	HNSIIYHWHTLP
30	HFNHNHRGFHLI
31	AASPHYSSSHSH

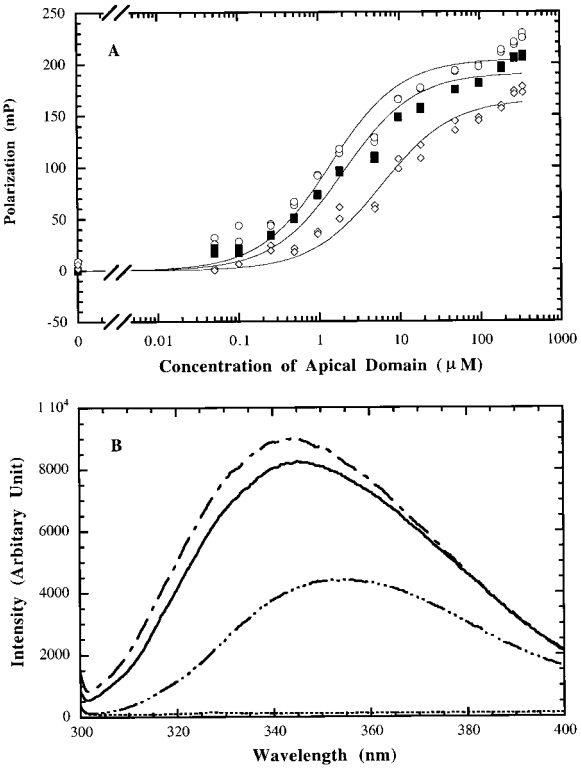

Figure 2. Effects of the Apical Domain on SBP in Fluorescence Anisotropy and Tryptophan Fluorescence Spectroscopy

(A) A fluorescein tag was attached to the N terminus of the peptide via a three-glycine spacer. The peptide concentration was constant (1 nM) throughout the experiments, and the polarization signals were measured as a function of the concentration of the apical domain. Multiple anisotropy measurements were taken at each temperature to assess the experimental errors; at 4°C, open circles; at 20°C, filled squares; and at 37°C, open diamonds. The lines are the fits of data to the single binding mode.

(B) The experiments were conducted near saturation condition, in which SBP was 30 mM. Fluorescence spectra from the buffer (dotted line); SBP (line with three dots); the mixture of SBP/apical domain in 1:1 molar ratio (continuous line); the mixture of SBP/apical domain in 1:5 molar ratio (line with single dots).

Interactions Monitored by Tryptophan Fluorescence Spectroscopy

The apical domain contains no tryptophan, while SBP has two; therefore, tryptophan fluorescence can reflect changes in the peptide's environment upon binding to the apical domain. Under near saturating condition, when the apical domain was added to a solution containing SBP, tryptophan fluorescence intensity increased 2-fold, and the fluorescence maximum shifted from 356 nm to 345 nm (Figure 2B). Both the enhancement of the intensity and the blue shift of the peak indicate that one or both tryptophan residues of the peptide were excluded from solvent when bound to the apical domain.

Crystal Structures

Crystals of the complex of the apical domain with either SBP or its fluorescein-labeled counterpart were obtained at 19°C by vapor diffusion and diffracted to 2.1 Å and 2.0 Å, respectively. Both crystals belong to space group $P2_1$, with four complexes per asymmetric unit, and are nearly isomorphous (Table 2). For comparison, the apical domain of GroEL alone (the "apo" form) was also crystallized (in the presence of a different salt; see Experimental Procedures) but in space group $P3_221$, with two molecules per asymmetric unit. These crystals also diffracted to 2.0 Å. All three structures were solved by molecular replacement. Table 2 summarizes the data collection and refinement statistics for these three structures.

As expected, the overall fold of the apical domain does not change upon binding SBP. As in all other structures, the apical domain mainly consists of three helices (helix H, I, and J) and seven β strands (Figure 1B). The major structural difference in the apical domain between the SBP complex and the other structures (Braig et al., 1994; Boisvert et al., 1996; Zahn et al., 1996; Buckle et al., 1997; Xu et al., 1997) occurs in four regions (see Structural Plasticity), including the SBP-binding site formed by a pair of parallel helices, helix H and helix I. This pair of helices is at the ends of the tetradecameric GroEL, facing into the openings of the central cavity; collectively, seven pairs of helix H and helix I from seven apical domains form a ring surrounding each opening of the central cavity (Figure 1D).

After one round of simulated annealing and group B factor refinements, the electron density map of the SBP/apical domain complex gave an unambiguous image of the bound peptide's backbone and recognizable density for most of the expected side chains. The peptide was readily fit and refined together with the apical domain. An omit difference electron density map of the SBP-binding region is shown in Figure 1C. The peptide adopts a β hairpin conformation over a groove formed by a pair of parallel helices H and I, making a β turn (Thr-Pro-Trp-Gly), stabilized by five intrapeptide hydrogen bonds (Figure 3B). Analysis of the torsion angles reveals that the peptide is bound without strain. Both the N and C termini of SBP face outward to the solvent, exiting the binding site in a manner consistent with the peptide being an accessible part of a bound nonnative protein. All components of the unlabeled SBP are clearly seen in the structure except for the N-terminal serine residue, which is observed only in one of the four representations in the asymmetric unit. The fluorescein-tagged peptide has the same structure as the untagged counterpart. The fluorescein moiety and the three-glycine linker are not visualized in the crystal structure. As the crystals appear bright yellow, we must assume that the tag is present but not bound in an ordered manner and that it does not interfere with, or contribute to, the peptide's mode of binding.

Crystals of SBP bound to tetradecameric GroEL in the absence of any nucleotides were also obtained, and data were collected to 3 Å. Preliminary refinement (R_{work} = 26%, R_{free} = 31%) shows that SBP binds with the same β hairpin conformation to each of the 14 subunits with unit occupancy in the same manner as in the isolated apical domain complex. Comparison of the refined structure of the SBP-bound tetradecamer and that of the orthorhombic unliganded GroEL tetradecamer refined by Braig et al. (1995) revealed no peptide-induced changes other than those seen in the SBP complex of the isolated apical domain. Smaller but noticeable differences are noted around residues 384

Table 2. Structure Determination and Refinement

	Complex of SBP/Apical Domain	Complex of Fluorescein-Tagged SBP/Apical Domain	"Apo" Apical Domain
Data Collection Statistics			
Space group	$P2_1$	$P2_1$	$P3_221$
Unique cell dimension (Å, degrees)	a = 42.042, b = 134.668 c = 59.734, β = 94.381	a = 43.127, b = 134.712 c = 60.142, β = 94.533	a = b = 84.859, c = 77.068
No. of molecule per asymmetric unit	4	4	2
Resolution	2.1 Å	2.0 Å	2.0 Å
Observed/unique reflections	231,895/37,111	422,667/44,965	22,0421/19,371
Completeness (last shell) (%)	97.2 (89.9)	97.4 (94.3)	90.9 (96.9)
[a]R_{sym} (last shell)	0.077 (0.303)	0.074 (0.496)	0.064 (0.327)
I/σ (last shell)	18.6 (4.1)	18.0 (2.9)	16.7 (4.9)
Refinement Statistics			
No. of reflections (working, test)	34,114/1,783	41,576/2,191	16,635/1,815
R_{work}/R_{free} (all data) (%)	21.5/26.5	24.6/28.7	22.5/24.1
Rms deviation from ideality			
Bond length (Å)	0.007	0.007	0.006
Bond angles (°)	1.32	1.31	1.32
Dihedral angles (°)	23.6	24.0	23.6
B factors (rmsd of bonded atoms-main/side chain)	2.26/3.76	2.38/3.85	1.62/2.83
No. of protein atoms in the final model	4,722	4,684	2,174
No. of H_2O in the final model	96	105	188

[a] $R_{sym} = \Sigma_{h'} |I_h - I_{h'}|/\Sigma_{h'} I_{h'}$, where $|I_h - I_{h'}|$ is the absolute deviation of a reflection $I_{h'}$ from the average I_h of its symmetry and Friedel equivalents.

(intermediate domain) and 338 (apical domain). Both of these regions are at the end of long α helices and reflect the en bloc variation in the orientation of the domain rather than local conformational adjustments to the bound peptide.

SBP/Apical Domain Interface

The center of the binding site consists of confluent hydrophobic pockets in the groove between helices H and I (Figure 4B) surrounded by a surface rich in polar and charged moieties. The SBP peptide takes advantage of this bimodal binding site by burying two bulky hydrophobic residues, Trp-7 and Phe-9, in two hydrophobic pockets and forming a network of hydrogen bonds between the backbone of the peptide and surrounding polar surface of the apical domain (see Figure 3 for details). If SBP were to adopt the same β hairpin conformation free in solution, the SBP/apical domain interaction would bury ~930 Å^2 of solvent-accessible molecular surface: ~430 Å^2 from the apical domain and ~500 Å^2 from the peptide. In fact, the buried molecular surface of the peptide, and thus the complex interface, is much greater because the peptide was found to be unstructured in solution by 1D or 2D NMR studies (data not shown); that is, the hairpin conformation, to a certain extent, is imposed on the bound peptide. The formation of the β hairpin would bury an additional 470 Å^2 upon binding, assuming (albeit unrealistically) a fully extended peptide in solution. Overall, the SBP forms a tight interface with the apical domain that is complementary both in shape and distribution of polar and nonpolar surfaces.

The mode by which SBP binds in the crystal structure is consistent with other binding experiments in solution. The importance of Trp-7 and Phe-9 in binding has been demonstrated by the fact that changing either Trp-7 or Phe-9 to an alanine leads to at least a 20-fold decrease in the affinity of the altered SBP for the apical domain (data not shown). In addition, the burial of Trp-7 in the hydrophobic pocket explains the changes in tryptophan fluorescence upon peptide binding.

Comparisons of Several Different Peptide Binding Modes

Interestingly, two other crystal structures reveal peptides bound to helices H and I of the apical domain. In the structure of the GroEL/GroES/$(ADP)_7$ (Xu et al., 1997), the GroES mobile loops (residues 18 through 30) interact with helices H and I. Accordingly, a peptide corresponding to this GroES mobile loop was synthesized. Fluorescence polarization indicates a detectable but weak interaction of this individual 13-residue mobile loop with the isolated apical domain ($K_D > 200$ μM) (data not shown); however, the affinity of the seven GroES mobile loops to the interhelical grooves of the seven apical domains of GroEL is almost certainly enhanced substantially by the perfect stereochemical match between the stable rotationally symmetrical assembly of seven GroES subunits and the ring of seven apical domains in the *cis* assembly of the GroEL/GroES complex. In another example of a crystalline peptide complex with the apical domain, Buckle et al. (1997) found seven residues near the N terminus binding in an extended conformation between helices H and I of a neighboring apical domain in the crystal lattice. The stability of this interaction is clearly enhanced by fortuitous crystal packing, as this segment in isolated form shows no detectable affinity for the apical domain by our fluorescence polarization assay (data not shown). In summary, steric restraints, either a rotationally symmetrical seven-fold cooperative alignment or crystal packing proximity, enhance the affinity of the GroES mobile loops and the near N terminus

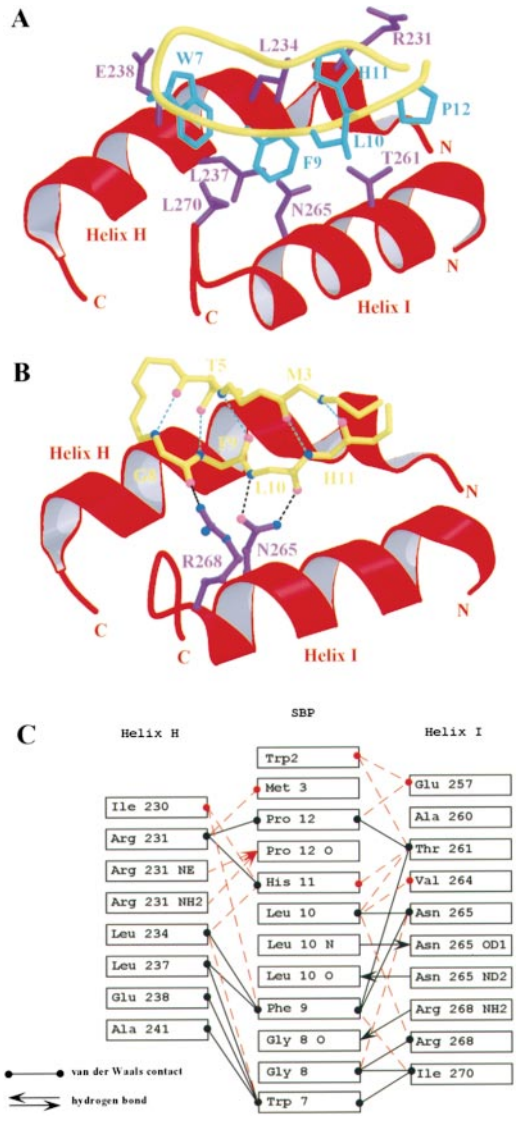

Figure 3. Interactions between SBP and the Apical Domain

SBP is in yellow, helices H and I of the apical domain are in red, side chains of SBP are shown in cyan, and side chains of the apical domain are in purple.

(A) Residues involved in significant van der Waals contact are shown along with their side chains.

(B) Hydrogen bonding network in SBP. Black dotted lines are hydrogen bonds between SBP and the binding site in the apical domain, and cyan dotted lines are the intrapeptide hydrogen bonds. Oxygen and nitrogen atoms are represented in pink and blue, respectively.

(C) Diagram of the interactions between SBP and helices H and I. Black solid lines denote the molecular contacts (atoms within 4 Å) observed in all four complexes within the asymmetric unit of the crystal, and red dashed lines indicate the additional contacts in some but not all four complexes.

(A) and (B) were produced using MOLSCRIPT and RASTER 3D (Kraulis, 1991; Merritt and Bacon, 1997).

peptide segment, respectively, to the SBP-binding site in the apical domain, enabling us to study their binding crystallographically.

Comparisons of the binding modes of three peptides—SBP, the GroES mobile loop, and the apical domain's N-terminal extension—reveal some common characteristics as well as distinct features of interactions between helix H and helix I of the apical domain and the peptide segments. First, although three peptides adopt different conformations (a classical β hairpin for SBP, a loop for the GroES mobile loop, and an extended conformation for the N-terminal extension), the overlapping segments of three interacting peptides provide most of the binding surface and share the same polarity; that is, they run unidirectionally along the binding groove (Figure 4A). Second, in all three structures, two hydrogen bonds connect the side chain of Asn-265 in the apical domain to the backbone of an adjacent residue in the bound peptide (for example, Leu-10 of SBP; Figure 3B). Third, the residues that form the most extensive side chain interactions with the central hydrophobic groove are all hydrophobic in three peptides. Nevertheless, SBP distinguishes itself by making the most intimate contact with the apical domain as noted above. SBP occupies the binding pockets in the central groove with bulky hydrophobic residues Trp-7 and Phe-9 (Figure 4B), each of which contributes about 6 and 5 kcal/mol, respectively, to the affinity, assuming a hydrophobic effect of 47 cal/mol/ Å^2 (Sharp et al., 1991). In contrast, only one residue with a smaller side chain, Val-26 from the GroES mobile loop or Leu-185 from the N-terminal extension, occupies the same binding pockets (Figures 4C and 4D). Overall, SBP occludes the most surface area of the apical domain (∼430 Å^2 by SBP, ∼380 Å^2 by the GroES mobile loop, and ∼330 Å^2 by the N-terminal extension). Finally, SBP in contrast to the more weakly bound peptides forms five intraloop hydrogen bounds and buries at least 400 Å^2 of solvent-exposed surface when bound. Thus, the larger interaction surface, greater shape complementarity, and internal stabilization of the bound conformation will contribute to a much higher affinity of SBP to the apical domain. It would seem beyond coincidence, however, that the three bound conformations that have been observed crystallographically so far all adhere with the same polarity and hydrogen bonding pattern to the same surface of the apical domain. In addition, Fersht and coworkers, using transferred NOE experiments, have very recently implicated this same region as a peptide-binding site, in this instance for an α helix (Kobayashi et al., 1999). It is highly likely that the surface comprising the H and I helices is a preferred substrate-binding site.

Structural Plasticity

The apical domain in the noncomplexed or apo form displays structural flexibility as demonstrated by significant backbone variations between the two molecules within the asymmetric unit of the crystal (Figure 5A). Structural deviations are clustered mainly in four regions: segment 207–211, segment 301–311, helix H (residue number 230–244), and helix I (residue number 254–268). Segment 207–211 is near the region where residues (Tyr-199, Ser-201, Tyr-203, and Phe-204) were shown to be important in substrate binding by mutagenesis (Fenton et al., 1994). Helices H and I contains all the other crucial residues (Leu-234, Leu-237, Leu-259, Val-263, and Val-264) implicated by mutagenesis in substrate binding. Except for these four flexible regions, the two apo apical domains superimpose very well. (Eliminating the four flexible regions, the rms deviations of

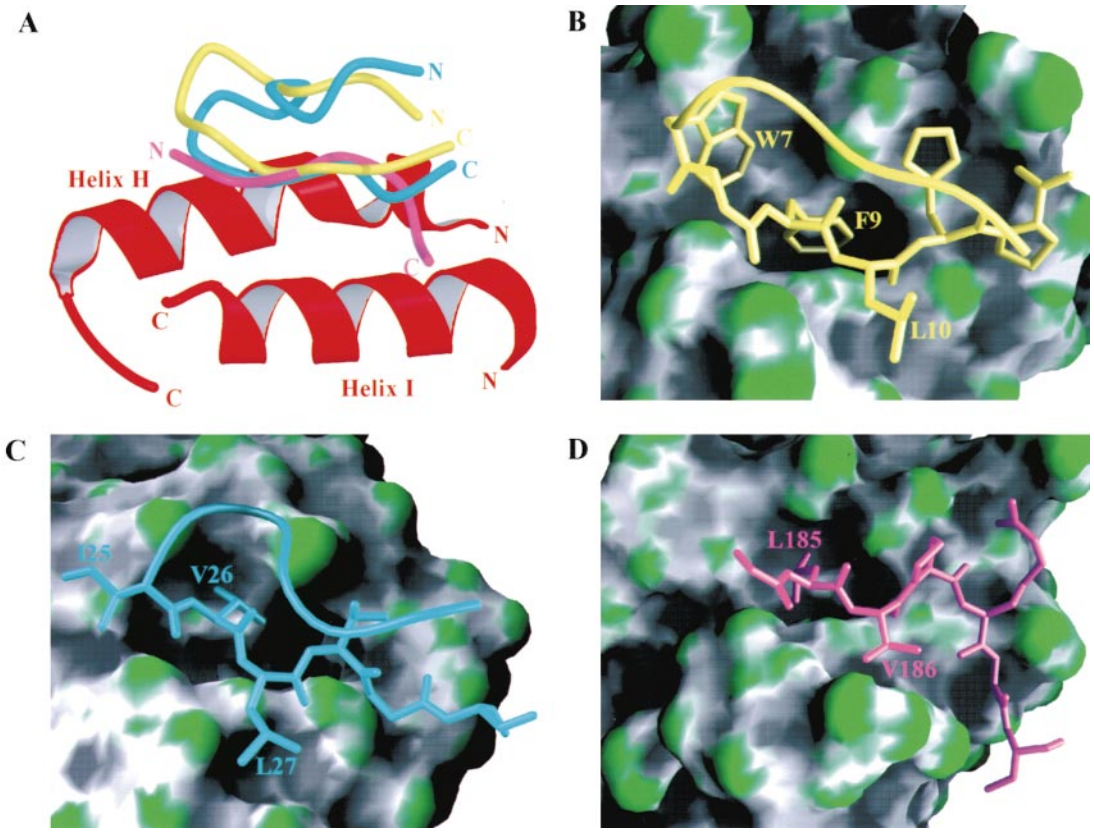

Figure 4. Structural Comparisons of Three Peptides Interacting with Helices H and I of the Apical Domain

SBP is yellow, the GroES mobile loop (Xu et al., 1997) is cyan, the N-terminal extension of the apical domain (Buckle et al., 1997) is magenta, and helices H and I are red.

(A) Superposition of Cα coordinates of the apical domain of three structures, showing the backbone of three different peptides bound over the peptide-binding groove formed by helix H and helix I. Structure of the helices displayed here is taken from the structure of the SBP/apical domain complex.

(B–D) Molecular surfaces color coded by curvature (green for convex, and gray for concave) of the binding groove in SBP/apical domain, GroEL/GroES/$(ADP)_7$, and N-terminal extension/apical domain, respectively. The orientation in these three figures is the same as in (A). For clarity, only side chains of residues located at the C-terminal arms of the β turn of the SBP (starting from W7) and the GroES mobile loop (starting from I25) are shown, as these segments form most of the contacts with the binding site. The N-terminal arms of the β turn of these two peptides are shown as a Cα trace. Residues in the peptides that form extensive side chain interactions with the binding site are labeled. (A) was produced using MOLSCRIPT and RASTER 3D (Kraulis, 1991; Merritt and Bacon, 1997), and (B)–(D) were generated with GRASP (Nicholls et al., 1991).

corresponding Cα atoms are 0.2 Å, within the estimated coordinate error 0.3 Å indicated by the Luzzati plot.)

In contrast, structures of the four SBP/apical domain complexes within the asymmetric unit of the crystal superimpose very well (Figure 5A), with rms deviations of *all* Cα atoms of ~0.3 Å. In particular, little conformational variation is found around the peptide-binding site formed by helices H and I. The common structure of the SBP/apical domain complex differs from that of the two apo forms in all four of the domain's flexible regions (Figure 5B). Thus, it appears as if helices H and I have adjusted to form a common binding site for the SBP. In addition, segments 207–211 and 301–311, which do not interact directly with SBP, also adjust locally to a new common structure in the SBP complex. Apparently, the flexibility of these two segments indirectly contributes to the plasticity of the peptide-binding site and is fixed in a specific way upon binding a peptide. It should be pointed out here that the average temperature factors of these four regions in both the apo and the peptide-bound domains are essentially the same as the overall temperature factors for their respective crystal structures. Thus, these regions are plastic enough to assume various conformations but not freely mobile. In summary, four flexible regions in the apical domain, including the peptide-binding site formed by helix H and helix I, adjust their backbone structure in response to the bound peptide and thereby reduce their conformational variability.

Previously determined crystal structures of the isolated apical domains (Zahn et al., 1996; Buckle et al., 1997) show the same flexible segments; that is, they differ from one another in the same four regions described above. These two previously reported structures of the apical domain also differ from both our apo forms and the common structure of the four SBP complexes in the same four flexible regions (Figure 5B). The N-terminal extension bound to the neighboring molecule in

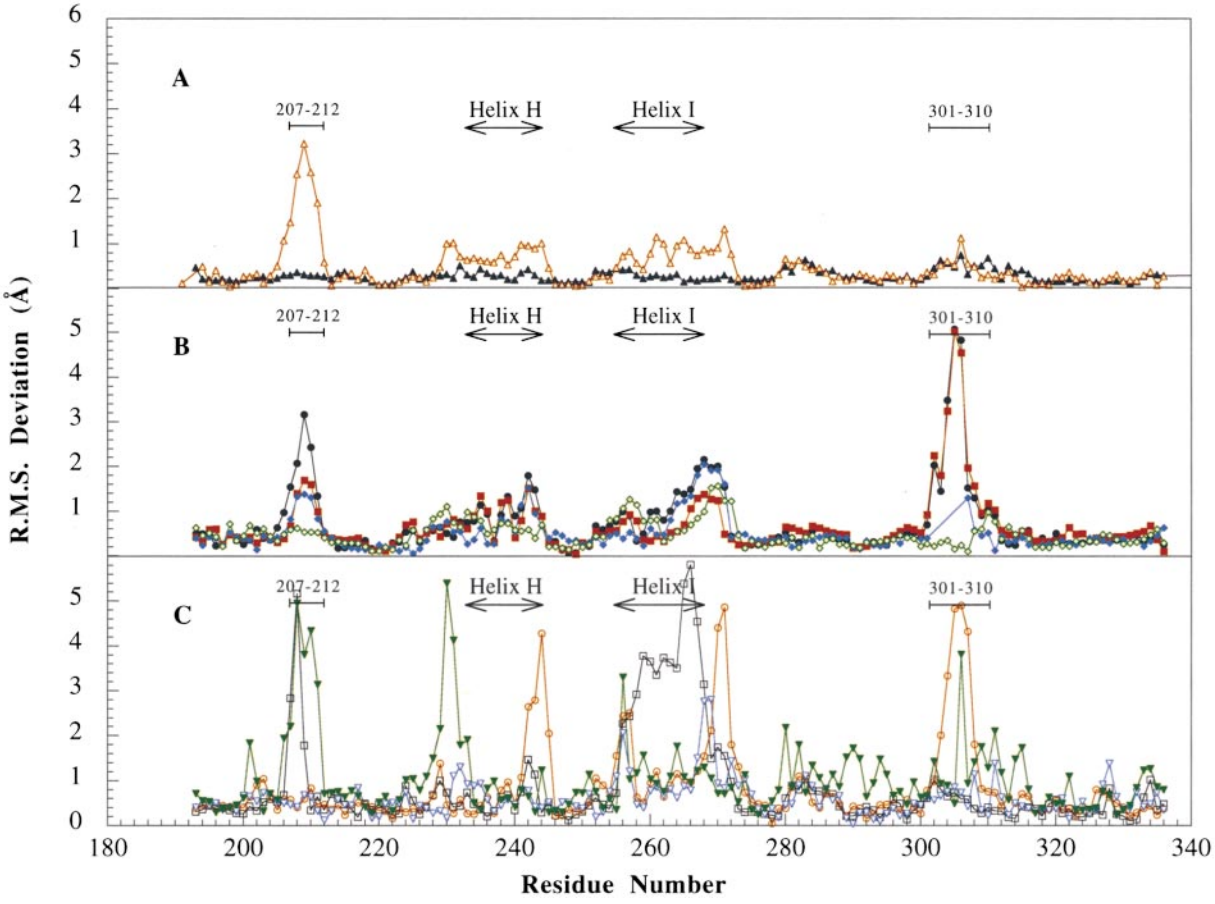

Figure 5. Plasticity within the Apical Domain

Plots of rms deviation of corresponding Cα atoms used to compare the structure of the apical domain in various functional states and crystalline systems.

(A) Comparison of the four representations of the SBP/apical domain complex in the crystalline asymmetric unit, filled triangles; comparison of the two representations of the apo apical domain in the crystalline asymmetric unit, open triangles.

(B) Comparison of the SBP-bound apical domain and each of the two apo domains determined in this study, filled circles and filled squares; comparison of the SBP-bound apical domain with the apo form of the apical domain determined by Zahn et al. (1996), filled diamonds; and with the apical domain in complex with the seven residue N-terminal extensions of a neighboring molecule in the lattice (Buckle et al., 1997), open diamonds.

(C) Comparison of the SBP-bound apical domain with apical domains of unliganded tetradecameric GroEL (Braig et al., 1995), open squares; with the apical domains of GroEL/$(ATP\gamma S)_{14}$ (Boisvert et al., 1996), open circles; and with the *trans* and *cis* rings of GroEL/GroES/$(ADP)_7$ (Xu et al., 1997), open and filled inverted triangles, respectively.

the crystal leads to a conformation of the binding site that differs from that of the apo and SBP-bound forms of the apical domain, again reflecting the flexibility of the binding site. It is important to note that the variations in structure occur only in these four regions and that the rest of the apical domain's structure is well conserved in all cases. It would appear that different peptides impose different structures on the apical domain's binding surface. These same segments also appear to be the most flexible in GroEL and GroEL/GroES complex (Figure 5C).

Figure 6 shows that the flexible peptide-binding grooves of seven subunits face the opening of the GroEL central cavity where nonnative protein substrates are bound. Together, they form what appears to be an adhesive ring lining the opening in which each of the flexible binding sites can adjust individually to accommodate various segments of the nonnative polypeptide. Thus, the structural plasticity in the peptide-binding site described in this study may be correlated with the broad substrate spectrum of GroEL.

The intrahelical groove is reminiscent of the peptide-presenting site of the MHC molecules. In their analysis of the structures of viral peptides complexed with MHC class I, Fremont et al. (1992) and Madden et al. (1993) propose that adjustments of the two somewhat longer α helices that form the peptide-binding groove account for the capacity to present a diverse set of extended peptides. In a like manner, upon the entry of the substrate into the central cavity, helix H and helix I adjust their conformations to optimize their interactions with the accessible oligopeptide segments presented by the nonnative protein to a ring of seven apical binding sites, thus allowing GroEL to bind with the entire substrate molecule through multiple sites. In addition to subtle

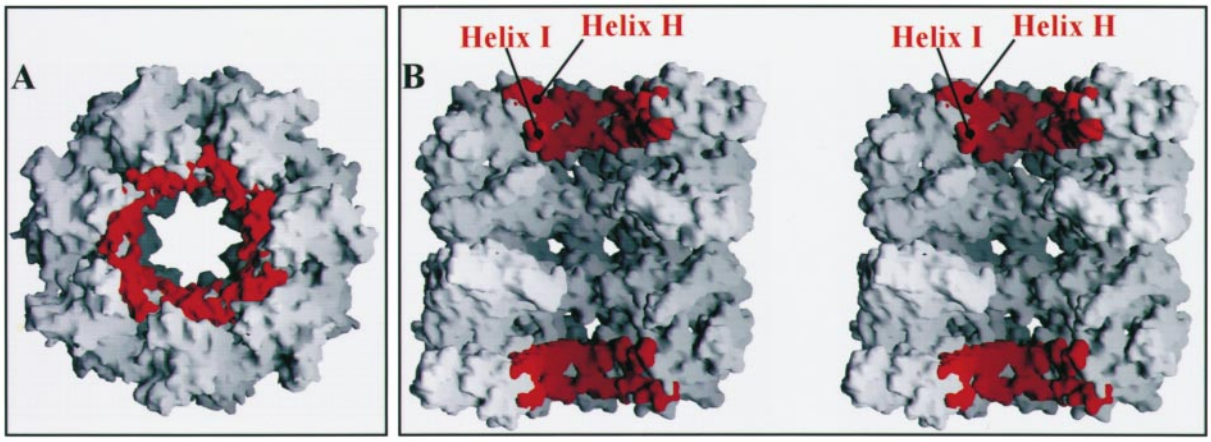

Figure 6. Molecular Surface Representation of the Substrate-Binding Sites in GroEL

The molecular surface of the binding sites formed by helices H and I is highlighted in red. The binding sites form "elastic rings" located on the opening of the GroEL central cavities.

(A) Top view of GroEL.

(B) Stereo view of the central cavities of binding-competent GroEL.

The three subunits from each of the rings nearest the reader were removed to show the inside of the central cavities. Figures were generated in GRASP (Nicholls et al., 1991).

modifications on the binding surface of the apical domain, the highly dynamic en bloc movements of the apical domain and that of the nonnative substrate protein itself add further degrees of structural adaptability to the GroEL–substrate interaction. Therefore, GroEL may acquire substrate promiscuity through the following: (1) local conformational adjustments of a flexible peptide-binding site formed by helices H and I, and two supporting regions; (2) en bloc movements of the apical domains; and (3) the inherent structural dynamics of the nonnative substrate.

Implications for Binding Nonnative Proteins

The mode of peptide binding revealed in this study may reflect features common to the binding of a partially folded or misfolded substrate with GroEL. First, the nature of GroEL/substrate interaction is mainly hydrophobic. Hydrophobic regions of SBP (consisting of Trp-7, Phe-9, and Leu-10) complement the hydrophobic pockets of the binding site formed by helices H and I. Such an interaction surface can select for the hydrophobic residues characteristically accessible in a nonnative protein, thus favoring nonnative over properly folded proteins as substrates for GroEL. The nonpolar contacts are supplemented in all cases by hydrogen-bonded contacts between flexible polar side chains of the binding site and the backbone amide function of the bound peptide. Since all extended strands, irrespective of sequence, must present to the binding surface the polar groups of the peptide backbone, these polar interactions are an ideal mechanism for supplementing affinity in a sequence-independent manner.

Second, substrate is captured in the central cavity of GroEL through multiple attachments. The observation that both termini of the bound SBP project into the central cavity of GroEL tetradecamer is consistent with the binding of an accessible element of misfolded or partially folded protein substrate trapped in the opening of the central cavity (Braig et al., 1993; Chen et al., 1994; Thiyagarajan et al., 1996). While a single peptide–apical domain interaction may not be strong enough to capture substrate, a nonnative protein may present several such accessible elements. Once one of these interactive elements is bound to one binding site, even transiently, the restraint enhances the probability that other substrate segments will interact with one or more of the remaining six binding sites. While affinity may be enhanced cooperatively by the increased likelihood of a subsequent encounter with a binding site, restriction of the transiently bound substrate's degrees of freedom might impede the exact fit of substrate normally required for productive binding. This likelihood is offset, however, by the local and global plasticity of the apical domains discussed above, coupled with the inherent dynamic flexibility of a nonnative substrate. Together, these factors enable appropriate structural adjustments to take place that are conducive to cooperative binding of the restrained polypeptide. In addition, this cooperative model of multiple flexible attachments allows the unbound portions of the polypeptide chain to maintain some regional structure while bound to GroEL. Evidence for residual structure of the substrate complexed with GroEL has been demonstrated (Martin et al., 1991; Schmidt and Buchner, 1992; Hayer-Hartl et al., 1994; Hlodan et al., 1995).

Third, local structure can be imposed on the substrate when bound to GroEL. Unstructured SBP free in solution (by 1D and 2D NMR, data not shown) assumes the form of a well-ordered β hairpin upon binding to the apical domain. Our results are consistent with the previous observations of Landry et al. (1992) that, in addition to providing an Anfinsin cage for substrate folding, GroEL can promote some secondary structure formation in nonnative substrate.

Finally, our results substantiate the mechanism by which the bound substrate is released from the inner

wall of GroEL into the central cavity, where substrate folding takes place. Previous work (Xu et al., 1997) showed that the residues of the apical domain implicated mutagenically (Fenton et al., 1994) in binding of nonnative substrates to GroEL are removed from the central cavity upon forming the stabilizing interfaces of the newly formed and hydrophilically lined folding chamber, or Anfinson cage. In this way, the peptide is "peeled" off its binding site and perhaps unfolded in the process (Shtilerman et al., 1999). A key component of this step is the interaction of the GroES mobile loop with the binding site of SBP. The "competition" between the bound peptide and the GroES mobile loop may be a misleading oversimplification. While GroES does share the same binding site with the peptide substrate, the relative orientation of these sites upon binding of the nonnative protein is radically different than that when binding GroES. Hence, in the formation of the *cis* assembly, this reorientation alone will displace the bound nonnative segments well in advance of the binding of the GroES mobile loop. Indeed, it is strain and presumed distortion of the nonnative peptides caused by this elevation and twist of helices H and I that may account for the unfolding of misfolded substrate, which is likely to occur prior to its release. This study shows directly that the binding surface formed by helices H and I, previously implicated by mutagensis, is indeed an important part of the binding and release mechanism that characterizes the GroE folding cycle.

Experimental Procedures

Expression and Purification of the GroEL Apical Domain

The gene encoding the GroEL apical domain (residues corresponding to 191–336 in full-length GroEL) was subcloned from *E. coli groE* cDNA (Xu et al., 1997) using the polymerase chain reaction (PCR), into pET15b (Novagen) NdeI and BamHI sites with the following primer sequences: 5′-GCCGATATACATATGGAAGGTATGCAGTT CGAC-3′ and 5′-GCAGCCGGATCCTCATCAACGGCCCTGGATTG CAGC-3′. The apical domain protein with a six-histidine tag at the N terminus (−21 MGSSHHHHHHSSGLVPRGSHM −1) was expressed in the *E. coli* strain BL21(DE3) (Novagen). Cells were grown at 37°C to an optical density of 0.8 at 600 nm and induced with 1 mM IPTG for 3–4 hr. Cells were lysed at 0°C by microfluidizer in 50 mM TrisCl (pH 8.2) 300 mM NaCl, 10 mM imidazole. Protein purification was performed at room temperature. The lysate was cleared by centrifugation and loaded onto a nickle-nitrilotriacetic acid (Ni^{2+}-NTA) column (Qiagen). The apical domain was eluted from the column with an imidazole gradient of 20–300 mM at ~100 mM imidazole. Protein was concentrated with an Amicon YM10 (Amicon) and buffer exchanged using Centriprep-10 (Amicon) to 150 mM NaCl, 50 mM TrisCl (pH 8.2), before loading onto a Superdex-75 size exclusion column (Pharmacia). The resulting protein was at least 95% pure, as assessed on silver-stained SDS–PAGE. Yields were typically ~15 mg of purified protein per liter of cell culture.

Biopanning Selection of Affinity Peptides

The selection procedure was modified from a protocol provided by New England Biolabs (NEB). A 12-mer peptide phage library (complexity of 10^9) was purchased from NEB. The observed amino acid distribution of this original library is close to the expected occurrence taking into account the library's codon frequency, except that Pro and Thr are overrepresented about 2-fold, and that Arg, Gly, and Cys are underrepresented by at least 2-fold (NEB). The apical domain was immobilized on Ni-NTA resin via a His_6-tag at the N terminus of the protein. Approximately 100 μl of Ni^{2+}-NTA resin, saturated with the His_6-tagged apical domain, was incubated with 10 μl of the original phage library (2×10^{11} pfu) in 1 ml of 50 mM TrisCl (pH 7.5), 150 mM NaCl (TBS) at 4°C with a Nutator for 1–2 hr. The column was extensively washed ten times with 1 ml 0.1% Tween-20, 30 mM imidazole in TBS to minimize the weak nonspecific interactions between the peptides displayed on phage surface and the apical domain immobilized on Ni^{2+}-NTA resin, and to disrupt the possible interactions of the histidine residues on the displayed peptides with the Ni^{2+} on the resin. The bound phage, still associated with the His_6-tagged apical domain, were then eluted two times with 1 ml 300 mM imidazole in TBS after a 15 min incubation at room temperature. The above procedure concluded one round of selection. The phage bound to the eluted apical domains were amplified by infecting log-phased F^+ *E. coli* culture (strain ER2537 from NEB), and the amplified phage library was purified by NaCl/PEG precipitation. For the subsequent rounds of selection, approximately the same number of selected amplified phage (2×10^{11} pfu) was input in the incubation with the immobilized apical domains. In the second, third, and fourth rounds of selection, the wash condition became increasingly more stringent by increasing the concentration of Tween-20 to 0.3%, 0.5%, and 0.7%, respectively. At the end of the fourth round of selection, a total of 41 phage were randomly chosen, amplified, and their DNA were purified. DNA segment encoding the peptide was sequenced.

Fluorescence Polarization Measurements

The measurements were performed using a Beacon 2000 instrument (PanVera). The excitation and emission wavelengths for fluorescein are 345 nm and 495 nm, respectively. Fluorescein-tagged peptides were dissolved in buffer solutions to a final concentration of 1 nM. Appropriate amounts of apical domain stock solution (2.34 mM) were added to a series of 100 μl aliquot of peptide solution. The samples were incubated at either 4°C, 20°C, or 37°C for 30 min prior to measurements. Each data point represents an average of ten readings of the same mixture. For each condition (salt concentration and temperature), the entire experiments were repeated once.

Tryptophan Fluorescence Spectroscopy

Tryptophan fluorescence experiments were carried out using a fluorescence spectrophotometer (Hitachi, Model F-4500) at 20°C. The excitation wavelength was set to 295 nm for tryptophan to minimize tyrosine absorbance, and the emission spectra were recorded from 285 nm to 400 nm with a 0.2 nm step size and a 1 s integration time. SBP and a control 9-mer peptide (GRFNWLKEG) were dissolved in 100 mM TrisCl (pH 8.2), 150 mM NaCl to final concentrations of 30 μM and 60 μM, respectively. The apical domain was added as a stock solution (2.34 mM) to the peptide at both 1:1 and 5:1 molar ratios.

One- and Two-Dimensional Proton NMR

NMR experiments were carried out on Varian Unity 500 MHz spectrometer at 15°C. For one-dimensional (1D) NMR measurements, SBP was dissolved in 150 mM NaCl, 100 mM sodium phosphate (pH 6.5), and the solution was spun down to remove insoluble debris. The final SBP concentration was 0.38 mM. D_2O was added to a final concentration of 10% v/v prior to data collection to give a locked NMR signal. To improve the signal:noise ratio, 1D NMR spectra were also collected on SBP in D_2O. To transfer SBP into D_2O solution, the above sample was recovered, lyophilized, dissolved in 500 μl D_2O, incubated at 37°C for 1 hr, lyophilized again, and finally redissolved in 500 μl D_2O. Two-dimensional nuclear Overhauser enhancement spectroscopy (2D NOESY) experiments were also carried on this sample with a 500 ms mixing time. NOESY spectra were acquired with 2048 increments in the t2 domain and 128 data points in the t1 domain with spectral widths of 6000 Hz. Eighty scans for each t1 increment were accumulated with a 3.5 s delay between scans. Water suppression protocols (WATERGATE) were employed in 1D and 2D data collections (Piotto et al., 1992; Lippens et al., 1995). Data were processed and analyzed using Felix 95 (Molecular Simulations).

Crystallization

Complexes of the apical domain and SBP were prepared by mixing SBP (either SBP alone or fluorescein-tagged SBP) with 2.34 mM apical domain stock (in a molar ratio of apical domain:SBP = 1:1.5),

both in 10 mM NaCl, 20 mM TrisCl (pH 8.2). The solution was incubated at room temperature for 30 min and concentrated by Centricon-3 (Amicon) to a complex concentration of ~10 mg/ml, assuming all the apical domain was converted to the complex. Crystals of the complex were obtained by vapor diffusion from drops (4 μl) containing a 1:1 (v/v) ratio of complex solution to well solution (100 mM TrisCl [pH 8.0], 200 mM $MgCl_2$, and 30% PEG4000). The cryoprotectant solution was 120 mM TrisCl (pH 8.0), 240 mM $MgCl_2$, and 30% PEG4000, and 10% glycerol. Typical crystals were rectangular rods, with typical dimensions of 50 μm × 50 μm × 500 μm. However, they appeared in clusters. Diffraction data were collected on manually isolated single crystals; nevertheless, a weak second lattice was observed in some orientations of the crystal. Both indexing and scaling were carried out successfully, and the subsequent search for molecular replacement solution as well as refinements was not affected by the presence of this extra lattice.

Crystals of the apical domain alone were also obtained by vapor diffusion from drops containing a 1:1 (v/v) ratio of 10 mg/ml of the apical domain to well solution (100 mM TrisCl [pH 8.0], 200 mM Li_2SO_4, and 30% PEG4000). The cryoprotective solution was 120 mM TrisCl (pH 8.0), 240 mM Li_2SO_4, and 35% PEG4000. The crystals appeared as single hexagonal rods with dimensions of 100 μm × 100 μm × 800 μm.

Data Collection and Structure Determination

All data (complex of the apical domain with SBP, complex of the apical domain with fluorescein-tagged SBP, and apical domain alone) were collected at CHESS F-1 station using a Brandies CCD detector. Parameters of data collection and structural determination are listed in Table 2. A solvent content of ~40% is estimated for both complexes, and ~48% for the apo form. All the data were processed and merged using DENZO and SCALEPACK programs (Otwinowski and Minor, 1997).

Models were refined with CNS (Brünger et al., 1998). For the complex, two noncrystallographic two-fold rotation axes were shown by a self-rotation search to be nearly parallel to the **a** and **c*** axes, each deviating by ~4°. In the apo form, the two molecules in the asymmetric unit are related by a translation of 0.5 U along the **c** axis. Structures of the complex (SBP/apical domain) and the apo form were solved by molecular replacement, using the structure of an isolated apical domain (Zahn et al., 1996) as the search model. For the complex, strict noncrystallographic symmetry (ncs) was imposed in the initial rounds of refinements (25–2.5 Å). From 2.5 Å to 2.3 Å, restraint of ncs was gradually relaxed, and finally it was totally removed when refined at 2.1 Å. The structure was then used as the starting model for the refinement of the complex containing the fluorescein-tagged SBP. Maximum likelihood was used as the refinement target during positional refinement and simulated annealing. Model rebuilding, using O (Jones et al., 1991), was performed between rounds of CNS refinement. Five percent of the SBP/apical domain complex data and 10% of the apo apical domain data were removed at the outset for cross-validation. The refinement statistics are shown in Table 2.

Acknowledgments

We thank members of the Sigler lab for their help with data collection; in particular, we thank Otis Littlefield for critical review and valuable input. Marieke Bloemink provided excellent assistance in NMR experiments. Peptides were synthesized at the Howard Hughes Medical Institute facility/W. M. Keck Foundation Biotechnology Research Lab at Yale University. Data of SBP/apical domain complexes and the unliganded apical domain were collected at F-1 station at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation under award DMR 97–13424. Data of SBP/GroEL complex were collected at Argonne National Laboratory Structural Biology Center Beamline 19-ID at Advanced Photon Source (APS), which is supported by the US Department of Energy, Office of Energy Research, under Contract No. W-31–109-ENG-38. This work was supported by a National Institutes of Health grant (GM15225). L. C. is a Helen Hay Whitney Postdoctoral Fellow.

Received September 30, 1999; revised November 24, 1999.

References

Boisvert, D.C., Wang, J., Otwinowski, Z., Horwich, A.L., and Sigler, P.B. (1996). The 2.4 Å crystal structure of the bacterial chaperonin GroEL complexed with ATP gamma S. Nat. Struct. Biol. *3*, 170–177.

Braig, K., Simon, M., Furuya, F., Hainfeld, J.F., and Horwich, A.L. (1993). A polypeptide bound by the chaperonin GroEL is localized within a central cavity. Proc. Natl. Acad. Sci. USA *90*, 3978–3982.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., and Sigler, P.B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. Nature *371*, 578–586.

Braig, K., Adams, P.D., and Brünger, A.T. (1995). Conformational variability in the refined structure of the chaperonin GroEL at 2.8 Å resolution. Nat. Struct. Biol. *2*, 1083–1094.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Buckle, A.M., Zahn, R., and Fersht, A.R. (1997). A structural model for GroEL-polypeptide recognition. Proc. Natl. Acad. Sci. USA *94*, 3571–3575.

Chatellier, J., Buckel, A.M., and Fersht, A.R. (1999). GroEL recognises sequential and non-sequential linear structural motifs compatible with extended beta-strands and alpha-helices. J. Mol. Biol. *292*, 163–172.

Chen, S., Roseman, A.M., Hunter, A.S., Wood, S.P., Burston, S.G., Ranson, N.A., Clarke, A.R., and Saibil, H.R. (1994). Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. Nature *371*, 261–264.

Ellis, R.J. (1996). The Chaperonins (San Diego, CA: Academic Press).

Evans, S.V. (1993). SETOR: hardware lighted three-dimensional solid model representations of macromolecules. J. Mol. Graph. *11*, 134–138.

Fenton, W.A., and Horwich, A.L. (1997). GroEL-mediated protein folding. Protein Sci. *6*, 743–760.

Fenton, W.A., Kashi, Y., Furtak, K., and Horwich, A.L. (1994). Residues in chaperonin GroEL required for polypeptide binding and release. Nature *371*, 614–619.

Fremont, D.H., Matsumura, M., Stura, E.A., Peterson, P.A., and Wilson, I.A. (1992). Crystal structures of two viral peptides in complex with murine MHC class I H-2K. Science *257*, 919–926.

Hartl, F.U. (1996). Molecular chaperones in cellular protein folding. Nature *381*, 571–579.

Hayer-Hartl, M.K., Ewbank, J.J., Creighton, T.E., and Hartl, F.U. (1994). Conformational specificity of the chaperonin GroEL for the compact folding intermediates of alpha-lactalbumin. EMBO J. *13*, 3192–3202.

Hlodan, R., Tempst, P., and Hartl, F.U. (1995). Binding of defined regions of a polypeptide to GroEL and its implications for chaperonin-mediated protein folding. Nat. Struct. Biol. *2*, 587–595.

Horwich, A.L., Low, K.B., Fenton, W.A., Hirshfield, I.N., and Furtak, K. (1993). Folding in vivo of bacterial cytoplasmic protiens: role of GroEL. Cell *74*, 909–917.

Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L., and Deisenhofer, J. (1996). The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. Nature *379*, 37–45.

Itzhaki, L.S., Otzen, D.E., and Fersht, A.R. (1995). Nature and consequences of GroEL-protein interactions. Biochemistry *34*, 14581–14587.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kobayashi, N., Freund, S.M.V., Chatellier, J., Zahn, R., and Fersht, A.F. (1999). NMR analysis of the binding of a rhodanese peptide to a minichaperone in solution. J. Mol. Biol. *292*, 181–190.

Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Landry, S.J., Jordan, R., McMacken, R., and Gierasch, L.M. (1992). Different conformations for the same polypeptide bound to chaperones DnaK and GroEL. Nature *355*, 455–457.

Lin, Z., Schwartz, F.P., and Eisenstein, E. (1995). The hydrophobic nature of GroEL-substrate binding. J. Biol. Chem. *270*, 1011–1014.

Lippens, G., Dhalluin, C., and Wieruszeski, J.M. (1995). Use of a water flip-back pulse in the homonuclear NOESY experiment. J. Bio. NMR *5*, 327–331.

Ma, J., and Karplus, M. (1998). The allosteric mechanism of the chaperonin GroEL: a dynamic analysis. Proc. Natl. Acad. Sci. USA *95*, 8502–8507.

Madden, D.R., Garboczi, D.N., and Wiley, D.C. (1993). The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptids presented by HLA-A2. Cell *75*, 693–708.

Mande, S.C., Mehra, V., Bloom, B.R., and Hol, W.G.J. (1996). Structure of the heat shock protein chaperonin-10 of mycobacterium leprae. Science *271*, 203–207.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L., and Hartl, F.U. (1991). Chaperonin-mediated protein folding at the surface of groEL through a "molten globule"-like intermediate. Nature *352*, 36–42.

Mayhew, M., da Silva, A.C.R., Martin, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. (1996). Protein folding in the central cavity of the GroEL-GroES chaperonin complex. Nature *379*, 420–426.

Merritt, E.A., and Bacon, D.J. (1997). RASTER3D: photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

Nicholls, A., Sharp, K., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–325.

Piotto, M., Saudek, V., and Sklenar, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J. Bio. NMR *2*, 661–665.

Schmidt, M., and Buchner, J. (1992). Interaction of GroE with an all-beta-protein. J. Biol. Chem. *267*, 16829–16833.

Scott, J.K., and Smith, G.P. (1990). Searching for peptide ligands with an epitope library. Science *249*, 386–390.

Sharp, K.A., Nicholls, A., Fine, R.F., and Honig, B. (1991). Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. Science *252*, 106–112.

Shtilerman, M., Lorimer, G.H., and Englander, S.W. (1999). Chaperonin function: folding by forced unfolding. Science *284*, 822–825.

Sigler, P.B., Xu, Z., Rye, H.S., Burston, S.G., Fenton, W.A., and Horwich, A.L. (1998). Structure and function in GroEL-mediated protein folding. Annu. Rev. Biochem. *67*, 581–608.

Tanaka, N., and Fersht, A.R. (1999). Identification of substrate binding site of GroEL minichaperone in solution. J. Mol. Biol. *292*, 173–180.

Thiyagarajan, P., Henderson, S.J., and Joachimiak, A. (1996). Solution structures of GroEL and its complex with rhodanese from small-angle neutron scattering. Structure *4*, 79–88.

Viitanen, P.V., Gatenby, A.A., and Lorimer, G.H. (1992). Purified chaperonin 60 (groEL) interacts with the nonnative states of a multitude of Escherichia coli proteins. Protein Sci. *1*, 363–369.

Weissman, J.S., Hohl, C.M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H.R., Fenton, W.A., and Horwich, A.L. (1995). Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. Cell *83*, 577–587.

Weissman, J.S., Rye, H.S., Fenton, W.A., Beechem, J.M., and Horwich, A.L. (1996). Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. Cell *84*, 481–490.

Xu, Z., and Sigler, P.B. (1998). GroEL/GroES: structure and function of a two-stroke folding machine. J. Struct. Biol. *124*, 129–141.

Xu, Z., Horwich, A.L., and Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature *388*, 741–750.

Zahn, R., Buckle, A.M., Perrett, S., Johnson, C.M., Corrales, F.J., Golbik, R., and Fersht, A.R. (1996). Chaperone activity and structure of monomeric polypeptide binding domains of GroEL. Proc. Natl. Acad. Sci. USA *93*, 15024–15029.

Protein Data Bank ID Codes

The ID codes for the structures described in this paper are 1DKD and 1DK7 for the SBP/apical domain complex and apo form, respectively.