Crystal Structure of the Heterodimeric Complex of the Adaptor, ClpS, with the N-domain of the AAA⁺ Chaperone, ClpA*

Received for publication, August 8, 2002, and in revised form, September 11, 2002 Published, JBC Papers in Press, September 15, 2002, DOI 10.1074/jbc.M208104200

Fusheng Guo, Lothar Esser, Satyendra K. Singh, Michael R. Maurizi, and Di Xia‡

From the Laboratory of Cell Biology, Center for Cancer Research, NCI, National Institutes of Health, Bethesda, Maryland 20892

Substrate selectivity and proteolytic activity for the E. coli ATP-dependent protease, ClpAP, is modulated by an adaptor protein, ClpS. ClpS binds to ClpA, the regulatory component of the ClpAP complex. We report the crystal structure of ClpS in complex with the isolated N-terminal domain of ClpA in two different crystal forms at 2.3- and 3.3-Å resolution. The ClpS structure forms an αβ-sandwich and is topologically analogous to the C-terminal domain of the ribosomal protein L7/L12. ClpS contacts two surfaces on the N-terminal domain in both crystal forms; the more extensive interface was shown to be favored in solution by protease protection experiments. The N-terminal 20 residues of ClpS are not visible in the crystal structures; the removal of the first 17 residues produces ClpSΔN, which binds to the ClpA N-domain but no longer inhibits ClpA activity. A zinc binding site involving two His and one Glu residue was identified crystallographically in the N-terminal domain of ClpA. In a model of ClpS bound to hexameric ClpA, ClpS is oriented with its N terminus directed toward the distal surface of ClpA, suggesting that the N-terminal region of ClpS may affect productive substrate interactions at the apical surface or substrate entry into the ClpA translocation channel.

ClpA, a member of the Hsp100/Clp family of molecular chaperones, promotes the ATP-dependent degradation of proteins (1, 2) by forming a complex with ClpP, a proteasome-like hollow tetradecamer (3) that degrades substrates into peptides of 5–15 amino acids (4). The crystal structure of ClpA reveals an N-terminal domain (N-domain) and two tandem, nonidentical AAA⁺ modules (D1 and D2) (5). The AAA⁺ superfamily (ATPases associated with various cellular activities) comprises energy-dependent molecular machines that participate in assembly and disassembly of a wide variety of macromolecular complexes (6, 7). AAA⁺ proteins are composed of modular functional domains (8): one or two AAA⁺ modules that use binding and hydrolysis of ATP to couple conformational changes in the AAA⁺ protein to structural changes in associated macromolecular substrates and additional N- or C-terminal domains that appear to interact with substrates and to transduce the motion of the AAA⁺ protein into movements of or conformational changes in the substrate.

AAA⁺ proteins also interact with adaptor proteins that mediate various functions. Escherichia coli ClpX needs Rsb/SprE for specific degradation of RpoS (9, 10) and uses SspB to enhance degradation of SsrA-tagged proteins (11). Bacillus subtilis MecA is needed for efficient degradation by ClpCP (12). In eukaryotes, the AAA⁺ chaperone, p97, uses several adaptor proteins: p47, to promote specific membrane fusion (13); Ufd1p, for ubiquitin-dependent proteolysis (14); and Npl4, for nuclear transport (14). Such adaptors allow the activity of the p97 to be directed toward different cellular targets. Recently, a small adaptor-like protein, ClpS, has been shown to modulate activity of E. coli ClpA (15). ClpS is a 106-residue protein encoded upstream of ClpA in most bacteria (16). ClpS protects ClpA from autodegradation and appears to redirect its activity away from soluble proteins and toward aggregated proteins (15). ClpS is thus an important regulatory factor for the activity of ClpA and ClpAP.

ClpA recognizes sequences near the N or C terminus of specific substrates (17, 18), unfolds, and translocates them into the degradation chamber of an associated ClpP (18–20). Substrates apparently bind to the apical surface of ClpA and follow an axial translocation pathway into the active site of ClpP (20). Additionally, in ClpA, there appears to be a substrate binding site within the small subdomain of the D2 AAA⁺ module (21), but the role of this site is not well understood. ClpA lacks the N-domain can interact with specific substrates and promote protein degradation nearly as efficiently as intact ClpA (22). Interestingly, ClpS cannot inhibit activity of ClpA lacking its N-domain, suggesting that ClpS interacts with the N-domain of ClpA (15). Together, ClpS and the N-domain of ClpA play a modulating role in substrate selection and protein unfolding and degradation activities of ClpAP.

The N-domain of ClpA connects to the D1 domain by a long flexible loop (5) and is expected to be mobile based on electron microscopy and biochemical studies (20, 23). We have found that the isolated N-domain forms a tight complex with ClpS, suggesting that ClpS might affect ClpA activity by altering N-domain orientation or interfering N-domain interaction with substrates. We further report the crystal structure of the heterodimeric complex of ClpS with the N-domain and present a model of ClpS in complex with intact hexameric ClpA. The model suggests that the N-domain helps orient ClpS, allowing it to block substrate binding to the surface of ClpA.
Expression, Purification, and Crystallization of the ClpS-N Complex—The ClpS-N complex coding region was amplified from pWPC3.1 (16) using the forward primer, TGATAACTGCTATGCTGTAACAAAAAGCCGACTG-GC, and the back primer, TTATGCGTATGTCGTTTTTACTGTAC-CAC. The product was cut with NdeI and PstI and inserted into a derivative of the vector pBAD33. DH5α cells containing the plasmid, pBAD33-clpS, were grown in Luria broth containing chloramphenicol (34), transformed with pBAD33-clpS with 0.2% arabinose at an A600 of 0.7, and collected by centrifugation after 3 h. After suspension in Bis-Tris buffer (20 mM, pH 6.5) containing 10% (v/v) glycerol, cells were broken at 20,000 p.s.i. in a French pressure cell. Cell debris was removed by centrifugation at 10,000 × g for 30 min. The supernatant extract was treated with 0.05% polyethyleneimine, and precipitated material was removed by centrifugation. Ammonium sulfate (40% saturation) was added to the supernatant solution, and the precipitated protein was collected by centrifugation and dissolved in buffer containing 50 mM, Tris-HCl, pH 8.5, and 10% glycerol. After dialysis against the same buffer, the protein was further purified by chromatography on a MonoQ (Amersham Biosciences) and a hydroxyapatite (Bio-Rad) column and by gel filtration on a Superdex 75 column. Purified ClpS was concentrated to 12 mg/ml and dialyzed against 20 mM Hepes, pH 7.5. Aliquots were stored at −80 °C. The production of the N-domain of ClpA followed the procedure previously described (5). ClpS and the N-domain of ClpA were mixed in an equal molar ratio and run over a Superdex 75 gel filtration column. A single peak corresponding to the ClpS-N complex was collected and concentrated to 10 mg/ml for crystallization.

Crystallization experiments with the ClpS-N complex resulted in several crystal forms. Crystals in the trigonal space group P3_21 were obtained by vapor diffusion after 1 week at 21 °C in hanging drops of 2 μl of protein solution mixed with 2 μl of reservoir solution consisting of 0.1 M Bis-Tris, pH 6.5, 32% glycerol, and 10–15 mM yttrium chloride. Crystals in the space group P4_3212 were grown at 4 °C from hanging drops of 2 μl of protein solution and 2 μl of well solution consisting of 0.1 M Bis-Tris, pH 6.5, 32% glycerol, and 2% polyethylene glycol 4000. Both forms of crystals can be frozen successfully without further manipulation and diffracted X-rays to 2.3 Å resolution for the P3_21 and 3.3 Å resolution for the P4_3212 space group. Heavy-atom derivatives for the trigonal crystal form were prepared by soaking the crystals in the crystallization solution containing 1–10 mM heavy atom compound for 5–10 min.

Crystallographic Data Collection and Structure Determination—The native x-ray diffraction data set used in structure determination for the P3_21 space group was collected at the beamlines BioCARS-CAT, Advanced Photon Source at the Argonne National Laboratory, and X9B of the National Synchrotron Light Source at Brookhaven National Laboratory, using Area Detector System Co. quantum-4 CCD detectors. A data set at the wavelength corresponding to the peak of the EXAFS spectrum (λ = 1.0057 Å) for the mercury derivative was collected to 2.8 Å resolution at the beamline BioCARS of the Advanced Photon Source, the program Phaser was performed on the CCP4 suite (28), and phases were calculated in SIR/AS runs of MLPhare (27), and the electron density was determined with SnB (25) and confirmed with ShelXD (26). Phases were calculated in SIRAS runs of MLPhare (27), and the electron density was improved by solvent flattening and 2-fold molecular averaging in DM (28). Starting with the phases computed by DM, the program REFMAC (29) was used to solve the phases. Data refinement, difference Fourier maps, and model building were done in O (30). The models of two heterodimers were refined to 2.3 Å resolution using CNS (31). Throughout the refinement, appropriate noncrystallographic symmetry (NCS) restraints were maintained. Difference Fourier maps revealed the positions of four yttrium Bis-Tris moieties attached to acidic residues and two chloride ions. The tetragonal crystal form was solved by molecular replacement using a ClpS-N heterodimeric model built using the A interface (for details, see “Results and Discussion”) in rotational and translational searches with MolRep (32). The final solution was obtained and subsequently refined. Details of the structure determination and refinement statistics are listed in Table I.

RESULTS AND DISCUSSION

Complexes of ClpS with ClpA Hexamers and ClpA N-domain—Recombinant ClpS (106 amino acid residues) was purified from an overexpressing E. coli strain (see “Experimental Procedures”) (Fig. 1A, lane 3). The protein produced a single band upon gel electrophoresis under nondenaturing conditions (data not shown) and had an approximate molecular mass of 12 kDa by gel filtration (Fig. 1B). ClpS is thus a monodisperse monomer in solution. In agreement with an earlier report (15), ClpS formed a complex with intact ClpA; upon Superdex 200 gel filtration in the presence of ATP-γ-S, both proteins were found in the 15-min fraction, corresponding to a molecular mass of 500 kDa (Fig. 1A, lane 2). In contrast, when a mutant ClpA lacking the N-domain, ClpAΔ153 (22), was used, much less ClpS was recovered in the high molecular weight fraction (data not shown), suggesting that the N-domain of ClpA is needed for tight binding to ClpS.

To assay ClpS binding to the N-domain, we purified the 143-amino acid residue N-terminal fragment of ClpA, which behaves as a monomer in solution (Fig. 1A, lane 4) and B). Isolated ClpA N-domain had previously been shown to block the inhibitory activity of ClpS on intact ClpA, suggesting that the N-domain of ClpS binds to ClpS (15). When N-domain was mixed with ClpS and run on a Superdex 75 gel filtration column, the two proteins emerged together in a single peak with a molecular mass of about 28 kDa (Fig. 1A, lane 5 and B), providing direct evidence of a stoichiometric heterodimeric complex of ClpS and N-domain and suggesting that the major site of interaction with ClpS is through the N-domain of ClpA. We then generated a variant of ClpS (ClpSAN) by removing the N-terminal 17 amino acid residues with lysylendopeptidase C (Fig. 1A, lane 6). When ClpSAN was mixed with ClpS, both proteins eluted from the gel filtration column in the same fraction (Fig. 1A, lane 7). Thus, ClpS lacking its N-terminal 17 amino acid residues retains the ability to bind tightly to ClpA.

Structure Solution of the Heterodimer of ClpS and ClpA N-domain (ClpS-N)—The complex of ClpS with the N-domain of ClpA, which we will call ClpS-N, was crystallized, and two crystal forms were obtained under slightly different growth conditions (see “Experimental Procedures”). The trigonal form in space group P3_21 diffracted X-rays to 2.3 Å resolution and had cell dimensions of a = b = 88.0 Å, and c = 210.2 Å. The tetragonal crystals in space group P4_3212 diffracted to 3.3 Å resolution and possessed cell dimensions of a = b = 91.2 Å, and c = 198.6 Å. Statistics for diffraction data sets, phase determination, and final atomic models are given in Table I. Both crystal forms are unusual in having very high solvent content; the trigonal crystal form has two molecules of ClpS-N heterodimer in the asymmetric unit with the Matthews’ coefficient of 4.3 (33), corresponding to a solvent volume of 71%; the tetragonal form contains just one molecule of the heterodimer in the asymmetric unit, having a surprisingly high Matthews’
The atomic model refined to 2.3 Å resolution in P3121 includes sets of ClpS-N helical chains in six layers related by NCS symmetry. In the trigonal crystal, six of them, the charges were neglected (Fig. 2C). The L7/L12 ribosomal protein has two domains: an N-terminal dimerization domain and a C-terminal phobic core. The conserved hydrophilic residues are shown as small spheres (Fig. 2A). In the P3121 crystal, each N-domain is in contact with three molecules of ClpS using three different interfaces, whereas in P422, only two ClpS molecules interact with one N-domain, using two of the binding interfaces observed in the P3121 space group. By using the two contacts observed in both crystal forms (interfaces A and C; see below), ClpS and N-domain heterodimers produce continuous helical chains in crystals (Fig. 2A). In the trigonal crystal, the two N-domains in the asymmetric unit are related by an NCS rotation of 167°, and the two ClpS molecules are related by an NCS rotation of 172°. The atomic model refined to 2.3 Å resolution in P3121 includes 456 residues, four yttrium ions, two chloride ions, and 128 water molecules. The yttrium ions, required for crystallization, bind to the third interface between the N-domain and ClpS (interface B), which is not involved in helical chain extension of the ClpS-N heterodimers. The two chloride ions were assigned based on their positively charged environments, anomalous signals, and refined B factors.

Despite the large solvent content, the ClpS-N crystals are remarkably stable and well ordered. In the trigonal crystal, six sets of ClpS-N helical chains in six layers 60° apart in orientation and perpendicular to the c axis (three sets of two chains related by the NCS symmetry) form an interconnected network. In the tetragonal crystal, four sets of ClpS-N helical chains in four layers 90° apart in orientation stack together along the 4₃ axis. These interactions provide sufficient strength and support for the crystals to accommodate their extraordinarily large solvent content.

**Structures of ClpS and the N-domain of ClpA—**In the trigonal crystal, 87 out of 166 residues in ClpS can be seen in the two NCS-related heterodimers; the missing residues are from the N terminus. ClpS has a single α/β sandwich folding motif consisting of three α helices (H1, H2, and H3) on one side of a three-stranded anti-parallel β-sheet (S1, S2, and S3) (Fig. 2, A and B). The β-stand S2 contains a conserved β-bulge of type C, subdividing it into strands S2a and S2b (Fig. 2, B and C) (34). The molecule is cone-shaped, with the helices H1 and H2 lying flat at the base of the cone and the twisted β-sheet and helix H3 converging at the tip of the cone. The N-terminal residues from 21 to 26 extend out of the top of the cone, but residues 1–20 were not all visible. Secondary structure assignment of ClpS to its sequence is provided in Fig. 2B. Sequence alignment of selected ClpS homologs from prokaryotic and eukaryotic sources (Fig. 2B) shows a highly conserved N-terminal region (S1 and H1), with most conserved residues forming the hydrophobic core. The conserved hydrophilic residues are shown as the ball-and-stick models (Fig. 2A) and are clustered at two locations that form close contacts with the N-domain (see below).

A search of the protein structure data base (35) identified only the ribosomal protein, L7/L12, with significant structural similarity to ClpS. The C-terminal domain of L7/L12 (36, 37) overlaps with ClpS with an r.m.s. deviation of 1.7 Å for 63 pairs of C-α atoms out of 68 residues in the structure (Fig. 2C). Structure-based sequence alignment of L7 to ClpS (Fig. 2B) shows a 13% sequence identity or 22% similarity between the two sequences, with most of the conservation concentrated at the N terminus of the sequence in S1 and H1. The β-bulge in ClpS between strands S2a and S2b is also present in the structure of L7/L12 (Fig. 2C). There are nine pairs of charged residues in the aligned structures; in six of them, the charges are inverted (Fig. 2C). The L7/L12 ribosomal protein has two domains: an N-terminal dimerization domain and a C-terminal

---

**Table 1**

| Resolution (Å) | No. of unique reflections | Wavelength (Å) | Rmerge* (outer bin) | Completeness, % (outer bin) | Sites | Ph. P⁺ (acentric) | Rwork (acentric) | Ph. P⁺ (centric) | Rwork (centric) |
|----------------|---------------------------|---------------|-------------------|---------------------------|-------|----------------|---------------|----------------|---------------|
| P3,21 Native   | 50–2.3                    | 42,630        | 1.0093            | 0.055                     | (0.579) | (100.0)       | 99.9          | 7.134          | 0.840         |
| Mercury        | 50–2.8                    | 24,574        | 1.0057            | 0.041                     | (0.594) | (97.3)        | 99.6          | 7.134          | 0.840         |
| Zinc           | 50–2.25                   | 45,812        | 1.569             | 0.045                     | (0.638) | (69.9)        | 95.9          | 7.134          | 0.840         |
| P4,2,2 Native  | 28–3.3                    | 13,160        | 1.0072            | 0.061                     | (0.594) | (100.0)       | 99.5          | 7.134          | 0.840         |

P4,2,2 

Refinement statistics

| P3,21 | P4,2,2 | P3,21 |
|-------|--------|-------|
| No. of residues | 456 | 228 | 456 |
| No. of nonhydrogen atoms | 3,851 | 1,813 | 3,864 |
| No. of metal ions | 4 | 0 | 4 |
| No. of Cl⁻ | 2 | 0 | 2 |
| No. of Zn²⁺ | 0 | 0 | 2 |
| No. of water molecules | 128 | 0 | 150 |
| Rfree (outer bin) | 0.251 (0.321) | 0.280 (0.442) | 0.236 (0.354) |
| Rwork (outer bin) | 0.217 (0.308) | 0.258 (0.571) | 0.214 (0.351) |
| Resolution (Å) | 20–2.25 | 12–3.3 | 20–2.3 |
| r.m.s. deviation of bond length (Å) | 0.0157 | 0.0184 | 0.0116 |
| r.m.s. deviation of bond angle (degrees) | 1.64 | 1.37 | 1.46 |

* Rmerge = Σ[Ih,i]−<Ih> / Σ[Ih]−<Ih>, where Ih,i is the intensity for i-th observation of a reflection with Miller index h, and <Ih> is the mean intensity for all measured Ih values and Freidel pairs.

* Ph. P (phasing power) = <fph> / residual, where <fph> is the r.m.s. of structure factor contribution in amplitude of heavy atoms, and residual is the r.m.s. of lack of closure error.
The structure of the N-domain of ClpA complexed with ClpS. A, a ribbon representation of the ClpS-N complex. One N-domain and two associated ClpS are shown. Two different interfaces are observed between ClpS and the N-domain and are referred to as interfaces A and C. The N-domain is a pseudodimer, shown with the N-terminal half in yellow and the C-terminal half in blue. The location of the Zn$^{2+}$ binding site is as labeled. In the ClpS $\alpha/\beta$ sandwich, the $\beta$-sheet is in red, and the $\alpha$-helices are in green. Secondary structure elements as well as conserved
domain, connected by a flexible loop. The influence of the C-terminal domain of L7/L12 on the elongation factor-dependent GTP hydrolysis (38, 39) suggests that it might physically interact with these proteins. Indeed, in cryoelectron microscopic images, the stalk region (i.e. presumably L7/L12) can be seen in contact with a ternary EF-TU complex (40). The significant similarity in structure of ClpS to the L7/L12 may imply analogous functions of ClpS in associating with and affecting ATPase activity of ClpA. In this context, we have found that whereas ClpS has minimal effect on the ATPase activity of native ClpA, it leads to a 2-fold increase in ATPase activity of a truncated mutant consisting of the N-terminal and D1 AAA+ domains. The structure of the ClpA N-domain consists of eight α-helices (Fig. 2A). The first four-helix bundle (N-H1 to N-H4) is related to the second four-helix bundle (N-H1’ to N-H4’) by a pseudo-2-fold axis and is connected by a long acidic loop (positions 66–79) that is disordered in the full-length ClpA structure (5) and partially stabilized upon binding of ClpS. Sequence conservation between the two halves of the pseudodimer is poor except in the hydrophobic core (41) and in a few conserved hydrophilic residues distributed at both ends of helix bundles (Fig. 2B). The N-domain is apparently not related to any known protein structure in the protein structure data base (35). The details of the N-domain structure, in isolation and as part of the full-length ClpA subunit, and of the tandem repeat in the N terminus have been reported recently (5). There is a piece of isolated electron density identified on the surface hydrophobic depression between the N-domain pseudorepeating units. A tentative peptide model of the N terminus of ClpS was built (data not shown), suggesting a potential peptide binding site for the N-domain.

Heterodimeric Association of ClpS and ClpA N-domain—In the trinodal crystal, an N-domain has three ClpS neighbors, related by either noncrystallographic or crystallographic symmetry. We designate the three contacting interfaces as A, B, and C. Buried surface areas of 142 aligned C-atoms. The four helices in the C-terminal half of the N-domain form a cleft into which S-Asp36 and S-Tyr37, both conserved within prokaryotes, penetrate to interact with N-Asn107 and N-Arg100, respectively (Table II). The N terminus of ClpS in this configuration extends in the opposite direction of the N-domain (Fig. 2A).

To decide which interface is more likely to be the one dominated under physiological conditions, we generated models of the three crystallographically observed S-N-S trimers as in Fig. 2A; two of the models are NCS-related and come from the trinodal crystal form, and one comes from the tetragonal crystal form. The three trimers are superimposed on their N-domains as shown in Fig. 3C. The three aligned N-domains give rise to an r.m.s. deviation of 0.44 Å for 142 aligned C-atoms. The three ClpS molecules that bind to the N-domain via interface A have a slightly larger r.m.s. deviation of 0.56 Å for 86 C-atoms, whereas those binding to the N-domain through the interface C show an r.m.s. deviation of 1.65 Å. The data demonstrate that the ClpS binding surface A provides a more constant and apparently more stable contact than C in the different crystal environments. Therefore, surface A appears to be more favored as the dimer interface under physiological conditions.

To identify the contact that predominates in solution, we probed exposed surfaces of the proteins by limited proteolysis. The ClpS, the N-domain of ClpA, or a mixture of the two in a 1:1 molar ratio was incubated with chymotrypsin for various times, cleavage products were separated on a SDS gel, and N-terminal sequences of the products were determined (Fig. 4). When incubated alone, ClpA-N was cut at Phe14, Tyr122, and Phe137, but when ClpS was present, the Tyr122 site was protected from digestion, suggesting that ClpS blocks access to Tyr122. In the N-domain structure, Tyr122 is exposed on the surface, but in the complex Tyr122 is shielded by residues 22–24 of ClpS making the A contact. Interestingly, Phe14 was also protected by ClpS. This residue was part of the yttrium-stabilized interface B discussed above, suggesting that the B interface may form transiently in solution. Phe137 was not protected and is thus not involved in contact with ClpS. Analysis of the

hydrophilic residues are labeled accordingly. The diagram is produced with Molscript (48), Bobscript (49), and Povray (available on the World Wide Web at www.povray.com) interfaced with GL-render (L. Esser, unpublished work). B, structure-based sequence alignment of ClpS homologs. Selected sequences of ClpS homologs from prokaryotes and eukaryotes are shown. The alignment and secondary structure elements assignment to the sequence are based on the structures of ClpS determined here and E. coli ribosomal protein L7/L12 (Protein Data Bank entry 1CTF (36)). The β-strands, represented as green arrows, are labeled S1, S2, and S3 sequentially as they appear in the sequence; strand S2 is subdivided into S2a and S2b. The α-helices, shown as wiggling curves, are labeled H1, H2, and H3. Coils are shown as straight line segments. C, superposition of the structures of ClpS and E. coli ribosomal protein L7/L12. ClpS is shown in red, and L7/L12 is shown in green. Residues in which there is a charge inversion are shown as stick models and are labeled.
Fig. 3. Interfaces between the ClpS and the N-domain. The secondary structure elements that contribute residues to subunit interactions are labeled. Residues that form hydrogen bonds and salt bridge pairs are shown in stick models. Dashed lines are drawn between atom pairs less than 3 Å apart. Water molecules are shown as gray balls. A, stereo pair showing hydrogen bonds and salt bridges between residues from ClpS (green) and the N-domain (yellow) at interface A. B, stereo pair showing hydrogen bonds and salt bridges between residues from the ClpS (dark blue) and the N-domain (yellow) at interface C. C, ClpS bound at interface C is conformationally more flexible than that bound at interface A. Three independent N-domains (two NCS-related in P3212 and one in P43212) with attached ClpS molecules at both interfaces A and C were superimposed. The ClpS-N-domain/ClpS molecules in red/purple/red are from the trigonal crystal; the green/light green/green molecules are NCS-related to the red/purple/red in the trigonal crystal, and the blue/cyan/blue molecules are from the tetragonal crystal. The A and C interfaces are labeled.
Structure of the Heterodimer of ClpS and ClpA N-domain

TABLE II
Hydrogen bonds and salt bridges between the ClpS and the N-domain of ClpA

| Interface A | N-domain residue | Atom   | Distance Å | ClpS residue | Atom   |
|-------------|------------------|--------|------------|--------------|--------|
| 1           | Glu<sup>23</sup> | OE1    | 2.45       | Lys<sup>98</sup> | NE     |
| 2           | Glu<sup>28a</sup> | OE1    | 2.66       | Glu<sup>92</sup> | OE2    |
| 3           | Glu<sup>117</sup> | OE1    | 3.25       | Tyr<sup>28a</sup> | OH     |
| 4           | Arg<sup>66</sup> | NH2    | 2.83       | Glu<sup>92</sup> | OE1    |
| 5           | Arg<sup>66</sup> | NH1    | 2.98       | Glu<sup>92</sup> | OE2    |
| 6           | Thr<sup>81a</sup> | OG     | 2.63       | Glu<sup>92</sup> | OE1    |
| 7           | Leu<sup>82</sup> | N      | 3.20       | Thr<sup>93</sup> | OG     |
| 8           | Arg<sup>76</sup> | NH2    | 3.57       | Lys<sup>83</sup> | NZ     |
| 9<sup>a</sup> | Ala<sup>55</sup> | O      | 2.67 and 3.00 | Thr<sup>77</sup> | N      |
| 10<sup>a</sup> | Ser<sup>95</sup> | N      | 3.07 and 2.74 | Glu<sup>92</sup> | O      |
| 11<sup>a</sup> | Gln<sup>119</sup> | N      | 2.92 and 2.77 | Glu<sup>92</sup> | O      |

**Interface C**

| 1 | His<sup>140</sup> | NE | 2.83 | Pro<sup>23</sup> | O |
| 2 | Asn<sup>107a</sup> | ND | 3.63 | Tyr<sup>27</sup> | OH |
| 3 | Arg<sup>109b</sup> | NH2 | 3.28 | Asp<sup>36</sup> | OD1 |
| 4 | Arg<sup>110b</sup> | NH1 | 3.58 | Asp<sup>36</sup> | OD2 |
| 5 | Ser<sup>95</sup> | OG | 2.60 | Asp<sup>36</sup> | O |
| 6 | Arg<sup>131</sup> | NH2 | 3.46 | Glu<sup>92</sup> | O |

<sup>a</sup> Conserved residues.  
<sup>b</sup> Water-mediated hydrogen-bonds.

---

Fig. 4. Sites in ClpA-N and ClpS protected from protease digestion in the complex. ClpS alone, ClpA-N alone, or an equimolar mixture of the two was incubated at 37 °C for 0, 10, 30, 60, and 90 min with 5% (by weight) of chymotrypsin. Reactions were quenched by addition to hot SDS sample buffer. Proteins were separated by SDS-PAGE and stained with Coomassie Blue. A parallel gel was run, and the proteins were blotted to polyvinylidene difluoride membranes and N-terminally sequenced. S/29, proteolytic products of ClpS fragments 1–22 and 1–28, respectively. The numbers refer to the position in the amino acid sequence corresponding to the N terminus of the indicated products. S/I, proteolytic products of ClpS fragment 29–106; S/23, proteolytic products of ClpS fragment 25–106.

ClpS cleavage pattern gave reciprocal results. ClpS was cut at Tyr<sup>28</sup> in the absence of the N-domain and was protected in the complex (Fig. 4). Tyr<sup>28</sup> is involved in contact A with the N-domain. These data further support the model that the A interface is used for the ClpS-N dimer in solution and probably under physiological conditions. We cannot rule out that the B and C interfaces could also represent modes of interaction between ClpS and ClpA under some circumstances. Alternatively, the surfaces of ClpS or N-domain involved in these interfaces might have other binding functions in solution. It is not uncommon for substrate-binding sites and protein-protein interfaces to provide sites for crystal lattice contacts (43). In fact, in our structure of the full-length ClpA (5), the N-D1 interface employs the same region of the N-domain that makes interface B in the trigonal crystal form of ClpS-N. Also, the limited but rather precise fit producing contact C in the ClpS-N crystal may mimic another substrate or adaptor binding site that specifically recognizes an exposed Asp-Tyr sequence motif.

The N-domain of ClpA Contains a Zn<sup>2+</sup> Binding Site—A weak but consistent zinc signal measured by x-ray fluorescent spectra of ClpS-N crystal (data not shown) prompted an investigation of potential zinc binding property of the ClpS-N complex. A crystal of ClpS-N grown in the presence of 2 mM of ZnCl<sub>2</sub> diffracted X-rays to 2.25 Å resolution using a home x-ray source and was isomorphic to the trigonal crystal form. The structure was refined to an R<sub>free</sub> of 0.236 and R<sub>work</sub> of 0.214, respectively (Table I). The Fourier synthesis using the coefficients of |F<sub>i</sub>| – |F<sub>c</sub>| located a 14σ peak in the N-domain, and the Fourier synthesis using the coefficients of |F<sub>i</sub>| – |F<sub>c</sub>| produced a 7.5σ anomalous peak overlapping the same site. A zinc ion was introduced into the ClpS-N model and was placed in the N-terminal half of the N-domain in a depression between N-H1 and N-H4 near the acidic loop connecting the two halves of the N-domain (Fig. 2A). The Zn<sup>2+</sup> ion is tetrahedrally coordinated by three residues and a water molecule; His<sup>20</sup>, His<sup>22</sup>, and Glu<sup>63</sup> are 2.26, 2.15, and 2.14 Å, respectively, from the Zn<sup>2+</sup>, and the water molecule is 2.14 Å away (Fig. 5). When superimposed, the N-domains with and without Zn<sup>2+</sup> produced an r.m.s. deviation of 0.242 Å for 142 Cα pairs. Superposition of the N-terminal halves of the N-domain with and without Zn<sup>2+</sup> yielded an even smaller r.m.s. deviation of 0.146 Å for 48 Cα atom pairs. Large conformational changes are observed for the Zn<sup>2+</sup> ligands; the imidazole ring of His<sup>22</sup> is rotated nearly 120°, and the side chain of Glu<sup>63</sup> is moved more than 3 Å (Fig. 5).

The Zn<sup>2+</sup> coordination and binding residues in the N-domain is highly reminiscent of the Zn<sup>2+</sup> binding sites observed in carboxypeptidase A (44), thermolysin (45), mitochondrial processing peptidase (46), and other zinc-metalloproteases, although preliminary assays of peptidase activities in the presence of added Zn<sup>2+</sup> showed no evidence for activity, and there is lack of secondary structure resemblance. Although it is unlikely for ClpA to have an uncontrolled peptidase activity, we cannot rule out the possibility that specific cleavage may occur when the N-domain is attached to the D1 domain. Indeed, in the full-length ClpA, the Zn<sup>2+</sup> binding site is facing the D1 small domain, forming a cleft that is highly positively charged.
Conformational Variation of the N-domain in the Presence or Absence of ClpS—The structures of the N-domain of ClpA determined as an isolated domain and as part of the full-length ClpA subunit were virtually identical (5). Superposition of N-domains with and without bound ClpS gives an r.m.s. deviation of 1.01 Å for 139 of 142 C-α atom pairs, indicative of an overall structural rigidity of the N-domain. There are two noticeable local conformational changes to the N-domain upon ClpS binding (Fig. 6). At the A contact, the helix N-H2, the small turn between H2’ and H3’, and the acidic loop connecting the two halves of the N-domain undergo large displacement. Residues whose side chains display large motions are shown in Fig. 6. At the C contact, there is a large conformational change involving helices N-H1’ and N-H4’ that form the cleft where the S-Tyr37 inserts; both helices are pushed apart to accommodate the intrusion of S-Tyr37. Residues Ser36’ and Gly37’ at the tip of the cleft between N-H3 and N-H4 are displaced by as much as 3 Å.

Models of Hexameric ClpAS Complex—In our recently reported crystal structure of the full-length ClpA subunit (5), the N-domain is attached to the D1 AAA’ module, making major contacts with the D1 small subdomain. In the hexamer model, the N-domains are located on the periphery of the D1 hexameric ring with the acidic loop between the halves of the pseudodimer projecting toward D2. To model the ClpS bound to the ClpA hexamer, we superimposed the N-domain of ClpS-N with ClpS attached at two possible interfaces (the A or the C interface) onto the N-domain of ClpA (Fig. 7). We refer to the ClpAS complex using interface A as the A model and that using interface C as the C model. In the A model, the ClpS wedges comfortably into an empty space between the D1 and D2 domain, making contacts with the transition helix connecting D1 and D2 and with the D2 small domain. The N-terminal extension of ClpS would protrude toward the apical surface of D1, making it possible that the missing N terminus of ClpS might reach the edge of the apical surface of the ClpA hexamer. As mentioned earlier, a peptide segment (8 residues) bound to the N-domain with tentative sequence assignment of the disordered N terminus of ClpS is in a position to make a connection to the ClpS in the ClpAS model, which would allow the N terminus of ClpS to extend 20–30 Å out from the tip of the conical domain. In the C model, ClpS is bound to the N-domain projecting outward from the apical surface of the D1 ring (Fig. 7). With the N terminus protruding from ClpS, this configuration gives the complex the appearance of a jellyfish with six tentacles extending out from the substrate binding surface. It is conceivable that both the A and C models may be present when concentrations of ClpS reach a given threshold. At low ClpS concentrations, the A model would be expected to be the dominant species.

Substrate binding sites in ClpA, ClpX, and related proteins have been demonstrated at the apical surface of the D1 domain (20) and within the sensor and substrate discrimination (SSD) domains, located in ClpA D2 and ClpX D1 (21). The six D1 domains in the apical ring of ClpA enclose a cavity visible in
TABLE III

| ClpS species | ClpAP activity (%) |
|--------------|-------------------|
|              | [3H] Casein | GFP-SsrA |
| None         | 100         | 100     |
| ClpS         | 13          | 5       |
| ClpSΔN       | 104         | 100     |

ClpAP proteolytic activity was measured with 0.8 μg/ml ClpA and 20 μg/ml ClpP in standard assay buffer (see “Experimental Procedures”). Activity in the absence of inhibitor was 25 μg of casein and 3 μg of GFP-SsrA/hr of ClpA. ClpS or ClpSΔN was added at 0.5 μg to casein degradation assays and 3 μM to GFP-SsrA assays. ClpSΔN is missing 17 N-terminal amino acids.

both the electron microscopic reconstructions (23) and in the hexameric ring model based on the crystal structure of the ClpA subunit (5). Both N-terminally and C-terminally tagged substrates bind to the apical surface and are translocated along an axial pathway through the D1 cavity and eventually into ClpP (20, 47). SSD domains of different substrate recognition components of ATP-dependent proteases have preferences for different substrates. For example, the SSD of ClpA binds favorably to the heat shock transcription factor, σ32, whereas that of ClpX is more specific for SsrA-tagged proteins (21). It is possible that both types of substrate binding sites play a role in the catalytic cycle of ClpA.

The ClpA SSD domain (21) corresponds to residues 600–758, encompassing the sensor 2 motif and the small domain of D2 in the recently determined ClpA full-length structure (5). The small subdomain has an αβ-sandwich topology with three helices and a mixed three-stranded β-sheet. The contacts between the small and large subdomains of the AAA⁺ module are affected by the nucleotide state. Thus, the SSD domain offers a site for substrate-dependent regulation of activity and might provide a means of communication with the apical surface, where bound substrate undergoes unfolding and translocation.

The binding of ClpS to ClpA via the interface A places ClpS in contact with D2 and in a position to affect either substrate binding or changes in D2 induced upon substrate binding. However, this interaction is not sufficient to explain the inhibitory effect of ClpS. When the N-terminal 17 amino acids of ClpS were removed by lysylendopeptidase C treatment, the resulting ClpSΔN could still bind to ClpA and the isolated ClpS were removed by lysylendopeptidase C treatment, the binding or changes in D2 induced upon substrate binding.

The binding of ClpS to ClpA via the interface A places ClpS in a position to affect either substrate binding or changes in D2 induced upon substrate binding. However, this interaction is not sufficient to explain the inhibitory effect of ClpS. When the N-terminal 17 amino acids of ClpS were removed by lysylendopeptidase C treatment, the resulting ClpSΔN could still bind to ClpA and the isolated ClpS were removed by lysylendopeptidase C treatment, the binding or changes in D2 induced upon substrate binding.

Where is the N-terminal extension of ClpS when it is bound to ClpA hexamers? In the model of ClpS bound to ClpA hexamers using the favored A interface, the N terminus of ClpS would be far from the apical surface and probably would not be able to make contact.

In addition to the crystallographic and protease protection data, another factor favoring the A interface model is the ability of ClpS to protect ClpAP from autodegradation (15). Although the mechanism by which ClpS exerts its protective effect is not known, ClpS in the A model makes much more extensive contact with ClpA and is in a better position to protect large areas of the most sensitive regions of ClpA, whereas the ClpS in the C model is not in an obvious position to make a difference in ClpA stability. Taking the A interface as the one favored in solution, we can propose a model in which the N terminus of ClpS blocks substrate binding by competing with the substrate protein for binding to sites on the apical surface of ClpA. Since ClpS blocks degradation of soluble substrates but has apparent activating effects on degradation of aggregated proteins, we speculate that the N terminus of ClpS may block some but not all of the substrate binding sites on ClpA. By occupying substrate-binding sites on the apical surface of ClpA, ClpS would prevent binding of soluble substrates with single recognition sites and would thus impede their unfolding and degradation. At the same time, the blocking of some but not all of the sites could have an activating effect with aggregated protein substrates, which are expected to have many exposed sites for interaction with ClpA. So far, none of the models of ClpS action takes into account the probable mobility of the N-domain of ClpA itself, which would lead to significant changes in the position of bound ClpS and in its ability to affect substrate access to ClpA or binding to the apical surface. In the future, we hope to obtain crystals of ClpS bound to hexameric ClpA in different nucleotide states that should shed further light on the mechanism of action of this adaptor protein.

Acknowledgments—We thank Dr. G. Navrotski for helping us with the data collection at the BIOCARS-CAT beamline, Advanced Photon Source, Dr. Z. Dauter for assistance in the data collection at X9B of the National Synchrotron Light Source, and Drs. S. Gottesman and M. Gottesman for reading the manuscript and providing helpful comments.

REFERENCES

1. Gottesman, S., and Maurizi, R. M. (1992) Microbiol. Rev. 56, 592–621
2. Schirmer, E. C., Glover, J. R., Singer, M. A., and Lindquist, S. (1996) Trends Biochem. Sci. 21, 289–296
3. Wang, J., Hartling, J. A., and Flanagan, J. J. M. (1997) Cell 91, 447–456
4. Thompson, M. W., Singh, S. K., and Maurizi, M. R. (1994) J. Biol. Chem. 269, 18209–18215
5. Guo, F., Maurizi, R. M., Esser, L., and Xia, D. (2002) J. Biol. Chem. 277, 46743–46752
6. Newuald, A. F., Arvindl, L., Spague, J. L., and Konnin, E. V. (1999) Genome Res. 9, 27–34
7. Vale, R. D. (2000) J. Cell Biol. 150, F13–F19
8. Ogura, T., and Wilkinson, A. J. (2001) Genes Cells 6, 575–597
9. Zhou, Y., Gottesman, S., Hoskins, J. R., Maurizi, M. R., and Wickner, S. (2001) Genes Dev. 15, 627–637
10. Pratt, L. A., and Silhavy, T. J. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 2488–2492
11. Levchenko, I., Seidel, M., Sauer, R. T., and Baker, T. A. (2000) Science 289, 2354–2356
12. Turgay, K., Hahn, J., Burghoorn, J., and Duhau, G. (1998) EMBO J. 17, 6720–6728
13. Kondo, H., Rabouille, C., Newman, R., Levine, T. P., Pappin, D., Freemont, P., and Warren, G. (1997) Nature 388, 75–78
14. Meyer, H. H., Shorter, J. G., Seeemann, J., Pappin, D., and Warren, G. (2000) EMBO J. 19, 2181–2192
15. Doughan, D. A., Reid, B. G., Horwich, A. L., and Bukau, B. (2002) Mol. Cell 9, 673–683
16. Gottesman, S., Clark, W. P., and Maurizi, M. R. (1990) J. Biol. Chem. 265, 7886–7893
17. Gottesman, S., Roche, E., Zhou, Y., and Sauer, R. T. (1998) Genes Dev. 12, 1338–1347
18. Hoskins, J. R., Kim, S. Y., and Wickner, S. (2000) J. Biol. Chem. 275, 33631–33637
19. Weber-Ban, E. U., Reid, B. G., Mirankar, A. D., and Horwich, A. L. (1999) Nature 401, 90–93
20. Ishikawa, T., Beuron, F., Kessel, M., Wickner, S., Maurizi, M. R., and Steven, A. C. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 4328–4333
21. Smith, C. K., Baker, T. A., and Sauer, R. T. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 6678–6682
22. Singh, S. K., Boyczyk, J., Ortega, J., Ishikawa, T., Lo, J., Sennett, A. C., and Maurizi, M. R. (2001) J. Biol. Chem. 276, 29420–29429
23. Beuron, F., Maurizi, M. R., Belnap, D. M., Kocis, E., Booy, F. P., Kessel, M., and Steven, A. C. (1998) J. Struct. Biol. 123, 248–259
24. Owczarzak, Z., and Minar, W. (1997) Methods Enzymol. 276, 307–326
25. Weeks, C. M., and Miller, R. (1999) J. Appl. Crystallogr. 32, 120–124
26. Sheldrick, G. M. (1998) Direct Methods for Solving Macromolecular Structures.
Structure of the Heterodimer of ClpS and ClpA N-domain

Kluwer Academic Publishers, Dordrecht, The Netherlands
27. Otwinowski, Z. (1991) in Isomorphous Replacement and Anomalous Scattering: Proceedings of the CCP4 Study Weekend 25–26 January 1991 (Wolf, W., Evans, P. R., and Leslie, A. G. W., eds) pp. 80–86, SERC Daresbury Laboratory, Warrington, UK
28. Cowtan, K. D., and Main, P. (1996) Acta Crystallogr. Sect. D Biol. Crystallogr. 52, 43–48
29. Terwilliger, T. C. (2000) Acta Crystallogr. Sect. D Biol. Crystallogr. 56, 965–972
30. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
31. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
32. Vagin, A., and Teplyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025
33. Matthews, B. W. (1968) J. Mol. Biol. 33, 491–494
34. Chan, E. A. W., Hutchinson, E. G., Harris, D., and Thornton, J. M. (1993) Protein Sci. 2, 1574–1590
35. Holm, L., and Sander, C. (1993) J. Mol. Biol. 233, 123–138
36. Leijenmarck, M., and Liljas, A. (1987) J. Mol. Biol. 195, 555–579
37. Wahl, M. C., Bourenkov, G. P., Bartunik, H. D., and Huber, R. (2000) EMBO J. 19, 174–186
38. Kiseha, K., Moller, W., and Stoffler, G. (1971) Nat. New Biol. 233, 62–63
39. Moller, W., Schier, P. L., Maasen, J. A., Zantema, A., Schop, E., Reinalda, H., Cremer, A. F., and Mellem, J. E. (1983) J. Mol. Biol. 163, 553–573
40. Stark, H., Rodnina, M. V., Rinke-Appel, J., Brimaconbe, R., Wintemayer, W., and van Heel, M. (1997) Nature 389, 403–406
41. Le, J. H., Baker, T. A., and Sauer, R. T. (2001) Protein Sci. 10, 551–559
42. Lee, B., and Richards, F. M. (1971) J. Mol. Biol. 55, 379–400
43. Littlefield, O., and Nelson, H. C. (2001) Proteins Struct. Funct. Genet. 45, 219–228
44. Rees, D. C., Lewis, M., and Lipscob, W. N. (1983) J. Mol. Biol. 168, 367–387
45. Holmes, M. A., and Matthews, B. W. (1981) Biochemistry 20, 6912–6920
46. Taylor, A. B., Smith, B. S., Kitada, S., Kojima, K., Miyaura, H., Otwinowski, Z., Ito, A., and Deisenhofer, J. (2001) J. Mol. Biol. 315, 615–625
47. Ortega, J., Lee, H. S., Maurizi, M. R., and Steven, A. C. (2002) EMBO J. 21, 4938–4949
48. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
49. Esnouf, R. M. (1997) J. Mol. Graph. 15, 133–138
Crystal Structure of the Heterodimeric Complex of the Adaptor, ClpS, with the N-domain of the AAA + Chaperone, ClpA
Fusheng Guo, Lothar Esser, Satyendra K. Singh, Michael R. Maurizi and Di Xia

J. Biol. Chem. 2002, 277:46753-46762.
doi: 10.1074/jbc.M208104200 originally published online September 15, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M208104200

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 47 references, 18 of which can be accessed free at http://www.jbc.org/content/277/48/46753.full.html#ref-list-1