

Crystal Structure of LexA: A Conformational Switch for Regulation of Self-Cleavage

Yu Luo,¹ Richard A. Pfuetzner,¹ Steve Mosimann,¹ Mark Paetzel,¹ Elizabeth A. Frey,¹ Maia Cherney,² Baek Kim,^{3,5} John W. Little,^{3,4} and Natalie C.J. Strynadka^{1,6}

¹Department of Biochemistry and Molecular Biology
University of British Columbia
2146 Health Sciences Mall
Vancouver, British Columbia, V6T 1Z3
Canada

²Department of Biochemistry
University of Alberta
432 Health Sciences Building
Edmonton, Alberta, T6G 2H7
Canada

³Department of Biochemistry and Molecular
Biophysics and

⁴Department of Molecular and Cellular Biology
University of Arizona
Tucson, Arizona 85721

Summary

LexA repressor undergoes a self-cleavage reaction. In vivo, this reaction requires an activated form of RecA, but it occurs spontaneously in vitro at high pH. Accordingly, LexA must both allow self-cleavage and yet prevent this reaction in the absence of a stimulus. We have solved the crystal structures of several mutant forms of LexA. Strikingly, two distinct conformations are observed, one compatible with cleavage, and the other in which the cleavage site is ~20 Å from the catalytic center. Our analysis provides insight into the structural and energetic features that modulate the interconversion between these two forms and hence the rate of the self-cleavage reaction. We suggest RecA activates the self-cleavage of LexA and related proteins through selective stabilization of the cleavable conformation.

Introduction

LexA controls the SOS response of *E. coli* to conditions that damage DNA or inhibit DNA replication (Friedberg et al., 1995; Little and Mount, 1982). LexA represses a set of ~20 genes during normal growth. When the SOS response is triggered, RecA protein is activated, probably by binding to single-stranded DNA, to a form that interacts with LexA and stimulates its cleavage at a site between the N-terminal DNA binding domain and the C-terminal dimerization domain (Little et al., 1980). Cleavage inactivates LexA, derepressing the SOS genes. Hence, LexA cleavage plays a central role in this pathway.

Although RecA formally acts as an enzyme to catalyze

cleavage, several lines of evidence indicate that the chemistry of cleavage is carried out by groups in LexA itself, and that activated RecA serves indirectly as a coprotease to stimulate cleavage (Little, 1991, 1993). Although cleavage requires RecA in vivo, LexA can cleave itself in vitro in the absence of RecA (a reaction termed autocleavage; Little, 1984). Autocleavage is relatively rapid at high pH and hydrolyzes the same bond as in RecA-stimulated cleavage. In addition, mutations in LexA that interfere with RecA-stimulated cleavage also prevent autocleavage (Lin and Little, 1989). Several other related proteins, including phage repressors such as λ cI and mutagenesis proteins such as UmuD, undergo completely parallel cleavage reactions; in each case, autocleavage and RecA-stimulated cleavage cut the same bond (Burckhardt et al., 1988; Little, 1984). Cleavage of cI inactivates it, leading to the process of prophage induction (Ptashne, 1992; Roberts et al., 1978); by contrast, cleavage of UmuD to yield UmuD' activates this protein for its role in mutagenesis (Opperman et al., 1999).

LexA contains two types of sites involved in autocleavage. The substrate cleavage site is composed of the bond to be cleaved (between Ala-84 and Gly-85) and presumably the neighboring groups that confer specificity on that site. The active site (contained within the C-terminal domain) is composed of a catalytic center that carries out the chemistry of cleavage and, presumably, a binding pocket that binds the substrate and positions it optimally with respect to the catalytic center. The chemical mechanism is thought to involve a serine nucleophile (Ser-119), activated by a neutral general base lysine (Lys-156). Such Ser-Lys dyad mechanisms are becoming increasingly recognized in diverse hydrolases (e.g., Paetzel et al., 1998; Patricelli et al., 1999; Liao et al., 2000; Lejal et al., 2000). Ser-119 and Lys-156 are completely conserved in the LexA superfamily (Figure 1). Substitution of either residue with Ala completely blocks LexA cleavage (Sliatly and Little, 1987), and Ser-119 reacts selectively with a serine protease inhibitor (Roland and Little, 1990). Recent structural evidence with UmuD' (Peat et al., 1996a) and the C-terminal fragment of λ cI (Bell et al., 2000) shows that the corresponding Ser and Lys residues in these proteins are within hydrogen-bonding distance and lie at the end of a cleft on the surface of the protein, a configuration typical of protease active sites.

Although LexA is capable of self-cleavage, this reaction is evidently restrained by the structure of the protein. Based on the properties of several genetically screened mutant proteins (termed Ind^o) with markedly elevated rates of autocleavage, a conformation-equilibrium model for cleavage was proposed (Roland et al., 1992). According to this model, LexA exists in two distinct conformations, one in which cleavage can occur (termed L*) and one in which cleavage is prevented (termed L). The ratio of [L*]/[L] defines an equilibrium constant K_{conf} . In native LexA, the relative concentration of the noncleavable L form is much greater than the cleavable L* form (~10,000:1 at neutral pH) such that

⁵Present address: Department of Microbiology and Immunology, School of Medicine and Dentistry, University of Rochester, 601 Elmwood Ave., Rochester, New York 14642

⁶Correspondence: natalie@byron.biochem.ubc.ca

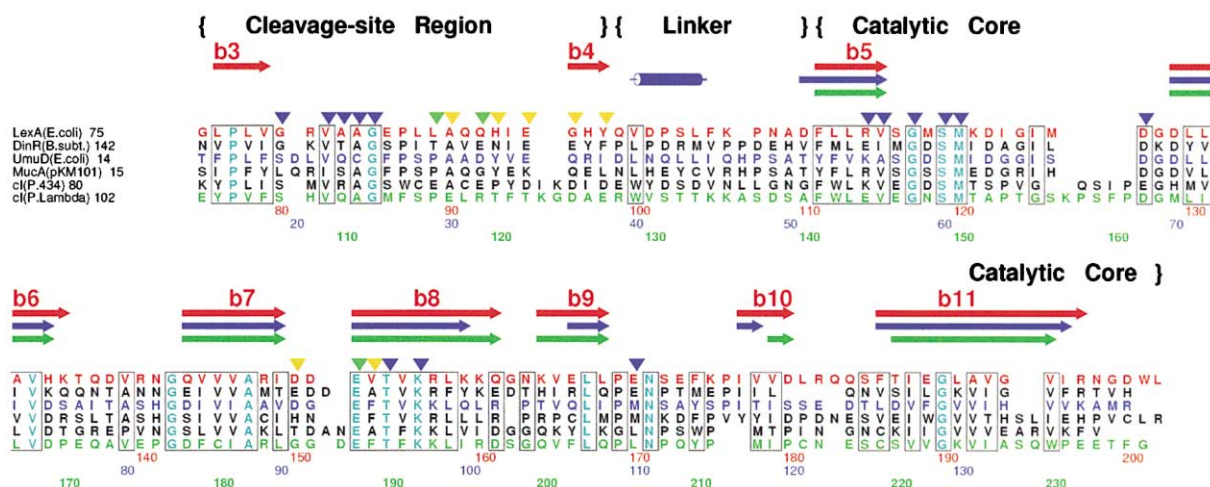


Figure 1. Sequence Alignment of the CTDs of the LexA Superfamily

Sequences of the C-terminal domains of 2 LexA repressors, 2 SOS response proteins, and 2 phage repressors, aligned by clustal-W (Thompson et al., 1994). Headings give the name of the protein, the organism name in parentheses, and the number of the starting sequence residue. Sequences, residue numbers, and β strands of *E. coli* LexA repressor, UmuD', and λ CI repressor are color coded in red, blue, and green, respectively. The CTDs contain a cleavage-site region (CSR), a structurally conserved catalytic core, and an intervening linker loop we propose to be involved in domain-swapping and intermolecular cleavage in this family of autoproteases. The CSR in LexA has a strand-loop-strand topology (b3-loop-b4). A segment of the LexA CSR assumes a variable conformation (residues Val-79 to Glu-95). The positions for RecA-specific mutants of λ CI, hypercleavable Ind^S mutants of LexA, and noncleavable Ind^N mutants of LexA are marked by brown, green, and blue triangles, respectively. The figure was created using Alscript (Barton, 1993). GenBank accession numbers are: *E. coli* *lexA*, J01643; *B. subtilis* *dinR*, M64684; *MucA*, M13388; *E. coli* *UmuD*, M13387; 434 *cI*, M12904; λ *cI*, J02459.

rates of autocleavage are very slow. It was proposed that the hypercleavable Ind^S mutations of LexA change this equilibrium toward the cleavable L* form, and that activated RecA does so to an even greater extent.

While it offers an attractive explanation for the mechanism by which cleavage is restrained, this proposal has not been supported by any biophysical evidence. Moreover, the cleavage-site region is missing from the previously determined UmuD' and λ CI fragment structures. Here we report the crystal structures of several forms of LexA. These structures all include the cleavage site region. Remarkably, we observe two distinct conformations of this region, which we believe represent the cleavable and noncleavable conformations.

Results and Discussion

Choice of Mutant Proteins

We sought to solve the structure of several mutant forms of LexA. As the native enzyme undergoes slow self-cleavage even in the absence of RecA, we designed constructs containing mutations that prevent cleavage such that the resulting structures would provide insight into the conformation(s) of the cleavage site region. The first, G85D, has a change in the Ala84-Gly85 cleavage site that blocks autocleavage (Lin and Little, 1989), but it should have a normal active site. The second, S119A, has a change in the active site Ser119 nucleophile that prevents cleavage (Siliaty and Little, 1987), but it should have a normal cleavage site. Collectively, these two mutants should allow for the reconstruction of the "native" LexA structure that is unamenable to structure determination (we acknowledge that subtle differences may still be possible). The third protein, here termed the quadru-

ple mutant (QM), contains three hypercleavable or Ind^S mutations (L89P, Q92W, and E152A), a combination which promotes extremely rapid RecA-independent cleavage *in vivo*, together with an additional mutation of the proposed general base, K156A, that blocks the chemistry of cleavage and appears to confer tight binding to RecA (Siliaty and Little, 1987). We rationalized such mutations would push the conformational equilibrium to favor the form required for cleavage and would allow us to trap this normally rare conformation in our structures.

Four structures were solved from three unique crystal forms: those of full-length S119A and G85D proteins, and those of tryptic fragments (residues 68–202) of the S119A and QM proteins (designated as Δ_{1-67}). In each of these structures, the cleavage site region is observed in one of two distinct and recurrent conformations (summarized in Table 1). As predicted by our design strategy, the structure of the Δ_{1-67} QM protein shows (for both molecules of the dimer in the asymmetric unit) a well-ordered cleavage site region in a conformation and location that is compatible with self-cleavage. On the other hand, the Δ_{1-67} S119A tryptic fragment shows a well-ordered cleavage site region in a conformation and location that is distant from the active site of the enzyme and incompatible with self-cleavage. Finally, the full-length S119A and G85D structures have one molecule in the asymmetric unit (molecule A) that is well-ordered throughout, including the cleavage site region which has the same noncleavable conformation as the Δ_{1-67} S119A tryptic fragment. The other molecule in the asymmetric unit (molecule B) is disordered in the N-terminal DNA binding domain, as well as either partially ordered (S119A) or fully disordered (G85D) in the cleavage site

Table 1. Crystallographic and Refinement Statistics

A: Crystallographic data				
Crystal	Δ_{1-67} S119A	S119A	G85D	Δ_{1-67} quadruple mutant
Space group	P43212	P43212	P43212	C2
Cell dimension (Å)	a, b = 49.88 c = 103.20	a, b = 89.75 c = 102.37	a, b = 90.02 c = 102.82	a = 124.7, b = 43.7, c = 49.5, β = 109.5
Resolution (Å)	2.0 (2.04–2.0) ^a	2.1 (2.15–2.0)	1.8 (1.86–1.8)	2.5 (2.54–2.50)
Observed reflections	90223	165809	143898	30913
Unique reflections	9882 (520)	24564 (1417)	38550 (3638)	8960 (486)
Completeness (%)	99.9 (100)	99.8 (99.5)	96.8 (93.8)	96.9 (95.1)
Redundancy	9.1 (8.5)	6.7 (5.9)	3.7 (3.3)	3.5 (3.0)
R _{merge} (%) ^b	7.2 (34.1)	7.4 (39.7)	4.8 (43.3)	7.4 (43.1)
I/ σ	32.5 (6.8)	26.6 (6.9)	28.4 (3.7)	18.2 (2.3)
B: Refinement				
Protomers/ASU	A	A, B	A, B	A, B
Residues observed	130	322	308	248
Ordered residues	A75–A204	A2–A198 B75–B87 B94–B199	A2–A198 B75–B79 B94–B199	A75–A198 B75–B198
CSR conformation ^c	NC	NC, C, partial disorder	NC, disorder	C, C
Resolution (Å)	15–2.0	15–2.1	15–1.8	15–2.5
R-factor/free R(%) ^d	23.0/27.8	23.0/27.3	24.0/26.3	22.0/28.4
Rmsd from ideality				
Bonds (Å)	0.0079	0.0075	0.0077	0.0070
Angles (degree)	1.51	1.41	1.35	1.52

^aValues in parentheses refer to the highest resolution shell.

^bR_{merge} = $\sum |I_h - \langle I \rangle_h| / \sum I_h$, where $\langle I \rangle_h$ is average over symmetry equivalents, h is reflection index.

^cCleavable conformation is listed as C; noncleavable conformation is listed as NC.

^dR-factor = $|F_{obs} - F_{calc}| / \sum F_{obs}$. The free R-factor is calculated using 10% of the observations not used throughout the refinement.

region. We predict the disorder is likely a result of crystal packing constraints in this region of the unit cell. Unexpectedly, the portion of the cleavage site region that is observed (with high temperature factors) in molecule B of S119A appears to adopt the cleavable conformation (very similar to that observed for the Δ_{1-67} QM protein). Again, an analysis of our structure in this region indicates that crystal packing constraints prevent the occurrence of the noncleavable conformation which is observed in the other S119A structures.

Overall Architecture

Full-length LexA is composed of a structurally distinct N-terminal DNA binding domain and a C-terminal catalytic domain (termed NTD and CTD, respectively), as supported from numerous earlier biochemical studies (Hurstel et al., 1986). In contrast to the typical textbook cartoons, which depict a long hinge region (Ptashne, 1992), our structures show the two domains are separated by only a short, highly hydrophilic and solvent-exposed region comprising residues Gln-70 to Glu-74 (Figures 1 and 2), with the subsequent cleavage site region (residues Gly-75 to Tyr-98) forming a structured and integral part of the globular C-terminal catalytic domain.

The NTD (Met-1 to Leu-69) contains 3 α helices (a1 to a3) followed by 2 antiparallel β strands (b1 to b2). The fold is essentially identical to that of the isolated NTD monomer determined previously by NMR (Fogh et al., 1994). The DNA binding domain of chain A shows minimal contact with the rest of the structure (~ 470 Å² of buried surface area). As mentioned, in our structures the NTD of chain B is disordered (Figures 2A and 2B;

Table 1), and thus, our structure provides no evidence for dimer contacts in the DNA binding domain.

The CTD (Gly-75 to Asn-198) is an exclusively β stranded structure (b3 to b11; Figure 2A). This domain consists of the cleavage-site region (Gly-75 to Tyr-98), an intervening linker (Gln-99 to Asp-110), and a catalytic core (Phe-111 to Asn-198; b5 to b11) containing the proposed Ser-119 nucleophile and Lys-156 general base. We identify this as the catalytic core based on its structural conservation in the crystal structures of UmuD' and the λ cl fragment (respective root-mean-square deviations are 0.96 Å and 1.36 Å for 81 superposed C α pairs).

LexA is observed as a dimer in all of our structures related by either crystallographic or noncrystallographic dyad symmetry (Table 1). In solution, a dimerization constant, K_d , of 5×10^{-5} M has been measured (Schnarr et al., 1985). The dimer interface is formed entirely by the CTD with a moderate buried surface area of 1380 Å². Portions of two loops (Gln-99-Asp-110 and Ser-116-Gly-128) and of the C-terminal β strand b11 participate in the dimeric interactions (Figure 2A). This dimer interface closely resembles those of the UmuD' (Ferentz et al., 1997) and λ repressor cl fragment structures (Bell et al., 2000).

The most striking and novel feature of the structures is that the cleavage site region, or CSR, is ordered and present in two distinct conformations (Figure 3). Both are highly structured (formed by the antiparallel β strands b3 and b4), and both are intimately associated with the catalytic core through parallel β sheet hydrogen bonds between b3 and b5. In one form, which we term the cleavable or C form (Figures 3C and 3D), the CSR (in red) lies in the active site, with the scissile bond between

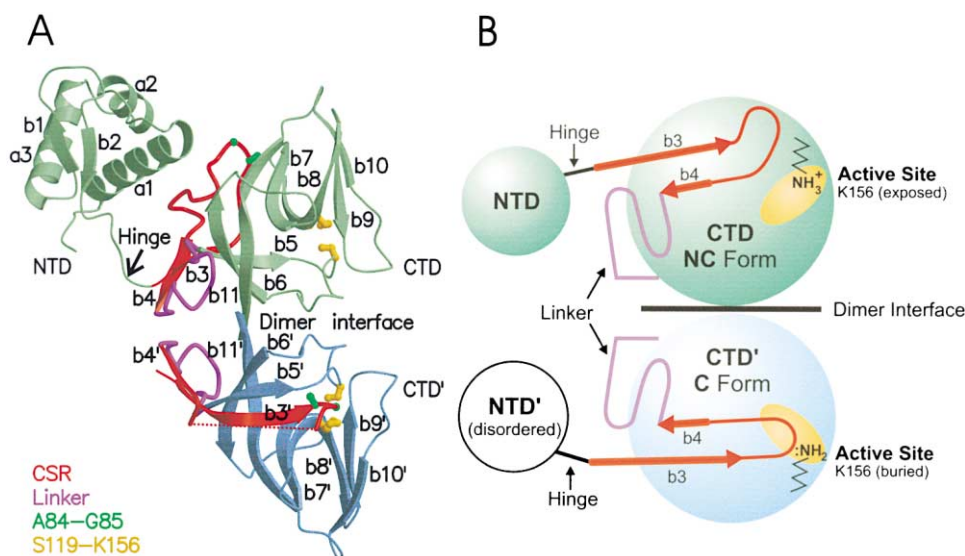


Figure 2. Overall Architecture of the LexA S119A Dimer

(A) One monomer is colored predominantly in light green, the other in blue. The dimer interface is marked and coincides with the 2-fold symmetry axis of the dimer. In monomer A (green), the NTD consists of three α helices (a1–a3), and two β strands (b1–b2), while the CTD consists of 9 β strands (b3–b11). Part of the CTD constitutes the catalytic core (b5 to b11) conserved between LexA, UmuD', and λ cl. In monomer B, the NTD is disordered and thus omitted from the figure. In both monomers, the CSRs, composed of b3, b4, and the intervening loop, are highlighted in red. The linkers between b4 (CSR) and b5 (catalytic core) proposed to allow for potential domain swapping and intermolecular cleavage in some members of the LexA superfamily are highlighted in purple. The active site Ser-119/Lys-156 dyad is colored in orange, the cleavage site Ala-84/Gly-85 in green. The Ser-119 side chain of the S119A mutant and the Lys-156 side chain of the quadruple mutant are reconstructed based on the structure of the G85D mutant.

(B) A schematic of the S119A dimer highlighting the key features of the structure. C refers to the cleavable conformation with the CSR (in red) bound in the active site (orange). NC refers to the noncleavable conformation with the CSR displaced from the active site. The linker loops proposed to potentially mediate intermolecular cleavage are shown in purple. Figures 2A to 6 were generated by Molscript (Kraulis, 1991) and Raster3D (Bacon and Anderson, 1988).

Ala-84-Gly-85 (in green) directly adjacent to the Ser-Lys active site dyad (in orange). In the other form, which we term the noncleavable or NC form (Figures 3A and 3B), the cleavage site of the partially restructured CSR lies ~ 20 Å away from the active site dyad. From the details of the structures (see below), we believe that the C and NC forms are, or closely resemble, the L* and L forms postulated previously on the basis of kinetics (Roland et al. 1992), and we propose that interconversion between C and NC represents the mechanism by which cleavage is controlled.

The LexA Active Site

The β -hydroxyl of the nucleophile Ser-119 forms a strong hydrogen bond to the ϵ -amino group of the proposed general base Lys-156 (2.9 Å and 3.1 Å in the 2 protomers of the G85D dimer). The conformations of the side chains in the Ser-Lys dyad are highly conserved between the LexA G85D structure and that of UmuD', λ cl, and the *E. coli* signal peptidase (which also has significant structural similarity in the catalytic core). In the C form of LexA, the Ser-119 O_γ sits ~ 2.7 Å from the Ala-84 C, with which it would react during cleavage (Figure 4A).

One feature of the conformation-equilibrium model is that the cleavable L* form has a markedly reduced pK_a for Lys-156 (estimated at ~ 5 – 6 ; Roland et al., 1992), compatible with its role as a general base in the reaction. In principle, the pK_a could be reduced by the presence

of adjacent positive charges, causing charge repulsion of the proton, or by placement in a buried and neutral environment, which would exclude charged atoms. The structure of C strongly favors the latter possibility. There are no charged amino acids within 8 Å of the ϵ -amino group of Lys-156. Lys-156 N_ϵ is somewhat exposed to solvent in NC, with a solvent-accessible surface of ~ 11 Å² (compared to 40–50 Å² for an entirely exposed N_ϵ), but becomes entirely buried in C (Figure 3C; Table 1). The side chain of Lys-156 forms strong van der Waals contacts with the hydrophobic residues Met-120 and Ile-177, which form part of the active site, and becomes further buried by additional residues from the cleavage site (Ala-84 to Pro-87) in C. The closest charge around the ϵ -amino group of Lys-156 is the γ -carboxylate of a highly conserved Glu-152 (~ 8 Å; Figure 4). As one might expect considering the long-range but complementary electrostatic charge, an E152A mutation at this position produces a 0.3 decrease in the apparent pK_a of cleavage and a 7-fold increase in the rate of autocleavage at pH 10 (Slilaty and Vu, 1991). A truncated LexA E152A enzyme also showed 10-fold elevated activity in the intermolecular cleavage reaction (Kim and Little, 1993).

In the mechanism for classic Ser-His-Asp serine proteases like chymotrypsin, a tetrahedral oxyanion intermediate is formed during catalysis which is favorably stabilized by a structural feature in the enzymes known as an "oxyanion hole" (Kraut, 1977). Our structure of LexA in the presence of substrate (the CSR) allows us

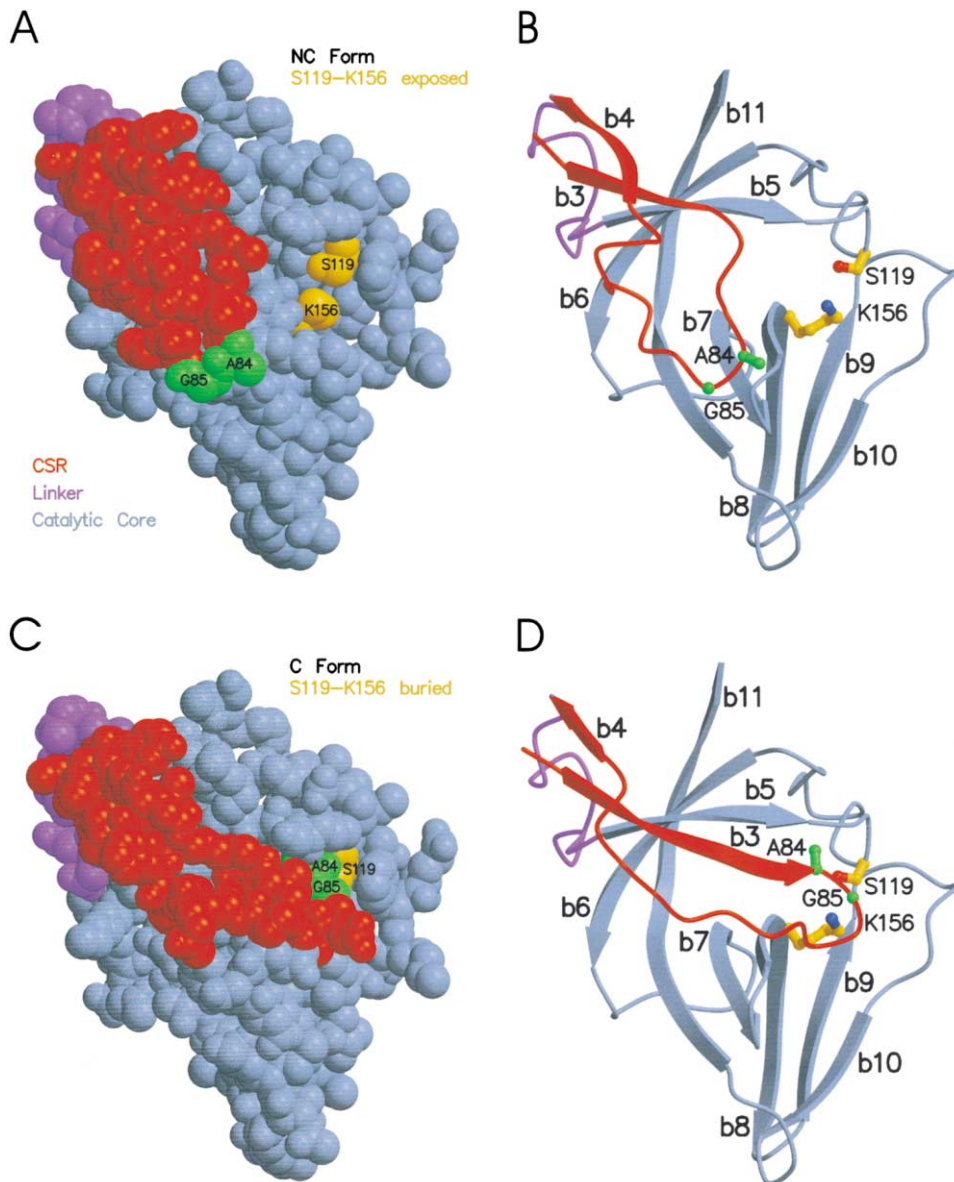


Figure 3. Two Conformational States of LexA

(A) A spacefilling representation of the CTD of the LexA Δ_{1-67} S119A mutant with the CSR in the NC form. Coloring is as for Figure 2.

(B) A ribbon representation of LexA Δ_{1-67} S119A mutant in the NC form in the same view as (A).

(C) A spacefilling representation of the CTD of the LexA Δ_{1-67} QM with the CSR in the C form. The Ser-Lys dyad is fully buried by the CSR and lies directly adjacent to the Ala-Gly cleavage site.

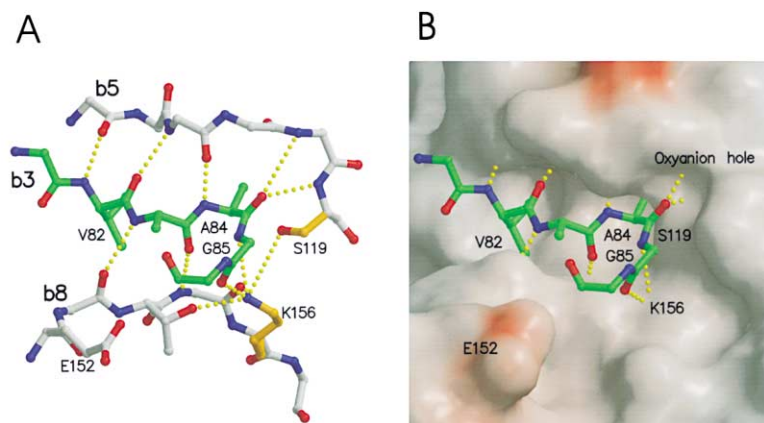
(D) A ribbon representation of the LexA Δ_{1-67} QM with the CSR in the C form in the same view as (C).

to uniquely define the oxyanion hole in the superfamily of Ser-Lys dyad proteases. In the C form of LexA, the main chain atom, Ala-84 O, lies within hydrogen-bond distance of two adjacent main chain amide nitrogens of Ser-119 (2.8 Å) and Met-118 (3.2 Å), and we suggest that this location is the oxyanion hole of LexA. The conformation of the oxyanion hole appears to be stabilized by hydrogen bonding of the adjacent glycine main chain atoms (Gly-117) to a solvent molecule and to the side chain of Asp-127. Mutations at Gly-117 in LexA and its counterpart in λ cl (Gly-147) have been shown to impair cleavage (Gimble and Sauer, 1986; Lin and Little, 1988),

as do mutations of Asp-127 in LexA (Silaty and Vu, 1991; Figure 5).

Substrate Specificity

The structure of the C form provides experimental evidence of substrate binding in the Ser-Lys dyad protease superfamily and reveals the basis for the substrate specificity of cleavage in LexA. The Ser-Lys dyad lies at the end of an extended hydrophobic cleft. In the C form, the CSR lies in this cleft with β strand b3 intercalated between β strands b5 and b8 of the catalytic core. The parallel β sheet hydrogen bonding observed between



the substrate CSR and the active site (Figure 4A) is unusual for a serine protease (typically the more energetically stable antiparallel hydrogen bonding is observed; Laskowski and Qasim, 2000).

The majority of members of the LexA superfamily cleave an Ala-Gly bond. In the C form, residues Arg-81 to Ala-84 of the substrate fit into the active site cleft, with the side chains of residues Val-82 and Ala-84 complementing the volume and hydrophobic nature of the corresponding hydrophobic pockets in the floor of the cleft (Figure 4B). Several previously isolated noncleavable (Ind⁻) mutations map to the CSR or to the substrate binding pockets, including V82M, V82E, A84T, A84D, G85D, and V115F (Lin and Little, 1988); in each case, the changes would create clashes with other atoms at the tightly packed substrate interface we observe in our structure, destabilizing the C form. At position 82, smaller side chains (V82A, V82G) or polar side chains with comparable size (V82S and V82T) autocleave poorly or not at all, but can undergo RecA-stimulated cleavage (Shepley and Little, 1996). Presumably, they destabilize C, but to a lesser extent that RecA can overcome.

Finally, selectivity on the universally conserved Gly-85 of the Ala-Gly cleavage site is likely due to its unusual conformation disfavored for other residue types ($\phi = +163$, $\psi = -142$). This glycol residue sits on the tip of a type IV turn with its N atom within hydrogen-bond distance to Lys-156 N ϵ . Our structure suggests its conformation would be critical for placing the glycol leaving group in close vicinity of the potential proton source of Lys-156. Mutant data support this notion; LexA G85D and the analogous λ cl G112E mutant proteins severely block cleavage, and λ cl G112A mutant autocleaves 5-fold slower than wild-type (Gimble and Sauer, 1986; Lin and Little, 1989).

Interconversion of the Cleavable and Noncleavable Conformations of LexA

Our structures suggest that C and NC represent two stable conformations that are stabilized by distinct interactions with the catalytic core. In both conformations, $\sim 1150 \text{ \AA}^2$ of solvent-accessible area is buried between the CSR and the conserved catalytic core with 12 and 7 mediating hydrogen bonds for C and NC, respectively. No buried salt bridges between the CSR and the catalytic core are present in either form. We infer that the C and NC forms are unlikely to be trapped by a high energy

Figure 4. Active Site Geometry of LexA

(A) A ball-and-stick model of the active site of the LexA Δ_{1-67} QM with the CSR in the C form. The Lys-156 side chain was modeled based on the refined model of the intact G85D mutant. The carbon atoms in the catalytic core and the CSR are colored in gray and green, respectively, while oxygen and nitrogen atoms are colored in red and blue regardless of their location. Hydrogen bonds are shown as dashed lines. Side chains not involved in substrate-recognition or catalysis have been omitted for clarity.

(B) The same view as (A) with the added electrostatic surface (GRASP; Honig and Nicholls, 1995). The electropositive surface is colored in blue, while the electronegative surface is colored in red.

barrier, and that the interconversion between the two forms may be rapid enough to reach equilibrium efficiently, compatible with earlier kinetic analysis of hypercleavable Ind⁺ LexA mutant proteins (Roland et al., 1992).

The conformational switch between C and NC appears to require only localized conformational changes in the CSR. In both forms, the catalytic core (Phe-111 to Asn-198) remains essentially identical (r.m.s. deviation 0.24 \AA for all 352 main chain atoms). The primary conformational change occurs in residues Val-79 to Glu-95. This change requires rotations around four consecutive main chain bonds (ψ of V79, ϕ and ψ of G80, and ϕ of R81) and extensive changes in the region from Gly-85 to Glu-95.

It has been estimated that the cleavable conformation accounts for $< 0.1\%$ of the total population of LexA at neutral pH in vitro (Roland et al., 1992). What forces modulate the interconversion between C and NC and determine their relative stabilities? The LexA structures suggest several possibilities, including a role for hydrophobic interactions, differential displacement of ordered solvent between the 2 forms, differential formation of an energetically unfavorable β bulge in the CSR, and importantly, the protonation of the general base Lys-156. We consider each of these contributions in turn.

The catalytic core provides an unusually extensive hydrophobic surface which accommodates the CSR in C and NC, respectively. In C, the substrate-recognition cleft provides hydrophobic pockets for binding the side chains of Val-82 and Ala-84 of the CSR (Figure 4B and 6B). In NC, this cleft is instead filled with ordered solvent molecules (bound to the main chain atoms of the flanking b5 and b8; Figure 6A), and the CSR binds to an adjacent hydrophobic surface on the catalytic core formed by the side chains of several conserved residues, including Phe-111, Leu-113, Val-146, and Val-153. Hydrophobic side chains of the CSR (Val-79, Val-82, Leu-88, and Leu-89) are buried at this site stabilizing NC (in C, the side chains of Leu-88 and Leu-89 are unfavorably exposed to solvent). We propose that the differential burial of hydrophobic residues, and the observation that the CSR must compete with ordered solvent molecules in order to gain access to the substrate-recognition cleft, may provide an added stability advantage to the NC form.

Previous mutagenesis studies also support the role of these hydrophobic interactions in modulating the

conformational interconversion. In particular, the hypercleavable Q92W stabilizes the C form, as observed in the structure of the Δ_{1-67} QM protein, in which the side chain of Trp-92 provides additional interactions with the exposed hydrophobic surface (Figure 6B). Furthermore, a favorable π -cation interaction (Gallivan and Dougherty, 1999) is observed between the aromatic ring of Trp-92 and the guanidiny group of Arg-148 (3.5 Å), explaining why aromatic residues at position 92 are particularly effective in elevating the rate of autocleavage.

An additional structural feature that may destabilize C is a β bulge formed by Val-79 N and Gly-80 N of the β strand b3 opposing Leu-112 O of β strand b5 (in NC, there is no such bulge). Even though the energy cost for forming a β bulge is estimated to be moderate by a statistical study of protein structures (Chan et al., 1993), it may contribute to the destabilization of C. Glycine has the highest propensity to form a β bulge. Mutations at Gly-80 would be expected to increase the energy cost to create a β bulge, therefore shifting the conformational equilibrium toward NC. Indeed, LexA G80D and G80V mutants block cleavage (Lin and Little, 1988, 1989).

Finally, a critical structural feature that will destabilize the C form is the protonation of the ϵ -amino group of Lys-156. In NC, our structures show this group is partially exposed to solvent, whereas in C it is completely buried by the bound substrate. It is highly likely that significant energy is required to drive the protonated lysine in the NC form to the fully buried neutral form required for catalysis in C. Earlier thermodynamic studies have shown that burial of a mutated lysine residue in the catalytic core of T4 lysozyme substantially destabilizes the enzyme (Dao-pin et al., 1991). We speculate that one advantage of a Ser-Lys dyad in the LexA superfamily over the more classic Ser-His-Asp triad lies in the intrinsically high pK_a of the lysyl ϵ -amino group. In this view, the energetic cost of burying the lysyl ϵ -amino group helps keep the autocleavage reaction in check.

Conservation of the CSR in the LexA Superfamily

Although the previously published structures of UmuD' and the λ cl fragment both lack the CSR, several lines of evidence suggest that the CSRs of these autocleavable proteins have a similar structure to that in LexA. β strand b3 is relatively well conserved in sequence (Figure 1). Pro-77 of LexA is highly conserved and the two subsequent residues are hydrophobic, suggesting a location in a buried environment similar to that in the interior β strand b3 of the LexA CSR. In addition, when P22 repressor is lightly treated with proteases like trypsin, a fragment spanning the CSR copurifies with the C-terminal core in several ion-exchange and size-exclusion steps (De Anda et al., 1983), suggesting both that the CSR interacts strongly with the rest of the CTD, and that a portion corresponding to b4 in LexA is relatively exposed and susceptible to protease. Finally, in all cases, RecA greatly stimulates cleavage at neutral pH (see also below). Taken together, these lines of evidence imply that, in addition to the same chemical mechanism of cleavage, members of the LexA superfamily share a common fold that includes the strand-loop-strand motif of the CSR. A folded CSR has clear advantages over a solvent-exposed and randomly mobile CSR, both in

terms of regulating cleavage and in minimizing nonspecific proteolysis in vivo.

The extensive hydrophobic surface, which alternatively stabilizes the CSR in LexA, is also relatively conserved amongst other members of the LexA superfamily, suggesting a similar general role in modulating the NC to C interconversion (Figure 1). However, there are differences in the size and specific nature of the hydrophobic surface, as well as the region corresponding to the conformationally variable region of the LexA CSR (Val-79 to Glu-95). Cleavage rates vary between members of the LexA superfamily, and have likely evolved to fit the biological role of each protein (Kim and Little, 1993). Our structures lead us to suggest that such "fine-tuning" of autocleavage rates could involve specific mutations within the CSR and the hydrophobic surface to which it binds.

If the CSR in other proteins resembles that of LexA, we note two other important implications from our structures. First, in vivo cleavage of UmuD is intermolecular, not intramolecular (McDonald et al., 1998). In the LexA structure, by contrast, the CSR interacts with the catalytic core in an intramolecular mode, and the conformation of the linker loop (Gln-99 to Asp-110) between the CSR and the catalytic core is the same in the C and NC forms. In the structure of the LexA dimer, this linker loop connects b4 of the CSR to b5 of the catalytic core (in purple in Figure 2). Our structure indicates that a relatively small conformational change of the Gln-99 to Asp-110 linker loop could result in a "domain swap" between CSR (monomer A) and CSR' (monomer B) such that an intermolecular cleavage would be facilitated. Three residues were shown to be critical for the UmuD intermolecular reaction (McDonald et al., 1999). The equivalent residues in LexA are 102, 106, and 109, respectively, which map to the linker loop (Figure 1). We propose that this linker can assume different conformations in various members of the LexA superfamily, thus favoring either intra- or intermolecular cleavage.

Second, as noted above, cleavage of UmuD activates its mutagenesis function. Cleavage of LexA should disrupt the structure of the remaining part of the CSR, since the buried strand b3 stabilizes b4 but is missing after cleavage, so that b4, and probably residues 99–110, would no longer be structured. A similar disruption in UmuD' offers a simple explanation for its altered function after cleavage. In particular, it would allow the "molecular dimer" interaction, permitting formation of UmuD' filaments (Peat et al., 1996b).

RecA-Stimulated Cleavage

How does RecA coprotease stimulate LexA cleavage at neutral pH? RecA may either act directly in the chemistry of cleavage, or it might act indirectly to favor a reactive conformation. The structure of LexA does not provide support for a direct role of RecA in the chemistry of cleavage. In the C form, the catalytic center is completely buried, and it would be unlikely for a side chain from RecA to gain access to it. If RecA contributes to catalysis by providing a positively charged side chain to reduce the pK_a of Lys-156, our structure indicates that the closest approach of this side chain would be limited to > 6 Å in the C form. For these reasons, we

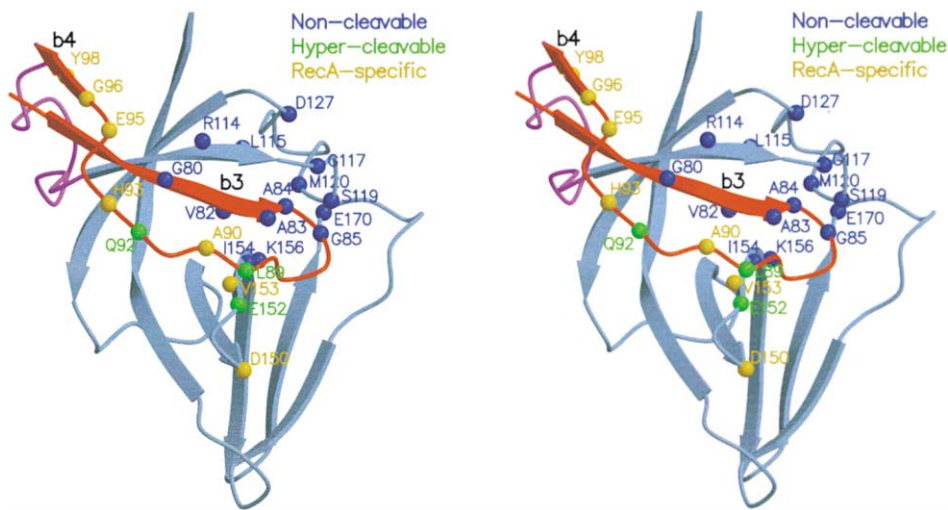


Figure 5. Mapping of Previously Characterized Mutants

A stereo ribbon representation of LexA (C form) with LexA Ind^- mutations (in blue), Ind^+ mutations (in green), and λ cl RecA-specific mutations (in brown) mapped on the structure (as based on Figure 1).

favor the possibility that RecA plays a more indirect role in catalysis by increasing the population of the cleavable conformation of LexA. Importantly, our structures show that RecA probably does not do so by binding at a distant site on LexA and favoring an allosteric change in the catalytic core, since this core is almost identical in C and NC. Instead, we propose that RecA binds preferentially to the C form, stabilizing it and promoting cleavage.

Sites of interaction with RecA are suggested by the properties of λ cl mutations that specifically interfere with RecA-stimulated cleavage while allowing RecA-independent autocleavage (RecA-specific mutations). The simplest interpretation of their properties is that they prevent cl from binding to RecA, although this has not been shown directly. Strikingly, several of the RecA-specific mutations in λ cl (Gimble and Sauer, 1986) are mapped to a region of our LexA structure (Figure 5) that includes surface exposed residues on the loop between b3 and b4 in the CSR.

The properties of several noncleavable Ind^- LexA mutant proteins are also consistent with preferential binding of RecA to the C form of LexA. In competitive inhibition assays, several mutant proteins were tested for their ability to inhibit RecA-stimulated cleavage of wild-type LexA (Lin and Little, 1989). Two mutant proteins (K156A and K156H) inhibit cleavage far better than does wild-type LexA, and bind tightly to RecA. Our structures suggest that these mutations favor the C form by removing the energetic cost of burying the charged Lys-156 ϵ -amino group. By contrast, several other Ind^- mutants are poor inhibitors and bind weakly to RecA. Our structures show that these mutations map to the oxyanion hole (G117E), the substrate binding pocket (V115F), or the CSR (G80D, G80V, V82M), each of which likely destabilize the C form. Clearly, the role of RecA will only be completely resolved when detailed kinetic data of the RecA-stimulated cleavage reaction and structural data of RecA in complex with a member of the LexA superfamily become available.

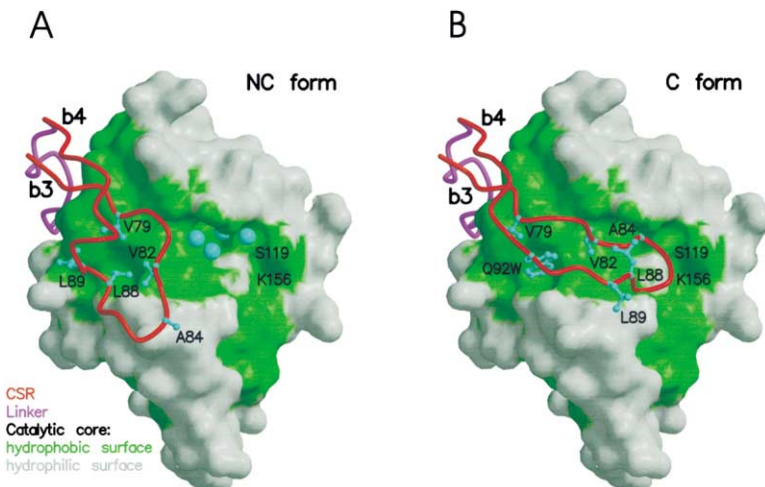


Figure 6. The Exposed Hydrophobic Surface of LexA

The catalytic core of LexA is shown in a molecular surface representation with the hydrophobic area highlighted in green (GRASP; Honig and Nicholls, 1995). The CSR and linker loop are shown as red and purple ribbons, respectively. The side chains of selected hydrophobic residues on the CSR that become differentially exposed to solvent are highlighted in a cyan ball and stick representation.

(A) NC form.
(B) C form.

Conclusion

This study provides the structural basis for both the mechanism of autocatalysis and the conformational switch that controls cleavage. This view of the cleavage site:active site complex provides insights into the serine/lysine dyad, oxyanion hole, and substrate-recognition features of the enzyme. Importantly, two distinct conformational states of a highly structured cleavage site region were captured in the crystal structures of noncleavable LexA mutants. The structural characteristics of the two states, and the structure-based explanation for biochemical properties of Ind^- , Adg^- , and Ind^{δ} LexA mutant proteins, collectively suggest that these states represent the cleavable and noncleavable forms in a previously proposed conformational-equilibrium model (Roland et al., 1992). The deprotonation of the general base Lys-156 required for catalysis appears to be coupled with its solvent accessibility. The conformational energy required to bury the lysine may create an energy barrier to keep the autocleavage reaction in check. Other structural features which may modulate the C and NC equilibrium include alternative hydrophobic interactions between the CSR and an extensive hydrophobic surface on the catalytic core, as well as a β bulge formed only in the C form of the CSR. These features may allow evolution of the conformational equilibrium to suit the biological requirements of each individual member of the LexA superfamily. Finally, we propose that RecA shifts the conformational equilibrium of LexA through a structural stabilization of the C form (including the requisite burial and neutralization of K156). A controlled shift of conformational equilibrium through the stabilization of a preexisting form is an emerging paradigm in a variety of signaling systems (Volkman et al., 2001).

Experimental Procedures

Protein Expression, Purification, and Crystallization

Intact LexA G85D mutant, S119A mutant, and quadruple mutant were overexpressed in BL21 (λ DE3) cells, and the resulting proteins were purified as described (Little et al., 1994). A C-terminal his-tagged S119A mutant protein was cloned (plus the C-terminal sequence VPRGSLEHHHHH) and overexpressed as above. The protein was purified using a Ni-affinity column (Pharmacia). Trypsin digestion was conducted with a 1:400 ratio of trypsin to LexA (wt/wt) for 30 min at room temperature; 5 mM benzamide was added to stop the reaction. Trypsin treatment removed residues 1–67, as well as the His-tag from variants with the tag. Δ_{1-67} fragments of the quadruple mutant and the His-tagged S119A mutant protein were further purified using a mono-Q anion-exchange column (Pharmacia). The Δ_{1-67} fragment of His-tagged S119A mutant protein retained the VPR portion of the C-terminal His-tag. Calculated molecular weights of the intact proteins and tryptic fragments were confirmed by mass spectrometry (not shown). Intact G85D and S119A proteins were crystallized in 1.8 M ammonium sulfate with 0.1 M BIS-Tris propane at pH 6.8. The Δ_{1-67} S119A mutant was crystallized in 1.6 M dipotassium monohydrogen phosphate with 0.3 M Tris-HCl at pH 8.5. The Δ_{1-67} quadruple mutant was crystallized in 30% PEG 1500 with 0.1 M MES buffer at pH 5.6. All four crystal forms were obtained by the hanging drop vapor diffusion method.

Data Collection

Room temperature X-ray diffraction data of intact LexA G85D and S119A mutant protein crystals were acquired at Brookhaven National Laboratory and Hamburg, respectively. The tryptic fragment data were collected in-house with crystals flash-cooled to 100K in an Oxford Cryosystem. The cryoprotectant for the Δ_{1-67} S119A mutant crystal contained 15% glycerol, 1.6 M potassium phosphate, and

0.2 M Tris-HCl buffer at pH 8.5; that for the Δ_{1-67} quadruple mutant crystal contained 8% glycerol, 30% PEG 1500, and 0.1 M MES buffer at pH 5.6. The in-house X-ray source was a Rigaku RU-200 generator equipped with Osmic mirrors operating at 50 KV/100 mA. Recorded images were processed by HKL (Otwinowski and Minor, 1997). Data collection statistics are provided in Table 1.

Structure Determination

The 2.0 Å crystal structure of the S119A tryptic fragment was solved by AMoRe (Navaza, 1994) using the coordinates of the published UmuD' model (PDB entry 1UMU) with appropriate ends trimmed according to the sequence alignment. The structures of the intact LexA G85D mutant protein (1.8 Å) and S119A mutant protein (2.1 Å) and the Δ_{1-67} quadruple mutant protein were solved using the refined model of the Δ_{1-67} S119A mutant protein. After 100 cycles of refinement and rebuilding using Arp/Warp (Lamzin and Wilson, 1997) against the 1.8 Å data of the G85D intact LexA, the majority of the model was satisfactorily traced, and the map became clearly interpretable for the rest of the structure. The models were fitted using XFIT (McRee, 1999) and O (Jones et al., 1991), and were subsequently refined using CNS (Brunger et al., 1998). Refinement statistics and geometry of the models are shown in Table 1. The coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 1JHC, 1JHE, 1JHF, and 1JHH.

Acknowledgments

We are grateful to Donald Shepley and Robin A. Roberts for excellent technical assistance, John Rupley for helpful discussions, and William Montfort for comments on the manuscript. Work from the Strynadka laboratory was supported by an NSERC operating grant and funding through the Canadian Institute of Health Research, the Burroughs Wellcome Fund, and the Howard Hughes Medical Institute International Scholar Program. Work from the Little laboratory was supported by NIH grant GM24178 and NSF grants DMB9004455 and MCB9305092. Fern Ness made Figure 2B.

Received April 23, 2001; revised July 24, 2001.

References

- Bacon, D.J., and Anderson, W.F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graphics* 6, 219–220.
- Barton, G.J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* 6, 37–40.
- Bell, C.E., Frescura, P., Hochschild, A., and Lewis, M. (2000). Crystal structure of the lambda repressor C-terminal domain provides a model for cooperative operator binding. *Cell* 101, 801–811.
- Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., and Pannu, N.S. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 54, 905–921.
- Burckhardt, S.E., Woodgate, R., Scheuermann, R.H., and Echols, H. (1988). UmuD mutagenesis protein of *Escherichia coli*: overproduction, purification, and cleavage by RecA. *Proc. Natl. Acad. Sci. USA* 85, 1811–1815.
- Chan, A.W., Hutchinson, E.G., Harris, D., and Thornton, J.M. (1993). Identification, classification, and analysis of beta-bulges in proteins. *Protein Sci.* 2, 1574–1590.
- Dao-pin, S., Anderson, D.E., Baase, W.A., Dahlquist, F.W., and Matthews, B.W. (1991). Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry* 30, 11521–11529.
- De Anda, J., Poteete, A.R., and Sauer, R.T. (1983). P22 c2 repressor. Domain structure and function. *J. Biol. Chem.* 258, 10536–10542.
- Ferentz, A.E., Opperman, T., Walker, G.C., and Wagner, G. (1997). Dimerization of the UmuD' protein in solution and its implications for regulation of SOS mutagenesis. *Nat. Struct. Biol.* 4, 979–983.
- Fogh, R.H., Ottleben, G., Ruterjans, H., Schnarr, M., Boelens, R., and Kaptein, R. (1994). Solution structure of the LexA repressor DNA

- binding domain determined by 1H NMR spectroscopy. *EMBO J.* **13**, 3936–3944.
- Friedberg, E.C., Walker, G.C., and Siede, W.S. (1995). DNA repair and mutagenesis (Washington DC: Am. Soc. Microbiol. Press).
- Gallivan, J.P., and Dougherty, D.A. (1999). Cation- π interactions in structural biology. *Proc. Natl. Acad. Sci. USA* **96**, 9459–9464.
- Gimble, F.S., and Sauer, R.T. (1986). Lambda repressor inactivation: properties of purified *ind*- proteins in the autodigestion and RecA-mediated cleavage reactions. *J. Mol. Biol.* **192**, 39–47.
- Honig, B., and Nicholls, A. (1995). Classical electrostatics in biology and chemistry. *Science* **268**, 1144–1149.
- Hurstel, S., Granger-Schnarr, M., Daune, M., and Schnarr, M. (1986). In vitro binding of LexA repressor to DNA: evidence for the involvement of the amino-terminal domain. *EMBO J.* **5**, 793–798.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. (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.
- Kim, B., and Little, J.W. (1993). LexA and lambda CI repressors as enzymes: specific cleavage in an intermolecular reaction. *Cell* **73**, 1165–1173.
- Kraulis, P.J. (1991). Molscript: a program to produce both detailed and schematic plots of protein structure. *J. Applied Crystallogr.* **24**, 946–950.
- Kraut, J. (1977). Serine proteases: structure and mechanism of catalysis. *Annu. Rev. Biochem.* **46**, 331–358.
- Lamzin, V.S., and Wilson, K.S. (1997). Automated refinement for protein crystallography. *Methods Enzymol.* **277**, 269–305.
- Laskowski, M., and Qasim, M.A. (2000). What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? *Biochim. Biophys. Acta* **1477**, 324–337.
- Lejal, N., Da Costa, B., Huet, J.C., and Delmas, B. (2000). Role of Ser-652 and Lys-692 in the protease activity of infectious bursal disease virus VP4 and identification of its substrate cleavage sites. *J. Gen. Virol.* **81 Pt 4**, 983–992.
- Liao, D.I., Qian, J., Chisholm, D.A., Jordan, D.B., and Diner, B.A. (2000). Crystal structures of the photosystem II D1 C-terminal processing protease. *Nat. Struct. Biol.* **7**, 749–753.
- Lin, L.L., and Little, J.W. (1988). Isolation and characterization of noncleavable (*Ind*-) mutants of the LexA repressor of *Escherichia coli* K-12. *J. Bacteriol.* **170**, 2163–2173.
- Lin, L.L., and Little, J.W. (1989). Autodigestion and RecA-dependent cleavage of *Ind*- mutant LexA proteins. *J. Mol. Biol.* **210**, 439–452.
- Little, J.W. (1984). Autodigestion of *lexA* and phage lambda repressors. *Proc. Natl. Acad. Sci. USA* **81**, 1375–1379.
- Little, J.W. (1991). Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease. *Biochimie* **73**, 411–421.
- Little, J.W. (1993). LexA cleavage and other self-processing reactions. *J. Bacteriol.* **175**, 4943–4950.
- Little, J.W., and Mount, D.W. (1982). The SOS regulatory system of *Escherichia coli*. *Cell* **29**, 11–22.
- Little, J.W., Edmiston, S.H., Pacelli, L.Z., and Mount, D.W. (1980). Cleavage of the *Escherichia coli* *lexA* protein by the *recA* protease. *Proc. Natl. Acad. Sci. USA* **77**, 3225–3229.
- Little, J.W., Kim, B., Roland, K.L., Smith, M.H., Lin, L.L., and Sillaty, S.N. (1994). Cleavage of LexA repressor. *Methods Enzymol.* **244**, 266–284.
- McDonald, J.P., Frank, E.G., Levine, A.S., and Woodgate, R. (1998). Intermolecular cleavage by UmuD-like mutagenesis proteins. *Proc. Natl. Acad. Sci. USA* **95**, 1478–1483.
- McDonald, J.P., Peat, T.S., Levine, A.S., and Woodgate, R. (1999). Intermolecular cleavage by UmuD-like enzymes: identification of residues required for cleavage and substrate specificity. *J. Mol. Biol.* **285**, 2199–2209.
- McRee, D.E. (1999). XtalView/Xfit—A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.
- Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. D* **50**, 157–163.
- Opperman, T., Muri, S., Smith, B.T., and Walker, G.C. (1999). A model for a *umuDC*-dependent prokaryotic DNA damage checkpoint. *Proc. Natl. Acad. Sci. USA* **96**, 9218–9223.
- Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
- Paetzel, M., Dalbey, R.E., and Strynadka, N.C. (1998). Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature* **396**, 186–190.
- Patricelli, M.P., Lovato, M.A., and Cravatt, B.F. (1999). Chemical and mutagenic investigations of fatty acid amide hydrolase: evidence for a family of serine hydrolases with distinct catalytic properties. *Biochemistry* **38**, 9804–9812.
- Peat, T.S., Frank, E.G., McDonald, J.P., Levine, A.S., Woodgate, R., and Hendrickson, W.A. (1996a). Structure of the UmuD' protein and its regulation in response to DNA damage. *Nature* **380**, 727–730.
- Peat, T.S., Frank, E.G., McDonald, J.P., Levine, A.S., Woodgate, R., and Hendrickson, W.A. (1996b). The UmuD' protein filament and its potential role in damage induced mutagenesis. *Structure* **4**, 1401–1412.
- Ptashne, M. (1992). A genetic switch: phage lambda and higher organisms (Cambridge, MA: Cell Press and Blackwell Scientific Publications).
- Roberts, J.W., Roberts, C.W., and Craig, N.L. (1978). *Escherichia coli* *recA* gene product inactivates phage lambda repressor. *Proc. Natl. Acad. Sci. USA* **75**, 4714–4718.
- Roland, K.L., and Little, J.W. (1990). Reaction of LexA repressor with diisopropyl fluorophosphate. A test of the serine protease model. *J. Biol. Chem.* **265**, 12828–12835.
- Roland, K.L., Smith, M.H., Rupley, J.A., and Little, J.W. (1992). In vitro analysis of mutant LexA proteins with an increased rate of specific cleavage. *J. Mol. Biol.* **228**, 395–408.
- Schnarr, M., Pouyet, J., Granger-Schnarr, M., and Daune, M. (1985). Large-scale purification, oligomerization equilibria, and specific interaction of the LexA repressor of *Escherichia coli*. *Biochemistry* **24**, 2812–2818.
- Shepley, D.P., and Little, J.W. (1996). Mutant LexA proteins with specific defects in autodigestion. *Proc. Natl. Acad. Sci. USA* **93**, 11528–11533.
- Sillaty, S.N., and Little, J.W. (1987). Lysine-156 and serine-119 are required for LexA repressor cleavage: a possible mechanism. *Proc. Natl. Acad. Sci. USA* **84**, 3987–3991.
- Sillaty, S.N., and Vu, H.K. (1991). The role of electrostatic interactions in the mechanism of peptide bond hydrolysis by a Ser-Lys catalytic dyad. *Protein Eng.* **4**, 919–922.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
- Volkman, B.F., Lipson, D., Wemmer, D.E., and Kern, D. (2001). Two-state allosteric behavior in a single-domain signaling protein. *Science* **291**, 2429–2433.

Accession Numbers

The coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 1JHC, 1JHE, 1JHF, and 1JHH.