

MIT Open Access Articles

Molecular basis of substrate selection by the N-end rule adaptor protein ClpS

The MIT Faculty has made this article openly available. *Please share* how this access benefits you. Your story matters.

Citation: Román-Hernández, Giselle et al. "Molecular basis of substrate selection by the Nend rule adaptor protein ClpS." Proceedings of the National Academy of Sciences 106.22 (2009): 8888-8893. © 2009 National Academy of Sciences

As Published: http://dx.doi.org/10.1073/pnas.0903614106

Publisher: United States National Academy of Sciences

Persistent URL: http://hdl.handle.net/1721.1/52560

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Molecular basis of substrate selection by the N-end rule adaptor protein ClpS

Giselle Román-Hernández^a, Robert A. Grant^a, Robert T. Sauer^a, and Tania A. Baker^{a,b,1}

^aDepartment of Biology, ^bHoward Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Tania A. Baker, April 1, 2009 (sent for review March 25, 2009)

The N-end rule is a conserved degradation pathway that relates the stability of a protein to its N-terminal amino acid. Here, we present crystal structures of ClpS, the bacterial N-end rule adaptor, alone and engaged with peptides containing N-terminal phenylalanine, leucine, and tryptophan. These structures, together with a previous structure of ClpS bound to an N-terminal tyrosine, illustrate the molecular basis of recognition of the complete set of primary N-end rule amino acids. In each case, the α -amino group and side chain of the N-terminal residue are the major determinants of recognition. The binding pocket for the N-end residue is preformed in the free adaptor, and only small adjustments are needed to accommodate N-end rule residues having substantially different sizes and shapes. M53A ClpS is known to mediate degradation of an expanded repertoire of substrates, including those with N-terminal valine or isoleucine. A structure of Met53A ClpS engaged with an N-end rule tryptophan reveals an essentially wild-type mechanism of recognition, indicating that the Met⁵³ side chain directly enforces specificity by clashing with and excluding β -branched side chains. Finally, experimental and structural data suggest mechanisms that make proteins with N-terminal methionine bind very poorly to ClpS, explaining why these high-abundance proteins are not degraded via the N-end rule pathway in the cell.

AAA+ ATPase adaptor | clpAP protease | degradation tag | degron

Regulated intracellular proteolysis is crucial in many physio-logical processes, including the elimination of misfolded or truncated proteins, the initiation of appropriate transcriptional responses to cellular stress, and the control of protein life span (1-4). In organisms ranging from bacteria to mammals, intracellular degradation is frequently catalyzed by multisubunit AAA+ proteolytic machines. The critical step in degradation is recognition. In bacteria, AAA+ proteases recognize appropriate substrates via short accessible peptide sequences, called degradation tags or degrons (4, 5). After binding, the protease uses repetitive cycles of ATP hydrolysis to unfold the substrate and then to translocate the denatured polypeptide into a sequestered chamber for degradation. The degrons of substrates can be recognized directly by the protease or recognized in an assisted fashion with the help of accessory proteins, called adaptors, which can deliver specific substrates to the protease and/or prevent degradation of other classes of substrates.

ClpS is a bacterial adaptor protein that delivers N-end rule substrates to the AAA+ ClpAP protease (6–11). The N-end rule proteolytic pathway relates the half-life of a protein in vivo to the identity of its N-terminal residue (12). In bacteria, 4 hydrophobic residues (Tyr, Phe, Trp, and Leu) serve as primary N-end rule degradation signals (13). ClpS facilitates degradation by binding directly to these signals and to the N-terminal domain of ClpA. A specific family of E3-ubiquitin ligases recognize eukaryotic N-end rule substrates and covalently modifies them with a polyubiquitin chain, which marks these substrates for degradation by the proteasome (14, 15). This family of E3 ligases shares a region of homology to ClpS (16), suggesting that eukaryotic and prokaryotic systems use a common mode of N-end rule recognition.

The crystal structure of ClpS bound to the N-terminal domain of ClpA is known (7, 8), as is the structure of ClpS in complex with a peptide containing an N-terminal tyrosine (17). Here, we present ClpS structures bound to peptides with N-terminal Trp, Phe, and Leu, the remaining primary N-end rule residues of bacteria. We also report the structure of the free ClpS adaptor. Together, these ClpS structures reveal the molecular principles of N-end rule recognition. In all cases, the unique α -amino group of the N-terminal amino acid is recognized via 3 highly-specific hydrogen-bonding interactions. Two of these hydrogen bonds are made directly by conserved ClpS residues, and a third involves a water molecule that bridges the α -amino group and a conserved ClpS side chain in all of the peptide-bound structures. The N-terminal side chains of the peptides bind in a deep hydrophobic pocket, which is also present in unliganded ClpS. Small changes in the structure of this pocket accommodate the binding of certain N-end rule side chains. Although interactions of ClpS with other parts of the bound peptide are observed in some structures, none of these contacts are conserved. It has been shown that M53A ClpS recognizes the standard set of N-end residues but also mediates degradation of substrates with N-terminal Val or Ile (17). We demonstrate that M53A ClpS recognizes an N-end rule peptide in exactly the same fashion as wild-type ClpS but has an expanded binding pocket that accounts for the extended substrate repertoire of this mutant. Model building shows that an N-terminal methionine could bind wild-type ClpS, but we find that the affinity of this interaction is extremely weak, providing protection for the enormous number of bacterial proteins that start with this amino acid.

Results

Crystal Structures. For structural studies, we used a truncated variant of *Caulobacter crescentus* ClpS (residues 35–119), which is stably folded and binds N-end rule peptides (17). We crystallized the free protein and complexes with peptides containing an N-terminal Trp, Leu, or Phe. The crystals belonged to space group P2₁ or P2₁2₁2₁ and diffracted to resolutions of 2.1 Å (apo), 1.5 Å (Trp), 1.85 Å (Leu), and 2.4 Å (Phe) (Table 1). Molecular replacement was used to obtain initial phases, and the structures were refined (Table 1). The quality of the electron-density maps ranged from very good to excellent (Fig. 1*A*).

The structure of the free *C. crescentus* ClpS adaptor was extremely similar to the peptide-bound structures of the same protein and the structure of *Escherichia coli* ClpS in complex with the N-terminal domain of ClpA (Fig. 1*B*) (rmsd < 0.5 Å for backbone atoms in all comparisons). This result indicates that no substantial rearrangements of the polypeptide backbone of ClpS occur upon binding of the adaptor to either N-end rule substrates or ClpA.

Author contributions: G.R.-H., R.T.S., and T.A.B. designed research; G.R.-H., R.A.G., and R.T.S. performed research; G.R.-H. contributed new reagents/analytic tools; G.R.-H., R.A.G., R.T.S., and T.A.B. analyzed data; and G.R.-H., R.T.S. and T.A.B. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 3G19, 3GQ0, 3GQ1, 3G1B, and 3GW1).

¹To whom correspondence should be addressed. E-mail: tabaker@mit.edu.

Table 1	. Cryst	tallograph	c data	and	refinement	statistics
---------	---------	------------	--------	-----	------------	------------

Protein data set	APO	Wt-Trp	Wt-Leu	Wt-Phe	M53A-Trp
d, Å	2.10	1.50	1.85	2.40	1.45
λ, Å	0.97927	1.5418	1.5418	0.97927	1.5418
Space group	P2 ₁	P2 ₁	P212121	P2 ₁	P21
Cell					
a, b, c, Å	27.7, 38.5, 62.6	33.5, 53.9, 45	28.6, 34.4, 71.7	27.6, 38.32, 62.86	33.6, 54.8, 44.9
α, β, γ, °	90, 88.4, 90	90, 110.3, 90	90, 90, 90	90, 90.1, 90	90, 110.5, 90
R _{sym} , %	14.1 (43.5)	4.6 (13.0)	4.3 (11.6)	11 (31.8)	5.7 (8.1)
No. of reflections	7,983	21,297	5,661	4,895	23,780
Completeness, %	98.2 (94.5)	87.6 (46.1)	87.1 (51.4)	96.1 (96.8)	88.7 (47.7)
Redundancy	10.5 (8.2)	7.2 (6.1)	6.0 (3.1)	2.8 (2.6)	5.2 (3.1)
R _{work}	0.221	0.160	0.185	0.277	0.173
R _{free}	0.264	0.180	0.212	0.308	0.194

Values in parentheses are for the highest resolution bin. $R_{sym} = \sum_h \sum_j |I_j(h) - \langle I(h) \rangle| / \sum_h \sum_j \langle I(h) \rangle$, where $I_j(h)$ is the *j*th reflection of index *h* and $\langle I(h) \rangle$ is the average intensity of all observations of I(h). $R_{work} = \sum_h |F_{obs}(h) - F_{calc}(h)| | / \sum_h F_{obs}(h)|$, calculated over the 95% of the data in the working set. R_{free} is equivalent to R_{work} except it is calculated over the 5% of the data assigned to the test set.

Peptide Recognition. In each of the 3 peptide-bound ClpS structures, the α -amino group of the peptide was coordinated by a set of 3 conserved hydrogen bonds: 1 with the side chain of His⁷⁹, 1 with the side chain of Asn⁴⁷, and 1 with a water molecule that also hydrogen-bonds to the side chain of Asp49 and the backbone carbonyl oxygen of the same peptide residue (see Trp example, Fig. 1C). The same set of hydrogen bonds was observed in the cocrystal structure of ClpS with an N-terminal Tyr peptide (17). Thus, ClpS recognizes the α -amino group of different N-end rule peptides by a completely-conserved mechanism. Similarly, in each cocrystal structure, the N-terminal peptide side chain (Trp, Phe, or Leu) packed into the deep hydrophobic pocket on the ClpS surface (Fig. 2), which was occupied by a Tyr in the structure reported by Wang et al. (17). Although additional contacts between ClpS and the bound peptides were present in some structures, none of these interactions were conserved. Collectively, this complete set of cocrystal structures demonstrates that ClpS binds to different N-end rule substrates in the same fashion, with the α -amino group of the peptide and the identity of the N-terminal side chain serving as the principal recognition determinants. Relatively poor electron density was observed for peptide residues other than the N-end position. Moreover, although contacts between ClpS and the second or third peptide residues were present in some structures, these interactions were not conserved among the different structures.

Interestingly, otherwise identical peptides with Leu, Phe, Tyr, or Trp at the N terminus all bind to ClpS with affinities between 150 and 500 nM (18). The cocrystal structures suggest that these similarities in binding affinity may occur, in part, because the binding pocket of ClpS accommodates variations in size, shape, and polarity between these N-end rule side chains.

The hydrophobic binding pocket of ClpS is comprised of side-chain and/or main-chain atoms from residues Ile⁴⁵, Leu⁴⁶, Asn⁴⁷, Asp⁴⁸, Asp⁴⁹, Thr⁵¹, Met⁵³, Val⁵⁶, Met⁷⁵, Val⁷⁸, His⁷⁹, and Leu¹¹² (Fig. 2). Subtle changes in the positions of these amino acids were observed in some structures, allowing accommodation of different N-end side chains. The largest changes were observed between the structure with the smallest N-end rule side chain, Leu, bound in the pocket and all of the other structures, including the peptide-free structure. Specifically, when the pocket was occupied by Leu, the conformations of several residues in the deepest part of the pocket changed to fill voids that would otherwise have been left. These changes include new side-chain rotamers for Ile⁴⁵ and Leu⁴⁶ and inward movements of ≈ 1 Å of the side chains of Val⁷⁸ and Leu¹¹² (Fig. 3*A*). The net effect of these structural rearrangements is that a Leu side chain

and a Trp side chain pack into the hydrophobic pocket almost equally tightly.

The side chains of Trp and Tyr also have polar atoms that need to form hydrogen bonds in the pocket to compensate for interactions that these groups would normally make with water in the unbound state. For Trp, the indole -NH of the side chain satisfies this requirement by donating a hydrogen bond to the main-chain carbonyl oxygen of Met⁷⁵ of ClpS (Fig. 3*B*). For Tyr, the phenolic -OH donates a hydrogen bond to the main-chain carbonyl oxygen of Leu⁴⁶ (Fig. 3*C*) (17).

Peptide Recognition by an Expanded Specificity Mutant. Wild-type ClpS does not recognize substrates with N-terminal Val or Ile. However, Wang et al. (17) showed that a mutant of *E. coli* ClpS could bind and deliver these classes of substrates for degradation if it contained a Met \rightarrow Ala substitution at the position corresponding to Met⁵³ in the *C. crescentus* adaptor. We expressed the M53A mutant of *C. crescentus* ClpS and obtained crystals in complex with a peptide with an N-terminal Trp. The structure revealed that the M53A variant recognizes the N-end rule peptide using the same fundamental mechanism as wild-type ClpS (Fig. 4). The α -amino group of the peptide formed the same conserved network of side chain- and water-mediated hydrogen bonds (Fig. 4*B*), and the Trp side chain packed into the hydrophobic pocket in the same manner observed in the wild-type structure (Fig. 4*A*).

The only significant differences between the wild-type and mutant structures was the presence or absence of the Met⁵³ side chain, which forms part of the hydrophobic pocket (Fig. 4*A*). Thus, the mutant has a larger binding pocket. Based on modeling studies, Wang et al. (17) predicted that the β -branched side chains of Val or Ile would clash with the Met⁵³ side chain. Indeed, when we modeled Val or Ile into the binding pocket of the M53A mutant, there were no steric clashes. Hence, our structural studies support the idea that the Met⁵³ side chain plays an important role in specificity by excluding β -branched side chains from the binding pocket.

Substrates with N-Terminal Methionine. Although β -brached side chains are sterically excluded from the ClpS binding pocket, this is not true of Met. Indeed, modeling studies showed that a Met side chain could be accommodated in the binding pocket observed in the Tyr-bound ClpS structure (17). Nevertheless, proteins with N-terminal Met are not recognized as N-end rule substrates (9, 13). This selectivity is biologically important, because the majority of bacterial proteins have an N-terminal Met (19, 20)

We considered the possibility that Met might be sterically



Fig. 1. Structures of ClpS reveal a common fold and conserved mechanism of peptide recognition. (*A*) Representative electron density (contoured at 1.25 σ) from a 2 $F_{\sigma} - F_{c}$ map for residues 78 and 79 of ClpS chain A and the N-terminal Trp-1 of the peptide in the Trp-bound ClpS structure. (*B*) Alignment of backbone C α positions for the apo structure (light blue), the Leu-bound structure (orange), the Trp-bound structure (green), and the Tyr-bound structure (red) of *C crescentus* ClpS and the *E. coli* ClpS structure (pink; Protein Data Bank ID code1R60) from a complex with the ClpA N-terminal domain. (*C*) Hydrogen bonds between the α -amino group of the N-end Trp residue, a water molecule (red sphere) and ClpS, pink for the peptide).

allowed in the larger pocket observed in the apo, Trp-bound, Tyr-bound, and Phe-bound structures but not in the smaller pocket that optimizes packing of Leu. However, we were able to model Met into this smaller pocket without significant steric clashes (Fig. 5A). Nevertheless, only a single Met rotamer (mmt) was allowed, and this rotamer is found in only 2% of the methionines in the structure database (21). By contrast, the Leu rotamer in our structure is observed in 59% of all cases. If we assume that the distribution of side-chain rotamers in native structures approximates the distribution in unstructured molecules in solution, then we can calculate the expected reduction



Fig. 2. The hydrophobic binding pocket of ClpS. Views of the pocket from the apo structure and structures with bound Leu, Phe, and Trp peptides are shown. In each case, the N-end rule peptide is shown in stick representation (pink carbons), and *C. crescentus* ClpS residues 45, 48, 49, 51, 53, 56, 75, 78, 79, and 112 are shown in surface representation.

in side-chain entropy upon binding of a peptide with an Nterminal Leu or Met to ClpS. The side-chain entropic penalty for Leu binding at 37 °C is 0.3 kcal/mol ($-RT \cdot ln[0.59]$), whereas that for Met is 2.4 kcal/mol ($-RT \cdot ln[0.02]$). Hence, the free energy



Fig. 3. Side-chain specific changes in the ClpS binding pocket. (A) Overlay of the binding pockets from the apo structure (red) and the Leu-bound structure (yellow). New side-chain rotamers are adopted by ClpS residues lle^{45} and Leu^{46} to fill voids in the deepest part of the pocket and several other groups move to improve packing with the Leu side chain (pink). A van der Waals surface for lle^{45} and Leu^{112} in the Leu-bound pocket is shown by the yellow clouds; the surface of Leu¹ of the peptide is in pink. (*B*) The indole – NH of the N-end Trp side chain donates hydrogen bond to the backbone carbonyl oxygen of Met⁷⁵ in the binding pocket of ClpS. (C) The phenolic – OH of the N-end Tyr side chain donates a hydrogen bond to the backbone carbonyl oxygen of Leu⁴⁶ in the binding pocket of ClpS (Protein Data Bank ID code 3DNJ; ref. 17).



Fig. 4. Structure of ClpS^{M53A} bound to the Trp peptide. (*A*) Overlay of the parts of the binding pockets in the Trp-bound ClpS^{M53A} structure (orange) and Trp-bound wild-type ClpS (blue). The N-end Trp is shown in pink. (*B*) Hydrogen bonds with the α -amino group of the N-end Trp (pink) in the ClpS^{M53A} structure (pink). Atom colors are: oxygen (red), nitrogen (blue), and carbon (purple for ClpS, pink for the peptide). The same contacts are observed as seen in Fig. 1C.

of Met binding to ClpS would be reduced by ≈ 2.1 kcal/mol compared with Leu and its affinity would be 30-fold lower based on considerations of side-chain entropy alone.

To probe for any ClpS-mediated ClpAP degradation of a substrate with an N-terminal Met, we assayed proteolysis of a



Fig. 5. Met binds poorly to the ClpS recognition pocket. (A) An N-terminal Met side chain (green) was modeled into the binding pocket from the Leubound ClpS structure. Note that the Leu (pink) side chain and Met side chain share common $C\alpha$, $C\beta$, and $C\gamma$ positions. The $C\delta 2$ of Leu and $S\delta$ of Met are also close. However, the Leu C $\delta 1$ and Met Ce occupy very different positions. (B) Kinetics of degradation of ³⁵S Met-titin-I27 degradation by ClpAPS. (C) ClpS binds poorly to a peptide containing N-terminal Met (red) or NIe (black) but binds tightly to an identical peptide with an N-terminal Leu (blue) ($K_D = 214$ nM). Peptide binding was measured by change in fluorescence anisotropy. (*Inset*) From top to bottom, Leu, Met, and NIe side chains.

³⁵S-labeled variant of titin-I27 by release of acid-soluble radioactivity (Fig. 5*B*). Degradation was extremely slow. The rate changed linearly with substrate concentration, allowing us to calculate $V_{\text{max}}/K_{\text{M}}$ but not values for either individual kinetic constant. We estimated a K_{M} value 200–600 μ M, by assuming a V_{max} value between 2 and 6 min⁻¹enz⁻¹, which have been reported for ClpS-mediated ClpAP degradation of titin-I27 substrates with authentic N-end residues (10). These estimated K_{M} values suggest that ClpS interacts with proteins carrying an N-terminal Met ~1,000-fold more weakly than it interacts with substrates carrying true N-end rule residues. For example, changing the N-terminal Met of the titin-I27 substrate used above to Tyr reduces the K_{M} for degradation to 600 nM (10).

Because changes in side-chain entropy appear to account for only approximately half of the expected reduction in binding of ClpS to a substrate with an N-terminal Met, we considered the possibility that the sulfur in the Met side chain is less hydrophobic than a methylene group or packs less well into the ClpS pocket. To test this model, we synthesized 2 fluorescein-labeled peptides that differed only in having Met or norleucine (Nle) at the N terminus and measured binding to ClpS by fluorescence anisotropy. Both peptides bound weakly to ClpS, and, in each case, binding was not saturated at the highest peptide concentration that could be tested (Fig. 5C). Nevertheless, the peptide with the N-terminal Nle bound no more tightly to ClpS than the peptide with the N-terminal Met. Because Met and Nle differ only in the presence of a sulfur atom or methylene group at the δ position of the side chain, we conclude that poor ClpS binding of the Met peptide is not a consequence of the presence of the sulfur.

Leu is branched at the γ position, whereas Nle and Met are unbranched. As a consequence, the C δ 1 methyl group of Leu occupies a very different position in the ClpS pocket than the position modeled for the C ε methyl group of Met or Nle (Fig. 5A). Hence, it seems likely that these differences in structure give rise to changes in van der Waals packing, which favor Leu over its straight-chain cousins. Thus, considerations of packing and side-chain entropy could plausibly account for the fact that an N-terminal Leu binds ClpS \approx 1,000-fold more tightly than an N-terminal Met or Nle.

Discussion

The ClpS adaptor recognizes substrates bearing N-terminal Tyr, Leu, Phe, or Trp side chains and delivers them to the ClpAP protease for degradation (9, 10). We previously reported the structure of C. cresentus ClpS bound to an N-end Tyr peptide (17), and, here, present structures of the peptide-free adaptor and complexes with peptides with N-terminal Leu, Phe, and Trp. Despite differences in size and shape, the N-terminal side chain of each peptide fits snugly into a hydrophobic pocket on ClpS. In each case, the peptide α -amino group also forms 3 conserved hydrogen bonds either directly with ClpS side chains or via an invariant water molecule, which itself hydrogen-bonds to the peptide and ClpS. Thus, ClpS binds all N-terminal residues of the N-end rule pathway in a very similar manner. In all of the peptide-bound ClpS structures, only a few contacts are observed with additional residues of the peptide and these interactions are not conserved. This result is consistent with biochemical studies that show that amino acids after the first peptide position play only minor roles in ClpS affinity (10, 18). ClpS is homologous to an E3 ubiquin ligase domain that mediates N-degron recognition in eukaryotes (16), suggesting that the mode of N-end substrate recognition is highly conserved in all organisms.

Although all primary N-end rule side chains are generally hydrophobic, they vary substantially in size and shape. For example, the side chain of Trp is substantially larger than that of Leu. Nevertheless, ClpS interacts with substrates carrying either of these N-terminal residues with similar affinities (10, 18). Our structures help understand this observation. Although no global structural changes occur upon ClpS binding to N-end rule peptides, we do observe minor adjustments in the hydrophobic pocket. For example, several ClpS side chains that form the pocket adopt new rotamer conformations or move inward to ensure tight packing around the smaller Leu side chain. These movements eliminate potentially unfavorable voids or cavities in the Leu-bound structure.

The hydrophobic pocket in peptide-free ClpS is essentially identical in structure to the pockets in the Trp-bound, Tyrbound, and Phe-bound structures. Thus, the only conformational adjustments that would be required for binding of peptides with these N-terminal residues would involve selection of the proper N-end rotamers, which were generally the most frequently observed in the Protein Data Bank. Thus, other than loss of translational and rotation degrees of entropic freedom, only a small additional conformational entropic cost would accompany binding of these N-end rule residues to ClpS.

β-Branched amino acids, such as Val, are not recognized as N-end residues by ClpS, and thus proteins bearing this Nterminal residue are not degraded by this pathway (9, 13, 17). Indeed based on predictions and direct measurements, between 2% and 3% of proteins in the E. coli proteome are thought to begin with Val, because methionine aminopeptidase removes the initiator Met from proteins that are synthesized with Met–Val at the N terminus (19, 22, 23). Wang et al. (17) demonstrated that the Met⁵³ side chain in ClpS plays a major role in excluding Val and Ile, because the restriction against these β -branched residues was relieved when the "gatekeeper" Met was replaced by Ala. Here, we show that M53A ClpS binds an N-end rule peptide using exactly the same recognition contacts as wild-type ClpS. Indeed, the only substantial difference in the M53A cocrystal structure is the absence of the Met⁵³ side chain (Fig. 4A). Because the binding pocket in the M53A adaptor is not distorted and modeling shows that the Met⁵³ side chain in the wild-type pocket would clash with β -branched side chains, simple steric exclusion ensures that ClpS does not target proteins with N-terminal Val or other β -branched residues for ClpAP degradation.

Previous studies indicate that bacterial proteins with Nterminal Met are not targeted for N-end rule degradation (9, 10, 13). Because Met has a relatively hydrophobic side chain, which often substitutes for Leu in protein structures (24), it was not immediately apparent how unwanted ClpS-mediated degradation of proteins with N-terminal Met is avoided. We find that substrates with N-terminal Met bind ClpS with \approx 1,000-fold reduced affinity when compared with authentic N-end rule residues. Two principal mechanisms appear to be responsible for feeble binding. First, only a single relatively rare Met rotamer can be accommodated in the ClpS pocket without significant steric clashes. Based on the low frequency at which this rotamer occurs in the Protein Data Bank, we estimate that, upon binding, the loss of conformational sidechain entropy alone would reduce Met affinity for ClpS \approx 30-fold compared with Leu. Moreover, this Met rotamer does not fill the pocket as efficiently as does Leu. Thus, we suspect that poorer van der Waals packing accounts for the remaining loss of binding affinity of ClpS for an N-terminal Met side chain. We found that Met and its isostere, Nle, bound

- 1. Gottesman S (2003) Proteolysis in bacterial regulatory circuits. Annu Rev Cell Dev Biol 19:565–587.
- Sauer RT, et al. (2004) Sculpting the proteome with AAA+ proteases and disassembly machines. Cell 119:9–18.
- Hanson PI, Whiteheart SW (2005) AAA+ proteins: Have engine, will work. Nat Rev Moll Cell Biol 6:519–529.
- Mogk A, Schmidt R, Bukau B (2007) The N-end rule pathway for regulated proteolysis: Prokaryotic and eukaryotic strategies. *Trends Cell Biol* 17:165–172.
- Baker TA, Sauer RT (2006) ATP-dependent proteases of bacteria: Recognition logic and operating principles. Trends Biochem Sci 31:647–653.

with similar affinity to ClpS. Thus, the weak binding of Met to ClpS does not appear to be a function of the sulfur atom in this side chain.

Our cocrystal structures of ClpS bound to each of the 4 primary N-end rule residues reveal a simple set of chemical and physical principles that govern N-end recognition by this proteolytic adaptor. These structures, mutant analysis, and biochemical experiments show that the landscape or N-end rule specificity depends both on positive and negative features of recognition and exclusion. These principles of specificity will likely be important to dictating substrate choice in other AAA+ protein adaptors.

Materials and Methods

Proteins and Peptides. Variants of *C. crescentus* ClpS^{35–119} were used for all structural experiments (17). *E. coli* ClpS and *E. coli* ClpAP were purified as described (10, 25, 26). Peptides used for binding studies (NH₂-MLYVQRDSKEC-COOH; NH₂-Nle-LYVQRDSKEC-COOH; and NH₂-LLYVQRDSKEC-COOH) were synthesized by standard Fmoc technique using a solid phase peptide synthesizer (Apex 396), labeled with fluorescein maleamide (Thermo Scientific), and purified by HPLC. Peptides used for crystallography were synthesized (NH₂-LL-COOH) or purchased from Sigma (NH₂-LLL-COOH and NH₂-FGG-CCOH).

Peptide Binding and Protein Degradation Assays. Binding of fluoresceinlabeled peptides to ClpS was measured by monitoring changes in fluorescence anistropy at increasing ClpS concentration by using a Photon Technology International instrument as described (18). Protein degradation was also performed as described (18). Briefly, rates of degradation of ³⁵S-labeled variants of titin-I27 were measured as a function of substrate concentration with 0.5 μ M *E. coli* ClpA₆, 1.0 μ M *E. coli* ClpP₁₄, and 4.5 μ M *E. coli* ClpS. Degradation was initiated by the addition of an ATP-regeneration mix (4 mM ATP, 50 mg/mL creatine kinase, and 5 mM creatine phosphate). Time points were removed at 10-min intervals for 30 min, and degradation was quantified by the release of ³⁵S-peptides that were soluble in trichloroacetic acid. Degradation rates were normalized by dividing by the total ClpA₆ concentration.

Crystallography. Crystals of *C. crescentus* ClpS^{35–119} or the M53A variant were grown with or without added peptide at 20 °C in hanging drops containing 0.5 μ L of ClpS (3–5 mg/mL) in 10 mM Hepes (pH 7.5), 200 mM KCl, and 1 mM DTT mixed with 1.5 μ L of reservoir solution containing 0.1 M bis-Tris-HCl (pH 5.5), 0.2 M MgCl₂ (0.025 M for peptide-free crystals), and 14–25% PEG 3350.

Crystals were frozen without additional cryoprotection. X-ray diffraction data for the Leu, Trp, and M53A mutant ClpS were collected on a Rigaku MicroMax007-HF rotating anode source equipped with Varimax-HR mirrors, an RAXIS-IV detector, and an Oxford cryosystem. Diffraction data for the free adaptor (Apo) and the Phe cocrystal structure were collected at the NE-CAT 24-ID-E beamline at the Argonne National Labs Advanced Photon Source (Argonne, IL). Data were processed with HKL-2000 (27). Initial phases for all structures were obtained by molecular replacement using the previously published *C. crescentus* structure (Protein Data Bank ID code 3DNJ) as a search model in PHASER (28). COOT was used for model building (29), and PHENIX was used for refinement (30).

ACKNOWLEDGMENTS. We thank Amy Keating, Scott Chen, and Sarah Bissonnette for helpful discussions and advice and Kevin Wang for advice, assistance, and support. This work was supported by the Howard Hughes Medical Institute and National Institutes of Health Grants GM-49924 and Al-16892. T.A.B. is an employee of the Howard Hughes Medical Institute. Studies at the NE-CAT beamlines of the Advanced Photon Source were supported by National Institutes of Health–National Center for Research Resources Award RR-15301 and the Department of Energy Office of Basic Energy Sciences under Contract DE-AC02-06CH11357.

- Dougan DA, Reid BG, Horwich AL, Bukau B (2002) ClpS, a substrate modulator of the ClpAP machine. *Mol Cell* 9:673–683.
- Guo F, Maurizi MR, Esser L, Xia D (2002) Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. J Biol Chem 277:46743–46752.
- 8. Zeth K, et al. (2002) Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. *Nat Struct Biol* 9:906–911.
- 9. Erbse A, et al. (2006) ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. Nature 439:753–756.
- Wang KH, Sauer RT, Baker TA (2007) ClpS modulates but is not essential for bacterial N-end rule degradation. *Genes Dev* 21:403–408.

- Hou JY, Sauer RT, Baker TA (2008) Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. Nat Struct Mol Biol 15:288–294.
- Varshavsky A (1996) The N-end rule: Functions, mysteries, uses. Proc Natl Acad Sci USA 93:12142–12149.
- Tobias JW, Shrader TE, Rocap G, Varshavsky A (1991) The N-end rule in bacteria. Science 254:1374–1377.
- Barter B, Wunning I, Varshavsky A (1990) The recognition component of the N-end rule pathway. EMBO J 9:3179–3189.
- 15. Tasaki T, Tae Kwon Y (2007) The mammalian N-end rule pathway: New insights into its components and physiological roles. *Trends Biochem Sci* 32:520–528.

v

- Lupas AN, Koretke KK (2003) Bioinformatic analysis of ClpS, a protein module involved in prokaryotic and eukaryotic protein degradation. J Struct Biol 141:77–83.
- Wang KH, Roman-Hernandez G, Grant RA, Sauer RT, Baker TA (2008) The molecular basis of N-end rule recognition. *Mol Cell* 32:1–9.
- Wang KH, Oakes E, Sauer RT, Baker TA (2008) Tuning the strength of a bacterial N-end rule signal. J Biol Chem 283:24600–24607.
- Frottin F, et al. (2006) The proteomics of N-terminal methionine cleavage. Mol Cell Proteomics 5:2336–2349.
- Link AJ, Robison K, Church GM (1997) Comparing the predicted and observed properties of proteins encoded in the genome of *Escherichia coli* K12. *Electrophoresis* 18:1259–1313.
- Lovell SC, Word JM, Richardson JS, Richardson DC (2000) The penultimate rotamer library. Proteins Struct Funct Genet 40:389–408.

- Liao YD, Jeng JC, Wang SC, Chang ST (2004) Removal of N-terminal methionine from recombinant proteins by engineered *E. coli* methionine aminopeptidase. *Protein Sci* 13:1802–1810.
- Bradshaw R, Brickey W, Walker K (1998) N-terminal processing: The methionine aminopeptidase and N-acetyl transferase families. *Trends Biochem Sci* 23:263– 267.
- 24. Lipscomb LA, et al. (1998) Context-dependent protein stabilization by methionine-toleucine substitution shown in T4 lysozyme. *Protein Sci* 7:765–773.
- Maurizi MR, Clark WP, Kim SH, Gottesman S (1990) Clp P represents a unique family of serine proteases. J Biol Chem 265:12546–12552.
- Kim YI, Burton RE, Burton BM, Sauer RT, Baker TA (2000) Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Mol Cell* 5:639– 648.
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected inoscillation mode. *Methods Enzymol* 276:307–326.
- Storoni LC, McCoy AJ, Read RJ (2004) Likelihood-enhanced fast rotation functions. Acta Crystallogr D 60:432–438.
- Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. Acta Crystallogr D 60:2126–2132.
- Adams PD, et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. Acta Crystallogr D 58:1948–1954.