Crystal structures of the DNA-binding domain of Escherichia coli proline utilization A flavoprotein and analysis of the role of Lys9 in DNA recognition

John D. Larson, Jermaine L. Jenkins, Jonathan P. Schuermann, Yuzhen Zhou, Donald F. Becker and John J. Tanner

Protein Sci. published online Sep 25, 2006;
Access the most recent version at doi:10.1110/ps.062425706

Crystal structures of the DNA-binding domain of *Escherichia coli* proline utilization A flavoprotein and analysis of the role of Lys9 in DNA recognition

JOHN D. LARSON,[1] JERMAINE L. JENKINS,[2] JONATHAN P. SCHUERMANN,[1] YUZHEN ZHOU,[3] DONALD F. BECKER,[3] AND JOHN J. TANNER[1,2]

[1]Department of Chemistry, University of Missouri—Columbia, Columbia, Missouri 65211, USA
[2]Department of Biochemistry, University of Missouri—Columbia, Columbia, Missouri 65211, USA
[3]Department of Biochemistry, Redox Biology Center, University of Nebraska—Lincoln, Lincoln, Nebraska 68588, USA

(RECEIVED July 3, 2006; FINAL REVISION August 3, 2006; ACCEPTED August 8, 2006)

Abstract

PutA (proline utilization A) from *Escherichia coli* is a 1320-amino-acid residue protein that is both a bifunctional proline catabolic enzyme and an autogenous transcriptional repressor. Here, we report the first crystal structure of a PutA DNA-binding domain along with functional analysis of a mutant PutA defective in DNA binding. Crystals were grown using a polypeptide corresponding to residues 1–52 of *E. coli* PutA (PutA52). The 2.1 Å resolution structure of PutA52 mutant Lys9Met was determined using Se-Met MAD phasing, and the structure of native PutA52 was solved at 1.9 Å resolution using molecular replacement. Residues 3–46 form a ribbon–helix–helix (RHH) substructure, thus establishing PutA as the largest protein to contain an RHH domain. The PutA RHH domain forms the intertwined dimer with tightly packed hydrophobic core that is characteristic of the RHH family. The structures were used to examine the three-dimensional context of residues conserved in PutA RHH domains. Homology modeling suggests that Lys9 and Thr5 contact DNA bases through the major groove, while Arg15, Thr28, and His30 may interact with the phosphate backbone. Lys9 is shown to be essential for specific recognition of *put* control DNA using gel shift analysis of the Lys9Met mutant of full-length PutA. Lys9 is disordered in the PutA52 structure, which implies an induced-fit binding mechanism in which the side chain of Lys9 becomes ordered through interaction with DNA. These results provide new insights into the structural basis of DNA recognition by PutA and reveal three-dimensional structural details of the PutA dimer interface.

Keywords: Proline utilization A; multifunctional PutA; ribbon–helix–helix DNA-binding domain; transcriptional repressor; X-ray crystal structure

Reprint requests to: John J. Tanner, Department of Chemistry, University of Missouri—Columbia, Columbia, MO 65211, USA; e-mail: tannerjj@missouri.edu; fax: (573) 882-2754.

Abbreviations: P5C, Δ^1-pyrroline-5-carboxylate; P5CDH, Δ^1-pyrroline-5-carboxylate dehydrogenase; PDB, Protein Data Bank; PRODH, proline dehydrogenase; PutA, proline utilization A; PutAK9M, Lys9Met mutant of *E. coli* PutA; PutA52, polypeptide corresponding to *E. coli* PutA residues 1–52; PutA52K9M, Lys9Met mutant of PutA52; RHH, ribbon–helix–helix; RMSD, root mean square difference.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.062425706.

Proline utilization A (PutA) is a large, membrane-associated bifunctional enzyme that catalyzes the sequential two-step oxidation of proline to glutamate (Menzel and Roth 1981a; Brown and Wood 1993; Surber and Maloy 1998; Becker and Thomas 2001; Vinod et al. 2002; Zhu and Becker 2003). In the first step, proline is oxidized to Δ^1-pyrroline-5-carboxylate (P5C) by the FAD-dependent PutA proline dehydrogenase (PRODH) domain. P5C is hydrolyzed nonenzymatically to glutamic semialdehyde, and the semialdehyde is oxidized

to glutamate by the NAD^+-dependent PutA P5C dehydrogenase (P5CDH) domain. PutAs typically contain 1000–1300 amino acid residues, with the PRODH domain located in the N-terminal half of the polypeptide chain and the P5CDH domain located in the C-terminal half. PutAs are exclusively found in prokaryotes, whereas PRODH and P5CDH are separate enzymes encoded by distinct genes in eukaryotes (Phang 1985).

In addition to their dual catalytic functions, some PutAs regulate transcription of the proline utilization genes, *putA* and *putP* (encodes the high affinity Na^+-proline transporter) (Menzel and Roth 1981b; Wood 1981; Maloy and Roth 1983; Ostrovsky De Spicer et al. 1991; Brown and Wood 1992; Ostrovsky De Spicer and Maloy 1993). Examples include PutA from *Escherichia coli*, *Salmonella typhimurium*, and *Pseudomonas putida*. These ''trifunctional'' PutAs repress expression of the *put* genes, which are transcribed in opposite directions, by binding to the *put* intergenic DNA region. PutA switches from a transcriptional repressor to a membrane bound enzyme by translocating from the cytoplasm to a peripheral position on the membrane, where proline is efficiently converted to glutamate. Reduction of the PutA FAD cofactor by the substrate proline triggers PutA to associate with the membrane, thus disrupting PutA–DNA association and activating transcription of the *put* genes (Brown and Wood 1993; Ostrovsky De Spicer and Maloy 1993; Zhu and Becker 2003; Zhang et al. 2004b).

E. coli PutA is a polypeptide of 1320 amino acids and purifies as a dimer with a molecular mass of ~293 kDa (Brown and Wood 1992). We previously determined crystal structures of the PRODH domain bound to various inhibitors using a truncated protein corresponding to residues 86–669 of *E. coli* PutA (Lee et al. 2003; Zhang et al. 2004a). The structures showed that the core of the PRODH domain is a unique $\beta_8\alpha_8$ barrel (residues 263–561), and prior to the work described below, the PRODH domain was the only PutA domain to be structurally characterized.

We previously showed that the DNA-binding domain of *E. coli* PutA is contained in residues 1–47 (Gu et al. 2004). A polypeptide corresponding to PutA residues 1–47 (PutA1–47) exhibited in vivo transcriptional repressor activity and bound specifically to *put* intergenic DNA in vitro with $K_d = 15$ nM. Furthermore, we proposed that PutA belongs to the ribbon–helix–helix (RHH) family of transcription factors based on amino acid sequence analysis (Gu et al. 2004).

As a follow-up to these studies, we report the first crystal structures of a PutA DNA-binding domain. Crystals of this domain were obtained from a polypeptide corresponding to *E. coli* PutA residues 1–52 (PutA52). The PutA52 structures show unequivocally that PutA belongs to the RHH family, thus establishing PutA as the largest protein to contain an RHH domain. The PutA RHH domain forms the tight, intertwined dimer that is characteristic of the RHH family. Based on analysis of the structure and comparison to other RHH proteins, we propose several residues that may be important for dimerization and DNA recognition. We tested one of these predictions by mutating Lys9 to Met in full-length PutA, which abrogated binding to *put* intergenic DNA. The PutA52 structures provide key information for understanding the three-dimensional structural basis for DNA recognition and dimerization.

Results

Overall fold

Two crystal structures of the *E. coli* PutA DNA-binding domain are reported here (Table 1). First, the structure of PutA52 mutant Lys9Met (PutA52K9M) was solved at 2.1 Å resolution using Se-Met multiwavelength anomalous diffraction (MAD) phasing. Crystals of Se-Met PutA52K9M grew in space group C2 with six protein chains per asymmetric unit. Second, the structure of native PutA52 was determined to 1.9 Å resolution using a tetragonal crystal form that has four chains per asymmetric unit. This structure was solved by molecular replacement using knowledge of the PutA52K9M structure.

PutA52K9M was used for initial structure determination by MAD phasing because of the higher Met content compared to native PutA52. The density-modified, MAD phasing electron density map for PutA52K9M exhibited very high quality (Fig. 1A) and allowed unambiguous modeling of residues 3–44 of all six chains (labeled A–F) in the asymmetric unit. After a few rounds of model building and refinement at 2.1 Å resolution, $2F_o$-F_c maps revealed the locations of residue 2 of chains C and E; residue 45 of chains A, C, E, and F; and residues 45–47 of chain D. The remaining residues of PutA52K9M are presumed to be disordered.

The 1.9 Å resolution structure of PutA52 was solved using molecular replacement phasing with a PutA52K9M monomer serving as the search model. Initial electron density maps resulting from molecular replacement calculations allowed modeling of residues 2–47 in all four chains (labeled A–D) of the asymmetric unit. After several rounds of model building and refinement, chain A was extended to include residue 48, and chain D was extended to residue 49. Alternative side chain conformations were modeled for Ser39 and Val8 (C and D chains).

Each PutA52K9M and PutA52 chain exhibits the RHH fold (Fig. 1), consisting of a β-strand formed by residues 3–11 (β1) followed by two α-helices formed by residues 12–25 (αA) and 29–46 (αB). The six PutA52K9M chains have pairwise root mean square differences (RMSDs) of

Table 1. *Data collection and refinement statistics*

	Se-Met PutA52K9M			
	Remote	Inflection	Peak	PutA52
Space group	C2	C2	C2	$P4_12_12$
Unit cell parameters (lengths in Å, angles in degrees)	$a = 72.1$ $b = 91.5$ $c = 69.6$ $\beta = 119.2$	$a = 72.1$ $b = 91.5$ $c = 69.6$ $\beta = 119.2$	$a = 72.1$ $b = 91.5$ $c = 69.6$ $\beta = 119.2$	$a = 55.7$ $c = 125.0$
Wavelength (Å)	0.987121	0.979553	0.979144	1.2398
Diffraction resolution (Å)	36.5–2.1 (2.17–2.10)	36.5–2.1 (2.17–2.10)	36.5–2.1 (2.17–2.10)	50.9–1.9 (1.97–1.90)
No. of observations	171,983	172,046	172,042	211,825
No. of unique reflections	23,035	23,020	23,013	16,253
Redundancy	7.5 (7.4)	7.5 (7.5)	7.5 (7.5)	13.0 (8.3)
Completeness (%)	99.9 (99.9)	99.9 (99.9)	99.9 (99.9)	99.9 (98.0)
R_{merge}	0.045 (0.335)	0.051 (0.367)	0.075 (0.398)	0.043 (0.317)
Average I/σ	21.4 (5.6)	18.4 (5.0)	13.8 (4.6)	25.5 (5.7)
Wilson B-factor (Å^2)	41	41	40	32
No. of protein chains	6			4
No. of protein residues	261			183
No. of protein atoms	2072			1493
No. of water molecules	67			68
R_{cryst}	0.204 (0.234)			0.203 (0.257)
R_{free}[a]	0.242 (0.295)			0.249 (0.364)
RMSD[b]				
Bond lengths (Å)	0.011			0.009
Bond angles (°)	1.1			1.09
Ramachandran plot[c]				
Favored (%)	97.5			99.4
Allowed (%)	2.5			0.6
Average B-factors (Å^2)				
Protein	41			36
Water	38			41
PDB accession code	2AY0			2GPE

Values for the outer resolution shell of data are given in parentheses.
[a]5% R_{free} test set.
[b]Compared to the Engh and Huber (1991) force field.
[c]The Ramachandran plot was generated with PROCHECK (Laskowski et al. 1993).

0.22–0.49 Å for C_α atoms. The four PutA52 chains have pairwise RMSDs of 0.22–0.53 Å for C_α atoms. RMSDs between PutA52 and PutA52K9M chains are in the range 0.39–0.69 Å. Visual inspection of all 10 superimposed chains verified that the conformations are very similar, even in the loop regions. These results indicate that PutA52 and PutA52K9M have nearly identical conformations.

In both crystal forms, the RHH subunits assemble into dimers. The asymmetric units were chosen such that there are three dimers in the C2 asymmetric unit and two dimers in the $P4_12_12$ asymmetric unit. Each of the five dimers features an intermolecular antiparallel two-stranded β-sheet (Fig. 1B), and there are eight to nine main chain hydrogen bonds between the two strands. The two subunits of the dimer intertwine such that αB of one subunit packs against all three secondary structural elements of the other subunit (Fig. 1B). As a result, both subunits contribute nonpolar side chains to the hydrophobic core of the dimer. The following residues are the largest contributors of intersubunit nonpolar buried surface area: Met6, Val8, Ile37, Tyr40, and Leu41. Each of these side chains bury 66–84 Å^2 of surface area, and all five residues are highly conserved in PutAs. The intertwined subunits, intermolecular β-sheet, and tightly packed hydrophobic intersubunit interface are commonly found in other RHH proteins, which shows that PutA exhibits the classic RHH dimer structure.

Conserved RHH sequence–structure relationships in PutA

Although pairwise amino acid sequence identities between members of different RHH subfamilies are only

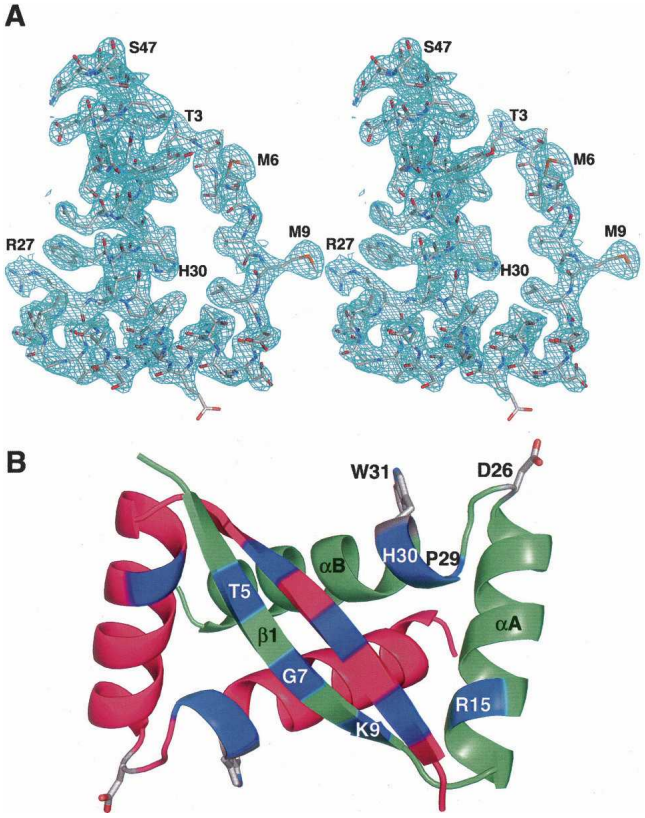

Figure 1. Structure of the PutA RHH domain. (*A*) Stereoscopic view of one PutA52K9M chain, covered by a figure of merit weighted experimental electron density map (1 σ). The map was calculated using amplitudes from the remote energy data set and density-modified MAD phases from RESOLVE. This map was used to build the initial model of PutA52K9M. The model shown is the final, refined structure of the D chain of PutA52K9M. (*B*) Ribbon drawing of the PutA RHH dimer. The two subunits are colored red and green. Secondary structure elements are labeled β1, αA, and αB for the green subunit. The blue patches and side chains indicate locations of residues mentioned in the text: Thr5, Gly7, Lys9 (Met9), Arg15, Asp26, Pro29, His30, and Trp31.

10%–15% (Chivers and Sauer 1999), there are several key sequence/structure relationships that are almost universally shared by RHH proteins. PutAs share some of these common features, which we now describe.

RHH domains usually have a basic residue (Arg/Lys) in β1 (Fig. 2A) that interacts directly with DNA bases through the major groove. RHH proteins have been classified according to whether this critical residue occurs at the beginning (type I) or end (type II) of β1 (Golovanov et al. 2003). PutA52 has only one basic residue in β1 (Lys9). Lys9 is located at the end of β1 (Fig. 1B), which shows that PutA has a type II RHH fold. For reference, the Arc repressor is a type II RHH protein (Raumann et al. 1994; Schildbach et al. 1999).

RHH proteins also have hydrophobic residues conserved at specific positions in β1, αA, and αB (Gomis-Ruth et al. 1998), as indicated in Figure 2A (see residues marked by ovals below the alignments). All six positions are occupied by hydrophobic residues in PutA (Met6, Val8, Leu10, Ile18, Ile33, Ile37). The side chains of these residues are directed toward the interior of the protein, where they contact other hydrophobic side chains, thus forming a tightly packed core. Note that these residues are highly conserved among PutAs (Fig. 2B), which suggests that all PutA RHH domains have the tightly packed hydrophobic core that is widely conserved in the RHH family.

RHH proteins typically have a small, usually polar, residue at the N_{cap} position of αB (Fig. 2A). The role of this residue in the Arc repressor (Ser32) has been extensively studied using mutagenesis (Anderson and Sauer 2003). Only small, uncharged residues at this position (Ser, Cys, Ala, Thr, Pro, Asn) preserve both stability and strong operator binding, which suggests that such side chains strike a balance between the thermodynamic requirements of protein stability and DNA recognition (Anderson and Sauer 2003). As in Arc, the αB N_{cap} of PutA is a small polar residue (Thr28) (Fig. 2). The hydroxyl of Thr28 forms a hydrogen bond with the backbone amine of Trp31, which is analogous to the hydrogen bond between Ser32 and the amine of Ser35 in Arc. In fact, the hydroxyl of PutA52 Thr28 perfectly overlaps the hydroxyl of Arc Ser32 after superposition of PutA52 and Arc (Fig. 3). This type of N_{cap} interaction is also observed in NikR (Asn27) (Schreiter et al. 2003), *Streptococcus pyogenes* omega protein (Asn50) (Murayama et al. 2001), and CopG (Ser27) (Gomis-Ruth et al. 1998), among others. Thr28 is identically conserved in PutAs (Fig. 2B), and thus the corresponding N_{cap} hydrogen bond is probably conserved as well.

Sequence–structure relationships unique to the PutA RHH subfamily

Amino acid sequence alignments suggest that PutA RHH domain sequences have several conserved residues that are not typically found in other RHH proteins: Gly7, Arg15, Asp26, Pro29, and Trp31 (see Fig. 2A, residues indicated by triangles). These PutA-specific residues could have important roles that enable the RHH domain to function within the unique context of the PutA polypeptide chain. Accordingly, these residues may be considered as defining features of the PutA RHH subfamily. We analyzed the PutA52 structures to gain insights into the roles that these residues might play in PutA structure and function.

PutA RHH domains have Gly as the middle residue (Gly7) in the important triad of residues in β1 that typically forms hydrogen bonds to DNA bases (Figs. 1B, 2). Gly is an unusual choice for this position, since it does not have a side chain capable of forming hydrogen bonds. The β-sheet structure of PutA52 is quite similar to

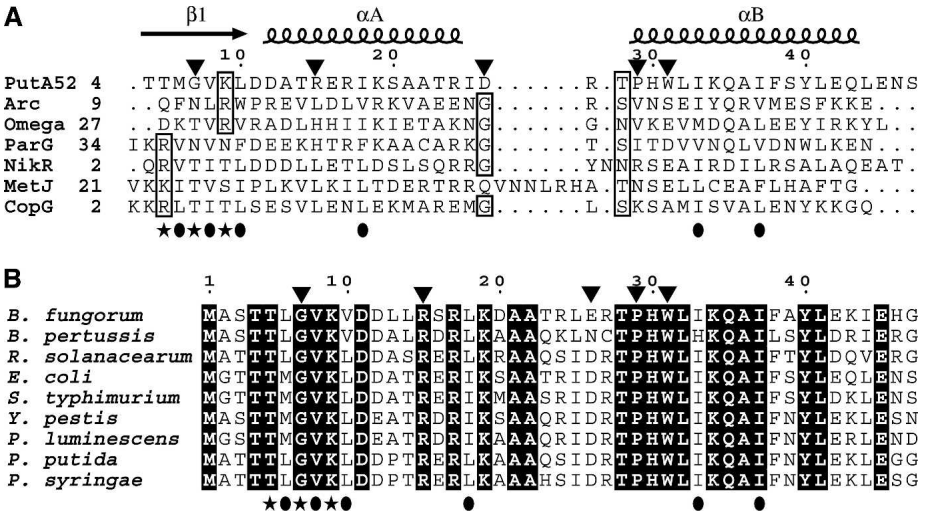

Figure 2. Amino acid sequence alignments of RHH and PutA proteins. (*A*) Structure-based sequence alignment of the *E. coli* PutA DNA-binding domain (PutA52) with other RHH domains. The numbers *above* the alignment correspond to *E. coli* PutA. The numbers on the *left* indicate the starting residue for each protein sequence. Boxes indicate conserved basic residues of β1, conserved glycine in the αA–αB turn, and the N_{cap} of αB. Stars *below* the alignment denote residues of β1 that typically interact with DNA. Ovals *below* the alignment denote conserved hydrophobic residues of the RHH family. Triangles *above* the alignment indicate residues unique to PutA RHH domain sequences. PDB accession codes for the sequences shown in this alignment are as follows: Arc (1BDT), Omega (1IRQ), ParG (1P94), NikR (1Q5V), MetJ (1CMC), and CopG (2CPG). (*B*) Sequence alignment of nine PutA RHH domains. Residues identically conserved in all nine sequences are indicated in black. Stars *below* the alignment denote residues of β1 that typically interact with DNA. Ovals *below* the alignment denote conserved hydrophobic residues of the RHH family. Triangles *above* the alignment indicate residues unique to PutA RHH domain sequences.

that of other RHH domains in terms of twist and interstrand hydrogen bonding. Thus, although glycine is somewhat unique at this position of the RHH fold, its presence does not disrupt the canonical RHH β-sheet conformation. We note that glycine at this position is not restricted to PutAs, as some NikR sequences also contain Gly at the analogous location (Chivers and Tahirov 2005). More typically, NikR contains Ser or Thr at this position (Fig. 2A).

Arg15 is conserved in PutAs, whereas hydrophobic residues typically occupy this position in other RHH domains (Fig. 2). In Arc, for example, the analogous residue is Leu19, which packs against the β-strand of the opposite subunit and contributes to the hydrophobic core. Arg15 of PutA adopts different conformations in the two crystal forms. In native PutA52, the density for Arg15 is rather weak but indicates that the side chain points out into solvent. In Se-Met PutA52K9M, Arg15 is more ordered and forms an intersubunit hydrogen bond with the side chain of Thr4. We note that intersubunit hydrogen bonding to this position of the β-sheet is not unprecedented in RHH domains. For example, NikR Gln2 forms analogous intersubunit hydrogen bonds to Asp10 and Arg28. As discussed in the next section, Arg15 may play a role in binding the DNA backbone.

PutA RHH domains have unique residues located in the αA–αB turn (Asp26) and N terminus of αB (Pro29). This turn is critical for orienting αB so that it packs against the opposite subunit. Non-PutA RHH domains almost always have a glycine in place of PutA Asp26 (Fig. 2A). There is considerable sequence variation within the RHH family at the N terminus of αB (Fig. 2A), but proline at this position seems to be unique to PutAs. Given these two unusual sequence features, it is somewhat surprising that the conformation of the αA–αB turn in PutA52 is strikingly similar to

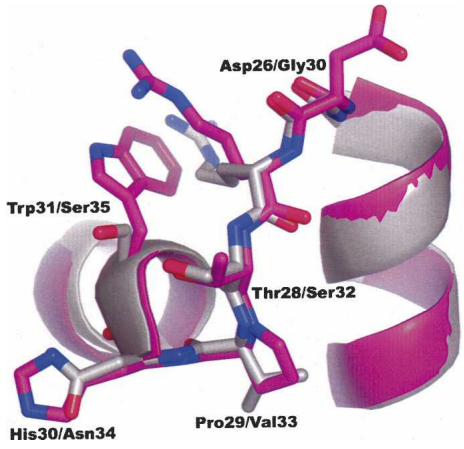

Figure 3. Comparison of the αA–αB loops of PutA52 and Arc. PutA52 is magenta, and Arc is white. Residues are labeled as PutA52/Arc. The orientation of this figure is similar to that of the green subunit in Figure 1B.

that of Arc (Fig. 3), which has Gly and Val at these two positions. Note that the main chain of Asp26 overlaps nearly perfectly with Arc Gly30, while the carboxyl of Asp26 is directed toward the solvent making no interactions with the protein (Fig. 3). Likewise, the main chain of PutA Pro29 superimposes perfectly that of Arc Val33 (Fig. 3). Importantly, the carbonyl of Pro29 forms a hydrogen bond to the amine of Ile33 (3.0 Å), showing that Pro29 does not disrupt the *i* to (*i* + 4) hydrogen bonding of αB. Thus, although PutA residues Asp26 and Pro29 are rather unusual in the RHH family, they cause no deviations from the classic RHH conformation in the αA–αB linker region.

Finally, PutA RHH sequences are unique in that Trp appears at position 31. This residue is typically polar or charged in other RHH proteins (Fig. 2A). Trp31 is located on a protein face opposite from the β-sheet, and its side chain points out into solvent (Figs. 1B, 3). The indole of Trp31 forms π-stacking interactions with a solvent imidazole molecule in the tetragonal crystal form. In the monoclinic crystal form, the indole N–H of Trp31 interacts with a chloride ion. Thus, Trp31 appears to have a high propensity for forming noncovalent interactions with moieties outside of the RHH domain. The analogous residue in Arc, Ser35, forms a water-mediated hydrogen bond to the DNA backbone (Raumann et al. 1994), which suggests that Trp31 could play a similar role in PutA.

Residues involved in binding DNA

Crystal structures of Arc (Raumann et al. 1994), MetJ (Somers and Phillips 1992), and CopG (Gomis-Ruth et al. 1998) complexed with DNA show that the β-sheet mediates the majority of protein–DNA interactions. The RHH β-sheet inserts into the major groove of DNA and typically two to three residues from each strand form hydrogen bonds with DNA bases and, in some cases, the phosphate backbone. These important residues are Gln9, Asn11, Arg13 in Arc; Lys23, Thr25, Ser27 in MetJ; and Arg4, Thr6, Thr8 in CopG (Fig. 2A). The analogous PutA residues are Thr5, Gly7, and Lys9. The side chains of Thr5 and Lys9 are directed toward the solvent in our structure, which is consistent with their putative role in binding DNA. The role of Gly7 is unclear at this time.

To test the importance of Lys9 for DNA recognition, we engineered and characterized the Lys9Met mutant of full-length *E. coli* PutA (PutAK9M). The biochemical properties of PutAK9M, such as enzyme activity, FAD absorbance spectrum, and dimeric structure, were consistent with wild-type PutA. However, the DNA-binding activity of PutAK9M was significantly disrupted. Gel mobility shift assays failed to show binding of PutAK9M to *put* control DNA (5 nM) even at 800-fold excess of PutAK9M to DNA (Fig. 4). Lack of DNA binding is not due to incorrect folding of the mutant because our crystallographic work shows that the structures of PutA52 and PutA52K9M are nearly identical. These results clearly show that Lys9 is essential for DNA binding in PutA. Lys9 is highly solvent exposed, which is consistent with it playing a direct role in binding DNA. Interestingly, the electron density for Lys9 in PutA52 was rather weak, and this side chain is disordered. Thus, binding of DNA to PutA likely induces structural ordering of Lys9.

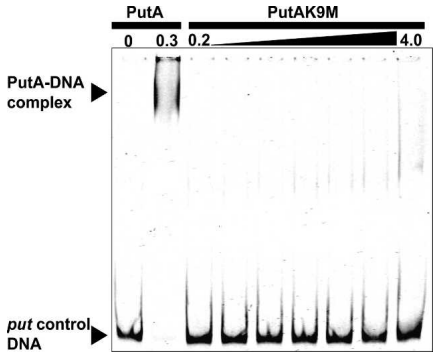

Figure 4. Gel mobility shift assay of wild-type PutA and PutA mutant K9M. Wild-type PutA (0.3 μM dimer) and PutAK9M (0.2–4 μM dimer) were incubated with IRdye-700 labeled *put* control DNA (5 nM) for 20 min at 20°C in 50 mM Tris (pH 8.0) containing 10% glycerol and 100 μg/mL nonspecific calf thymus DNA. Protein–DNA complexes were separated using a nondenaturing polyacrylamide gel (4%).

Our homology models of PutA52/DNA show the ammonium group of Lys9 within 3.8 Å of polar atoms of DNA bases (Fig. 5A). In addition, the model based on the CopG structure has the hydroxyl of Thr5 within 3.0 Å of polar atoms of DNA bases. Thus, we predict that Thr5 and Lys9, which are identically conserved in PutAs, likely form direct hydrogen bonds with DNA bases. The models are therefore consistent with gel-shift data for PutAK9M.

Homology models of PutA52 complexed with DNA also suggest that Arg15, Thr28, and His30 could interact with the phosphate backbone (Fig. 5A). All of our models have the side chain of His30 within 3.5 Å of a DNA phosphate group, while models based on Arc and MetJ show the Arg15 guanidinium within 3.4 Å of a phosphate group and the models based on Arc have the side chain of Thr28 3.3 Å from a phosphate group (Fig. 5A).

Electrostatic effects likely play an important role in orienting the PutA52 dimer in the major groove of DNA. The DNA-binding surface of PutA52 displays an overall positive electrostatic potential with high positive potential concentrated near Lys9, Arg15, and His30 (Fig. 5B). As a result, the β-sheet and flanking residues (Arg15, His30) form a ridge with high positive potential, which is predicted to bind the major groove of DNA (Fig. 5B). There is also a patch of high negative electrostatic potential on the

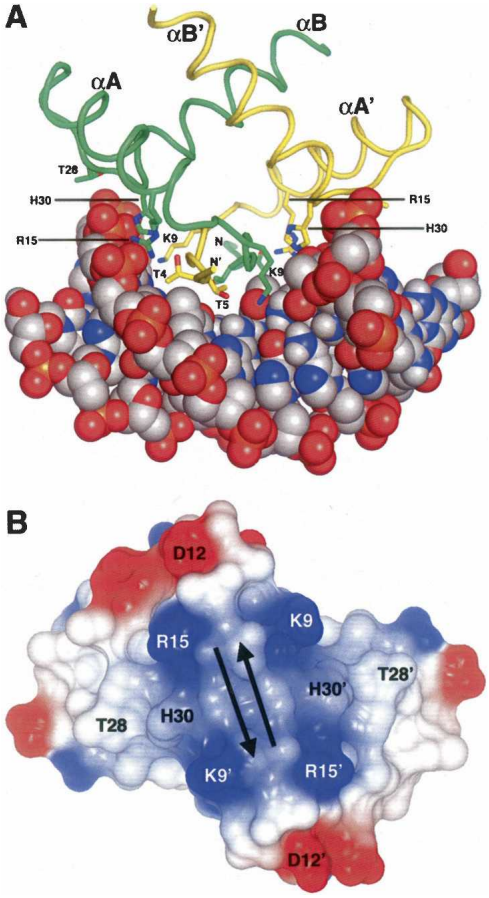

Figure 5. Model of the PutA RHH domain interacting with DNA. (*A*) Subunits of the PutA52 dimer are shown in green and yellow. DNA is shown in CPK mode. Helices of PutA52 are labeled αA, αB for the green subunit and αA′, αB′ for the yellow subunit. Note that the β-sheet is inserted into the major groove of DNA. N termini of the green and yellow subunits are labeled N and N′, respectively. Side chains are drawn for Thr4, Thr5, Lys9, Arg15, Thr28 (N_{cap} of αB), and His30. This model is based on homology with the Arc repressor as described in Materials and Methods. (*B*) Electrostatic potential of the DNA-binding surface of the PutA52 dimer. The potential ranges from −0.25 V (red) to +0.25 V (blue). The orientation of this panel is related to that of the panel *A* by a rotation of 90° around the horizontal axis, such that the DNA-binding surface faces the viewer. The arrows indicate the location of the β-sheet.

periphery of the binding surface near Asp12, which presumably helps steer the major groove onto the ridge by repelling the phosphate backbone (Fig. 5B).

Discussion

Relationship of PutA to other RHH DNA-binding proteins

Other members of the RHH family that have been structurally characterized include MetJ (Somers and Phillips 1992), Arc (Raumann et al. 1994), CopG (Gomis-Ruth et al. 1998), NikR (Schreiter et al. 2003), ParG (Golovanov et al. 2003), and more recently, the HP0222 protein from *Helicobacter pylori* (Popescu et al. 2005). These proteins are much smaller than PutA, ranging from the 45-residue CopG repressor to the 133-residue NikR protein from *E. coli* with variations due to additional N-terminal or C-terminal domains. The C-terminal domains of different RHH members have regulatory functions that enhance DNA binding. In MetJ, binding of the corepressor *S*-adenosylmethionine to the C-terminal domain regulates DNA binding and repression of the *met* operon by MetJ. NikR is responsive to nickel availability via a C-terminal nickel-binding domain. In response to increasing nickel concentrations, NikR binds to promoter regions of the *nik* operon and represses expression of the nickel ABC transporter. Activation of NikR DNA binding has been proposed to involve nickel-dependent conformational changes that are propagated from the nickel regulatory domain to the RHH domain (Schreiter et al. 2003). PutA adds a unique sensory function to the RHH family with an appendage of a flavin redox regulatory domain to the RHH fold.

Comparison of the PutA52 structure to other RHH proteins shows that PutA shares many of the sequence–structure relationships that are widely found in RHH domains. For example, the PutA RHH domain forms an intertwined dimer mediated by a two-stranded antiparallel intermolecular β-sheet, PutA has a basic residue (Lys9) in β1 that is predicted to interact with DNA bases, PutA displays the conserved pattern of hydrophobic residues that pack in the dimer interface, and there is a small polar residue (Thr28) at the αB N_{cap} position of PutA.

The PutA subfamily is distinguished from other RHH proteins at the primary sequence level by Gly7, Arg15, Asp26, Pro29, and Trp31. The PutA52 structures showed that these residues do not perturb the overall RHH fold and dimer structure.

In particular, replacement of the conserved glycine in the αA–αB turn by Asp26 and introduction of a proline (Pro29) at the N terminus of αB result in a conformation for the αA–αB turn that is almost identical to that of the Arc repressor (Fig. 3). Accommodation of Asp26 in the RHH fold can be understood by considering the Ramachandran plot probabilities of the 20 amino acids. The conserved Gly of RHH proteins occupies the left-handed α-helix region of the Ramachandran plot ($\phi = \psi = 260°$), and its unique backbone conformation is presumably essential for establishing the orientation of αB relative to that of αA. Similarly, the left-handed α-helix region of the Ramachandran plot is allowed for Asp (Hovmoller et al. 2002), and Asp26 of PutA52 indeed occupies this region. Interestingly, only Gly and Asn are found more frequently than Asp in this region of the Ramachandran plot (Hovmoller et al. 2002). It has been noted that the complexity of the Ramachandran plot for Asp is due to its role as a helix terminator (Hovmoller et al. 2002), which is a role that Asp26 plays in PutA (Fig. 3). Considering

Pro29 as the first residue of αB, although proline is known as a helix breaker, it has, in fact, a high propensity for occurring at the α-helix N terminus, as first discovered by Richardson and Richardson in 1988 (Richardson and Richardson 1988) and later described in more detail by Kim and Kang (1999). Thus, whereas PutA residues Asp26 and Pro29 are unusual at the primary sequence level compared to other RHH proteins, these residues are well suited for the structural roles that they play in the RHH fold.

Gly7 of PutA is interesting because RHH proteins typically have polar side chains at this position that interact with DNA bases. In the absence of a PutA/DNA crystal structure, it is difficult to predict a role for Gly7 in DNA recognition. Lack of a side chain at this position may allow more intimate contact between DNA and PutA. Alternatively, it is possible that Gly7 allows increased flexibility of the β-sheet, which might be important for induced-fit binding to DNA and/or the conformational changes that underlie the mechanism by which PutA switches between enzymatic and regulatory roles in response to cofactor reduction.

Arg15 and Trp31 are also unique to PutA RHH sequences. The PutA52 structures show that these side chains are solvent-exposed and therefore poised to interact with moieties outside the RHH domain. Homology modeling implicates Arg15 in binding the DNA backbone (Fig. 5A). By analogy to the Arc repressor structure, Trp31 may participate in DNA recognition. Since the indole N atom of Trp31 does not superimpose well with the hydroxyl of Arc Ser35 (Fig. 3), adjustments in the protein structure may be required for Trp31 to interact with DNA. An alternative hypothesis is that Trp31 interacts with another domain of PutA, which would be consistent with Trp31 being conserved in PutAs but absent in other RHH domains. The observation that Trp31 forms π-stacking interactions with solvent imidazole suggests that Trp31 may form analogous interactions with residues in another domain of PutA. The work presented here provides a foundation for mutagenesis studies that will be done to gain new insights into how PutA-specific residues, such as Gly7, Arg15, and Trp31, might be involved in proline-linked transcriptional regulation of proline utilization genes.

Structural basis of DNA recognition

Our models of the PutA RHH domain complexed to DNA suggest that the RHH domain spans 5–7 base pairs (bp) via major groove binding (Fig. 5A). This result is qualitatively consistent with the discovery of several PutA-binding sites containing the 6-bp sequence element GTTGCA in the *put* control DNA (Zhang et al. 2004b). Synthetic oligonucleotides of 21 bp that contain GTTGCA bind specifically to PutA with $K_d \leq 200$ nM (Zhang et al. 2004b), which is only fourfold higher than the K_d for PutA binding to the entire *put* control DNA ($K_d = 45$ nM) (Becker and Thomas 2001).

We showed here that mutation of Lys9 to Met eliminates binding of PutA to *put* control DNA, and our modeling studies predict that Lys9 and Thr5 contact DNA bases directly. The models also predict roles for Arg15, Thr28, and His30 in binding the DNA backbone (Fig. 5A). Arg15 was discussed in the previous section as being one of the PutA-specific residues. Thr28 is the N_{cap} of αB. This residue is highly conserved among PutAs (Fig. 2B), and it is structurally analogous to Arc Ser32 and MetJ Thr52. These residues form both direct (3.4 Å) and water-mediated electrostatic interactions with the phosphate backbone in the Arc and MetJ repressor/operator complexes (PDB codes 1BDT, 1CMA).

His30 is an interesting candidate for making a direct interaction with DNA because it forms intersubunit hydrogen-bonding interactions with β1. In the monoclinic crystal form (Se-Met PutA52K9M), His30 forms a hydrogen bond with the carbonyl of Gly7 of the opposite subunit. This interaction is structurally analogous to the intersubunit hydrogen bond in Arc between Asn34 and the carbonyl of Arg13. Since the side chain of Arc Arg13 forms hydrogen bonds to DNA bases, Asn34 of Arc plays an indirect role in DNA recognition. It remains to be seen whether His30 of PutA plays a direct or indirect role in binding DNA.

Structural basis of PutA dimerization

PutAs studied to date purify as apparent homodimers, whether they are trifunctional repressor/enzymes like *E. coli* PutA (Becker and Thomas 2001) or bifunctional enzymes such as *Bradyrhizobium japonicum* PutA (Krishnan and Becker 2005). A truncated form of *E. coli* PutA that starts at residue 86 (PutAΔ85) purifies as an apparent monomer, which suggests that the N-terminal RHH domain is essential for dimerization of full-length *E. coli* PutA (Gu et al. 2004). We find that the PutA DNA-binding domain forms a classic RHH dimer, and we assume that a similar dimeric substructure is present in full-length *E. coli* PutA. The structure reported here provides a basis for probing the importance of individual RHH residues in stabilizing the *E. coli* PutA dimer.

PutA contains PRODH and P5CDH domains in addition to the RHH domain studied here. The structure of *E. coli* residues 86–669 is a $\beta_8\alpha_8$ barrel, and it forms an apparent monomer in solution. PutA constructs containing only the P5CDH domain have not yet been successfully studied because of solubility problems (J.J. Tanner and D.F. Becker, unpubl.). However, *Thermus thermophilus* P5CDH, as well as other aldehyde dehydrogenases such as sheep liver aldehyde dehydrogenase (PDB entry

1BXS) (Moore et al. 1998) and retinal dehydrogenase (PDB entry 1BI9) (Lamb and Newcomer 1999), form dimers mediated by intermolecular β-sheets. Thus, it is possible that *E. coli* PutA dimerization involves both RHH and P5CDH domains. In bifunctional PutAs, which lack the RHH domain, the dimeric structure may be stabilized entirely by the P5CDH domain. The questions of how the RHH domain is integrated into the PutA architecture and whether trifunctional PutAs contain a second dimerization domain await structure determination of a full-length trifunctional PutA.

Materials and methods

Materials

Unless stated otherwise, all chemicals and buffers were purchased from Fisher Scientific and Sigma-Aldrich, Inc. Restriction endonucleases and T4 DNA ligase were purchased from Fermentas and Promega, respectively. BCA and Coomassie reagents used for protein quantitation were obtained from Pierce. All experiments used 18 MΩ water. *E. coli* strains XL-Blue and BL21 (DE3) pLysS were purchased from Novagen. The *put* intergenic DNA (419 bp) used for DNA-binding assays was prepared as described previously using genomic DNA from *E. coli* strain JT31 (Gu et al. 2004).

Subcloning and site-directed mutagenesis of PutA and PutA52

The region of the *putA* gene that codes for the N-terminal 52 amino acids of *E. coli* PutA was PCR-amplified using the construct PutA-pET23b (Zhu et al. 2002) as a template. The T7 promoter primer and the primer 5′-GCGATACTCTGCTCGAGC TACCTGCGC-3′, which introduced an XhoI site, were used for the PCR reaction. The PCR-amplified *putA52* gene product was subsequently subcloned into pET23b using NdeI and XhoI to produce the plasmid PutA52-pET23b. The encoded protein includes residues 1–52 and a noncleavable C-terminal His_6 tag.

A PutA52 construct containing a cleavable N-terminal His_8 tag was generated by mutagenesis of the plasmid PutA-pKA8H. Using QuickChange, the codon for Leu53 was mutated to an ochre (UAA) stop codon using the following oligonucleotide and its complement: 5′-GCGATACTCTGCCGGAG TAACCTGCGCTGCTTTCTG-3′. The encoded protein includes a His_8 tag, followed by a tobacco etch virus protease (TEVP) cleavage site and PutA residues 1–52.

Plasmids coding for *E. coli* PutA mutant Lys9Met (PutAK9M) and PutA52 mutant Lys9Met (PutA52K9M) were prepared using the QuikChange (Stratagene) site-directed mutagenesis kit with PutA-pET23b and PutA52-pET23b serving as the templates. The following oligonucleotide and its complement were used for Lys9Met mutagenesis: 5′-ACCATGGGGGTTATGCTGGAC GACGCG-3′. The clones and site-directed mutants were confirmed by nucleic acid sequencing.

Preparation of PutA proteins for DNA-binding studies

E. coli PutA, PutAK9M, and PutA52 were expressed as C-terminal His_6-tag fusion proteins from pET23b in *E. coli* strain BL21(DE3) pLysS using protocols similar to those used previously for PutA proteins (Gu et al. 2004; Zhang et al. 2004a), except that expression of PutA52 was induced with IPTG at 25°C for 6 h. After purification, PutAK9M was dialyzed into 70 mM Tris (pH 8.1) containing 2 mM EDTA, whereas PutA52 was dialyzed into 50 mM potassium phosphate buffer (pH 7.5) containing 0.2 M NaCl. The C-terminal His_6 tags were retained after purification. The purity of each protein was >95% as judged by SDS-PAGE analysis. Protein concentration was determined using the BCA method with bovine serum albumin as the standard and spectrophotometrically using a molar extinction coefficient of 12,700 M^{-1} cm^{-1} ($\lambda = 451$ nm) for PutA and PutAK9M, and an extinction coefficient of 6970 M^{-1} cm^{-1} ($\lambda = 280$ nm) for PutA52 (Becker and Thomas 2001).

DNA-binding assays

Nondenaturing gel electrophoretic mobility shift assays were used to test the binding of C-terminally His-tagged PutA proteins (PutA, PutAK9M, PutA52) to *put* control intergenic DNA as described previously (Gu et al. 2004). Fluorescently labeled *put* intergenic DNA was used in the binding reactions. This reagent was prepared by PCR amplification of the *put* intergenic DNA using a synthetic oligonucleotide labeled at the 5′-end with IRdye-700 (LI-COR, Inc.). The PutA–DNA complexes were separated using a polyacrylamide (4%) nondenaturing gel at 4°C as described previously (Gu et al. 2004). Calf thymus competitor DNA (100 μg/mL) was added to the binding mixtures to prevent nonspecific PutA–DNA interactions. The gels were visualized using a LI-COR Odyssey Imager. Measurement of the dissociation constant for binding of PutA52 to *put* intergenic DNA gave a value of 10.3 ± 2.4 nM (data not shown), which is similar to previously determined values for PutA ($K_d = 45$ nM) and PutA47 ($K_d = 15$ nM) (Becker and Thomas 2001; Gu et al. 2004).

Preparation of selenomethionyl PutA52K9M

The selenomethionyl (Se-Met) derivative of PutA52K9M was expressed as a C-terminal His_6-tag fusion protein from pET23b in *E. coli* strain BL21(DE3)pLysS using the metabolic inhibition method (Doublie 1997). The cells were grown in 1.0 L of M9 media at 37°C and 250 rpm until the optical density ($\lambda = 600$ nm) reached 0.8. An amino acid mixture containing selenomethionine (Doublie 1997) was added to the cultures, and incubation continued at 37°C and 250 rpm for an additional 30 min. The cultures were induced by addition of 0.5 mM IPTG, and incubation was continued for 8 h at 22°C and 200 rpm.

Pelleted *E. coli* cells were resuspended on ice in binding buffer (20 mM Tris, 5 mM imidazole, 0.5 M NaCl at pH 7.9) supplemented with Benzonase nuclease (2.5 U/mL), 1 mM dithiothreitol, and five protease inhibitors (0.1 mM TPCK, 0.05 mM AEBSF, 0.1 μM Pepstatin, 0.01 mM Leupeptin, 5 μM E-64). The resuspended cells were disrupted with two passes through a French pressure cell at 110 MPa and placed on ice. The disrupted cells were centrifuged at 35,000*g* and 4°C for 45 min. The supernatant fraction was collected, placed on ice, and subsequently added to a Ni-NTA affinity column equilibrated with binding buffer. Contaminating proteins were removed from the column by successive washes with the binding buffer followed by the binding buffer supplemented with 20 mM imidazole. PutA52K9M was eluted with the binding buffer supplemented with 500 mM imidazole. The purified protein sample was dialyzed overnight against 2.0 L of 70 mM Tris

(pH 8.1) buffer containing 2 mM EDTA, 5 mM dithiothreitol, and 5% glycerol. The protein was concentrated to 18 mg/mL using a centrifugal concentrator. Protein concentration was determined using the Bradford method (Pierce Coomassie Plus). SDS-PAGE analysis using a Coomassie protein stain method demonstrated that the protein preparations were purified to >95% homogeneity. The C-terminal hexahistidine affinity tag was not removed. Mass spectral analysis of PutA52K9M and Se-Met PutA52K9M indicated molecular masses of 6662.9 Da and 6754.42 Da, respectively, which suggested an average incorporation of two Se per polypeptide chain. For reference, the predicted sequence of PutA52K9M has three Met residues (Met1, Met6, Met9).

Preparation of PutA52 for crystallization trials

PutA52 for crystallization trials was expressed as an N-terminal His_8-tag fusion protein in *E. coli* strain BL21(DE3)pLysS. Cells were grown in 4.0 L of LB media at 37°C with a rotation rate of 250 rpm until the optical density reached $A_{600} = 0.5$. The temperature and rotation rate were reduced to 25°C and 220 rpm, respectively, and IPTG was added to a final concentration of 0.5 mM. After 4 h of induction, the cells were harvested and disrupted as described above for PutA52K9M.

The protein was purified using Ni-affinity chromatography as described above for PutA52K9M. The His_8 tag was removed by incubating the purified protein with His-tagged TEVP for 2 h at 30°C followed by dialysis overnight into 50 mM Tris-HCl (pH 8.0) and 300 mM NaCl. The mixture was loaded onto the Ni-affinity column, and the flow-through was collected. Tag-free PutA52 was then dialyzed into the pre-crystallization buffer (20 mM Tris at pH 8.0, 125 mM NaCl, and 5 mM Imidazole) and concentrated using a centrifugal concentrator to 25 mg/mL. The protein concentration was measured using the BCA method.

Crystallization of PutA52 and Se-Met PutA52K9M

Crystallization trials for both proteins were performed at 22°C using the hanging-drop method of vapor diffusion with drops formed by mixing equal volumes of the reservoir and protein solutions. Commercially available crystal screen kits (Hampton Research and Decode Genetics) were used to identify initial crystallization conditions.

Rod-shaped crystals of PutA52 grew in Index condition 57 (50 mM ammonium sulfate, 50 mM Bis-Tris at pH 6.5, 35% pentaerythritol ethoxylate [15/4 EO/OH]). The mother liquor provided cryoprotection; thus, the crystals were simply picked up with Hampton mounting loops and plunged into liquid nitrogen. The crystals belong to the space group $P4_12_12$ and have unit cell dimensions of $a = 55.73$ Å, $c = 125.02$ Å. There are four PutA52 chains per asymmetric unit, with 39% solvent and a Matthews coefficient of 2.0 $Å^3$/Da.

A monoclinic crystal form of Se-Met PutA52K9M was obtained from Index reagent 7. After optimization, crystals were grown over reservoir solutions of 1.7–3.0 M NaCl, 0.1 M Na citrate (pH 3.0–5.5), and 5 mM dithiothreitol. In preparation for low-temperature data collection, the crystals were equilibrated in a solution of 3.2 M NaCl, 0.1 M Na citrate (pH 3.0), 5 mM dithiothreitol, and 30% glycerol. The crystals have space group C2 with unit cell dimensions of $a = 72.1$, $b = 91.5$, $c = 69.6$, $\beta = 119.2°$. There are six PutA52K9M chains per asymmetric unit, with 50% solvent and a Matthews coefficient of 2.5 $Å^3$/Da (Matthews 1968).

Structure determination of PutA52K9M

The structure of PutA52K9M was solved using Se-Met MAD phasing from data collected at the Advanced Light Source beamline 4.2.2. Following acquisition of an absorption spectrum from one crystal, three data sets were collected from a second crystal at wavelengths corresponding to the peak of absorption for Se, ascending inflection point, and a remote low energy. Each data set consisted of 360 frames collected with an oscillation angle of 1° per frame, an exposure time of 1 sec per frame, and a detector distance of 125 mm. The data frames were accumulated by iterating through the three collection energies in 2° wedges in order to minimize errors due to radiation damage. Integration and scaling were performed using d*trek (Pflugrath 1999). I^+ and I^- were treated as equivalent reflections during scaling. Data collection and processing statistics are listed in Table 1.

SOLVE (Terwilliger and Berendzen 1999) was used to determine the Se substructure, and RESOLVE (Terwilliger 2003) was used for density modification and automated electron density map interpretation. SOLVE found all 12 of the expected Se sites in the asymmetric unit, and the resulting MAD phases had a figure of merit = 0.48 for 20–2.2 Å data. After density modification, the figure of merit increased to 0.54 for all reflections. The backbone tracing from RESOLVE consisted of six chains having a total of 256 residues with the sequence assigned to 229 residues. The model was improved with several rounds of model building in COOT (Emsley and Cowtan 2004) followed by refinement with REFMAC5 (Winn et al. 2001) against the remote energy data set. Noncrystallographic symmetry restraints were not used during refinement. In addition to the six protein chains, the asymmetric unit contains six chloride ions and 67 water molecules. The chloride ions were modeled at the locations of strong spherical electron density features that appeared in F_o-F_c maps at levels of 11–16 σ. When water molecules were modeled at these sites, the F_o-F_c density was observed at levels of 6.5–8.0 σ. Each chloride ion binds between two PutA52K9M protomers and interacts with the guanidinium of Arg27, the indole N–H of Trp31, and the backbone amine of Met9. Refinement statistics for the final model are listed in Table 1. Coordinates and structure factors for Se-Met PutA52K9M have been deposited in the Protein Data Bank (PDB) (Berman et al. 2000) under accession code 2AY0.

Structure determination of PutA52

A data set consisting of 180 frames was collected at the ALS beamline 4.2.2 with an oscillation angle of 1° per frame, an exposure time of 5 sec per frame, and a detector distance of 110 mm. Integration and scaling calculations were performed with d*trek. Data collection and processing statistics are listed in Table 1.

The structure of PutA52 was solved with molecular replacement using a PutA52K9M monomer as a search model. Phaser (McCoy et al. 2005) was used for molecular replacement calculations using data in the range 51–2.5 Å resolution. The best solution was found for space group $P4_12_12$ as indicated by an absence of steric clashes, least likelihood gain value of 1133.63, and Z-score of 25. The model from molecular replacement was improved with rigid-body refinement followed by several alternating rounds of model building with COOT and restrained refinement with REFMAC5 (with TLS). Noncrystallographic symmetry restraints were not used during refinement. A strong electron density feature consistent with a small planar

molecule was present in all four chains near Gln35. The pre-crystallization buffer contained 5 mM imidazole, which led us to build one imidazole molecule into each location. Each imidazole molecule is sandwiched between the aromatic rings of Trp31 and Phe38 and forms a hydrogen bond with the side chain of Gln35. After refinement, the imidazole molecules had an average B-factor of 39 $Å^2$. The solvent model for PutA52 also includes 68 water molecules. Coordinates and structure factors for PutA52 have been deposited in the PDB under accession code 2GPE. See Table 1 for refinement statistics.

Modeling of PutA52/DNA complex

Qualitative models of PutA52 complexed with DNA were built using structures of Arc/DNA (PDB code 1BDT) (Raumann et al. 1994), MetJ/DNA (PDB code 1MJ2) (Garvie and Phillips 2000), and CopG/DNA (PDB code 1B01) (Gomis-Ruth et al. 1998) as templates. For each template, a PutA52 dimer was superimposed onto the template structure by fitting the backbone of PutA52 β1 residues 5–9 (both subunits) to the equivalent residues of the template using CNS (residues 9–13 of Arc, residues 23–27 of MetJ, residues 4–8 of CopG) (Brünger et al. 1998). Missing atoms of Lys9 were added with COOT. The resulting models of PutA52 were combined with the DNA coordinates from the respective template and input to energy minimization to relieve steric clashes. The minimization protocol consisted of 200 steps of conjugate gradient energy minimization in CNS using polar hydrogen atoms, a 13 Å nonbonded cutoff, and the continuum solvent model (dielectric constant = 80). We note that our models do not account for conformational changes induced in the protein or DNA upon binding, although induced-fit binding has been observed for RHH proteins. For example, binding of DNA to Arc induces conformational adjustments in the β-sheet (Raumann et al. 1994), and CopG causes significant bending of DNA (Gomis-Ruth et al. 1998).

Structure analysis

The programs CNS, CCP4i, and PyMOL (DeLano 2002) were used for structural analysis. Superimposition of structures and RMSD calculations were done using SSM (Krissinel and Henrick 2004) and CE (Shindyalov and Bourne 1998). The electrostatic potential of the PutA52 dimer was calculated from the linearized Poisson-Boltzmann equation (Davis and McCammon 1989) as implemented in CCP4mg (Potterton et al. 2004).

Acknowledgments

This work was supported by NIH Grants GM065546 (J.J.T.) and GM061068 (D.F.B.), NSF Grant MCB0340912 (D.F.B), and the Nebraska Agricultural Research Division, Journal Series No. 15001. We thank Jay Nix and Darren Sherrell of the ALS beamline 4.2.2 for help with data collection and processing. The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory.

References

Anderson, T.A. and Sauer, R.T. 2003. Role of an N(cap) residue in determining the stability and operator-binding affinity of Arc repressor. *Biophys. Chem.* **100:** 341–350.

Becker, D.F. and Thomas, E.A. 2001. Redox properties of the PutA protein from *Escherichia coli* and the influence of the flavin redox state on PutA–DNA interactions. *Biochemistry* **40:** 4714–4722.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. 2000. The Protein Data Bank. *Nucleic Acids Res.* **28:** 235–242.

Brown, E. and Wood, J.M. 1992. Redesigned purification yields a fully functional PutA protein dimer from *Escherichia coli*. *J. Biol. Chem.* **267:** 13086–13092.

———. 1993. Conformational change and membrane association of the PutA protein are coincident with reduction of its FAD cofactor by proline. *J. Biol. Chem.* **268:** 8972–8979.

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

Chivers, P.T. and Sauer, R.T. 1999. NikR is a ribbon–helix–helix DNA-binding protein. *Protein Sci.* **8:** 2494–2500.

Chivers, P.T. and Tahirov, T.H. 2005. Structure of *Pyrococcus horikoshii* NikR: Nickel sensing and implications for the regulation of DNA recognition. *J. Mol. Biol.* **348:** 597–607.

Davis, M.E. and McCammon, J.A. 1989. Solving the finite difference linearized Poisson-Boltzmann equation: A comparison of relaxation and conjugate gradient methods. *J. Comput. Chem.* **10:** 386–391.

DeLano, W.L. 2002. The PyMOL molecular graphics system. http://www.pymol.org.

Doublie, S. 1997. Preparation of selenomethionyl proteins for phase determinations. *Methods Enzymol.* **276:** 523–530.

Emsley, P. and Cowtan, K. 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60:** 2126–2132.

Engh, R.A. and Huber, R. 1991. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **47:** 392–400.

Garvie, C.W. and Phillips, S.E. 2000. Direct and indirect readout in mutant Met repressor–operator complexes. *Structure* **8:** 905–914.

Golovanov, A.P., Barilla, D., Golovanova, M., Hayes, F., and Lian, L.Y. 2003. ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon–helix–helix structure. *Mol. Microbiol.* **50:** 1141–1153.

Gomis-Ruth, F.X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R., Gonzalez, A., Espinosa, M., del Solar, G., and Coll, M. 1998. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J.* **17:** 7404–7415.

Gu, D., Zhou, Y., Kallhoff, V., Baban, B., Tanner, J.J., and Becker, D.F. 2004. Identification and characterization of the DNA-binding domain of the multifunctional PutA flavoenzyme. *J. Biol. Chem.* **279:** 31171–31176.

Hovmoller, S., Zhou, T., and Ohlson, T. 2002. Conformations of amino acids in proteins. *Acta Crystallogr. D Biol. Crystallogr.* **58:** 768–776.

Kim, M.K. and Kang, Y.K. 1999. Positional preference of proline in α-helices. *Protein Sci.* **8:** 1492–1499.

Krishnan, N. and Becker, D.F. 2005. Characterization of a bifunctional PutA homologue from *Bradyrhizobium japonicum* and identification of an active site residue that modulates proline reduction of the flavin adenine dinucleotide cofactor. *Biochemistry* **44:** 9130–9139.

Krissinel, E. and Henrick, K. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60:** 2256–2268.

Lamb, A.L. and Newcomer, M.E. 1999. The structure of retinal dehydrogenase type II at 2.7 Å resolution: Implications for retinal specificity. *Biochemistry* **38:** 6003–6011.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26:** 283–291.

Lee, Y.H., Nadaraia, S., Gu, D., Becker, D.F., and Tanner, J.J. 2003. Structure of the proline dehydrogenase domain of the multifunctional PutA flavoprotein. *Nat. Struct. Biol.* **10:** 109–114.

Maloy, S. and Roth, J.R. 1983. Regulation of proline utilization in *Salmonella typhimurium*: Characterization of *put*:Mu *L*(Ap, *lac*) operon fusions. *J. Bacteriol.* **154:** 561–568.

Matthews, B.W. 1968. Solvent content of protein crystals. *J. Mol. Biol.* **33:** 491–497.

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. 2005. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* **61:** 458–464.

Menzel, R. and Roth, J. 1981a. Enzymatic properties of the purified *putA* protein from *Salmonella typhimurium*. *J. Biol. Chem.* **256:** 9762–9766.

———. 1981b. Regulation of genes for proline utilization in *Salmonella typhimurium*: Autogenous repression by the putA gene product. *J. Mol. Biol.* **148:** 21–44.

Moore, S.A., Baker, H.M., Blythe, T.J., Kitson, K.E., Kitson, T.M., and Baker, E.N. 1998. Sheep liver cytosolic aldehyde dehydrogenase: The structure reveals the basis for the retinal specificity of class 1 aldehyde dehydrogenases. *Structure* **6:** 1541–1551.

Murayama, K., Orth, P., de la Hoz, A.B., Alonso, J.C., and Saenger, W. 2001. Crystal structure of omega transcriptional repressor encoded by *Streptococcus pyogenes* plasmid pSM19035 at 1.5 Å resolution. *J. Mol. Biol.* **314:** 789–796.

Ostrovsky De Spicer, P. and Maloy, S. 1993. PutA protein, a membrane-associated flavin dehydrogenase, acts as a redox-dependent transcriptional regulator. *Proc. Natl. Acad. Sci.* **90:** 4295–4298.

Ostrovsky De Spicer, P., O'Brian, K., and Maloy, S. 1991. Regulation of proline utilization in *Salmonella typhimurium*: A membrane-associated dehydrogenase binds DNA in vitro. *J. Bacteriol.* **173:** 211–219.

Pflugrath, J.W. 1999. The finer things in X-ray diffraction data collection. *Acta Crystallogr. D Biol. Crystallogr.* **55:** 1718–1725.

Phang, J.M. 1985. The regulatory functions of proline and pyrroline-5-carboxylic acid. *Curr. Top. Cell. Regul.* **25:** 92–132.

Popescu, A., Karpay, A., Israel, D.A., Peek Jr., R.M., and Krezel, A.M. 2005. *Helicobacter pylori* protein HP0222 belongs to Arc/MetJ family of transcriptional regulators. *Proteins* **59:** 303–311.

Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G.N., Cohen, S., Perrakis, A., and Noble, M. 2004. Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D Biol. Crystallogr.* **60:** 2288–2294.

Raumann, B.E., Rould, M.A., Pabo, C.O., and Sauer, R.T. 1994. DNA recognition by β-sheets in the Arc repressor-operator crystal structure. *Nature* **367:** 754–757.

Richardson, J.S. and Richardson, D.C. 1988. Amino acid preferences for specific locations at the ends of α helices. *Science* **240:** 1648 1652.

Schildbach, J.F., Karzai, A.W., Raumann, B.E., and Sauer, R.T. 1999. Origins of DNA-binding specificity: Role of protein contacts with the DNA backbone. *Proc. Natl. Acad. Sci.* **96:** 811–817.

Schreiter, E.R., Sintchak, M.D., Guo, Y., Chivers, P.T., Sauer, R.T., and Drennan, C.L. 2003. Crystal structure of the nickel-responsive transcription factor NikR. *Nat. Struct. Biol.* **10:** 794–799.

Shindyalov, I.N. and Bourne, P.E. 1998. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **11:** 739–747.

Somers, W.S. and Phillips, S.E. 1992. Crystal structure of the met repressor-operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. *Nature* **359:** 387–393.

Surber, M.W. and Maloy, S. 1998. The PutA protein of *Salmonella typhimurium* catalyzes the two steps of proline degradation via a leaky channel. *Arch. Biochem. Biophys.* **354:** 281–287.

Terwilliger, T.C. 2003. SOLVE and RESOLVE: Automated structure solution and density modification. *Methods Enzymol.* **374:** 22–37.

Terwilliger, T.C. and Berendzen, J. 1999. Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55:** 849–861.

Vinod, M.P., Bellur, P., and Becker, D.F. 2002. Electrochemical and functional characterization of the proline dehydrogenase domain of the PutA flavoprotein from *Escherichia coli*. *Biochemistry* **41:** 6525–6532.

Winn, M.D., Isupov, M.N., and Murshudov, G.N. 2001. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* **57:** 122–133.

Wood, J.M. 1981. Genetics of L-proline utilization in *Escherichia coli*. *J. Bacteriol.* **146:** 895–901.

Zhang, M., White, T.A., Schuermann, J.P., Baban, B.A., Becker, D.F., and Tanner, J.J. 2004a. Structures of the *Escherichia coli* PutA proline dehydrogenase domain in complex with competitive inhibitors. *Biochemistry* **43:** 12539–12548.

Zhang, W., Zhou, Y., and Becker, D.F. 2004b. Regulation of PutA-membrane associations by flavin adenine dinucleotide reduction. *Biochemistry* **43:** 13165–13174.

Zhu, W. and Becker, D.F. 2003. Flavin redox state triggers conformational changes in the PutA protein from *Escherichia coli*. *Biochemistry* **42:** 5469–5477.

Zhu, W., Gincherman, Y., Docherty, P., Spilling, C.D., and Becker, D.F. 2002. Effects of proline analog binding on the spectroscopic and redox properties of PutA. *Arch. Biochem. Biophys.* **408:** 131–136.