

## NOTES

# CheZ Phosphatase Localizes to Chemoreceptor Patches via CheA-Short

Brian J. Cantwell,<sup>1</sup> Roger R. Draheim,<sup>1</sup> Richard B. Weart,<sup>1†</sup> Cameran Nguyen,<sup>1</sup>  
Richard C. Stewart,<sup>2</sup> and Michael D. Manson<sup>1\*</sup>

*Department of Biology, Texas A&M University, College Station, Texas 77843,<sup>1</sup> and Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland 20742<sup>2</sup>*

Received 5 July 2002/Accepted 4 November 2002

**We have investigated the conditions required for polar localization of the CheZ phosphatase by using a CheZ–green fluorescent protein fusion protein that, when expressed from a single gene in the chromosome, restored chemotaxis to a  $\Delta cheZ$  strain. Localization was observed in wild-type,  $\Delta cheZ$ ,  $\Delta cheYZ$ , and  $\Delta cheRB$  cells but not in cells with *cheA*, *cheW*, or all chemoreceptor genes except *aer* deleted. Cells making only CheA-short (CheA<sub>S</sub>) or CheA lacking the P2 domain also retained normal localization, whereas cells producing only CheA-long or CheA missing the P1 and P2 domains did not. We conclude that CheZ localization requires the truncated C-terminal portion of the P1 domain present in CheA<sub>S</sub>. Missense mutations targeting residues 83 through 120 of CheZ also abolished localization. Two of these mutations do not disrupt chemotaxis, indicating that they specifically prevent interaction with CheA<sub>S</sub> while leaving other activities of CheZ intact.**

CheZ is a phosphatase that accelerates the removal of the intrinsically labile phosphoryl group from CheY-P (19, 28). Zhao et al. (30) recently determined the crystal structure of the CheZ dimer complexed with two CheY monomers containing the phosphoryl analog BeF<sub>3</sub><sup>-</sup>. CheZ can also be isolated in a complex with CheA-short (CheA<sub>S</sub>) (26), a short form of CheA produced from an internal translation initiation site at codon 98 of *cheA* (9). This complex was first identified in coimmunoprecipitation experiments, but it can also form with purified proteins. In experiments done in vitro at 4°C to slow the spontaneous dephosphorylation of CheY-P, the CheZ/CheA<sub>S</sub> complex showed a 2.3-fold-higher phosphatase activity than free CheZ (26). Binding of CheA<sub>S</sub> to CheZ may be inhibited by CheW, as indicated by the decreased level of coimmunoprecipitation when CheW is overexpressed (27). CheA<sub>S</sub> lacks the phosphoryl-accepting His-48 residue of CheA-long (CheA<sub>L</sub>), but it is catalytically active and can phosphorylate CheA<sub>L</sub> in *trans* (29).

Using CheZ fused to yellow fluorescent protein (1), Sourjik and Berg (23) found that CheZ localizes to the subpolar chemoreceptor clusters identified by Maddock and Shapiro (12). We had been using CheZ fused to green fluorescent protein (GFP) (1) to study the distribution of CheZ in cells and, after learning of Sourjik and Berg's work, asked whether CheA<sub>S</sub> is needed for this localization and what part of CheZ is required.

**Construction of a CheZ-GFP chimera.** A *cheZ-gfp* gene fusion was created by PCR with primers encoding a 7-amino-acid flexible linker (GGSSAAG). The fusion gene was cloned into

the pBAD18 vector, and the resulting plasmid, pBJC101 (Table 1), was shown to enable the  $\Delta cheZ$  strain RP1616 to form wild-type chemotactic swarms in tryptone semisolid agar containing 0.002% arabinose. The fusion gene was subsequently cloned into the vector plasmid pCJ30 to facilitate its insertion into the chromosome by using the  $\lambda$ InCh system (5). The resulting pBJC104 plasmid was also able to restore chemotactic swarming to strain RP1616, even in the absence of the inducer IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside). The *cheZ-gfp* gene was inserted from plasmid pBC104 into the chromosome of strain RP1616. The BC200 strain created in this way made wild-type chemotactic swarms when transcription of *cheZ-gfp* was induced with 1 mM IPTG. In the absence of IPTG, little or no swarming was observed.

Extracts from cells containing plasmid pBJC104 or the chromosomal insertion of *cheZ-gfp* contained a protein of the size expected (~54 kDa) for CheZ-GFP, whether the immunoblots were developed with anti-CheZ or anti-GFP antibody. Only a small amount of normal-length CheZ was detected. CheZ expressed from plasmid pBJC100 also supported the best swarming in strain RP1616 at 0.002% arabinose. Since the amounts of CheZ and CheZ-GFP, estimated from immunoblots, appeared to be about the same under these conditions of induction, we concluded that CheZ-GFP is functional and responsible for the observed complementation of  $\Delta cheZ$ .

**Subcellular localization of CheZ-GFP.** Subcellular localization of GFP fluorescence was examined in strains containing the *cheZ-gfp* gene in single copy on the chromosome. In strain BC200 ( $\Delta cheZ$ ), CheZ-GFP localized to patches as was previously observed by Sourjik and Berg (23) for plasmid-encoded CheZ-yellow fluorescent protein. All cells exhibited diffuse cytoplasmic fluorescence, but bright, localized patches of fluorescence were seen in 85% of the cells, primarily near the

\* Corresponding author. Mailing address: Department of Biology, Texas A&M University, College Station, TX 77843. Phone: (979) 845-5158. Fax: (979) 845-2891. E-mail: mike@mail.bio.tamu.edu.

† Present address: Division of Biology and Biomedical Sciences, Washington University School of Medicine, St. Louis, MO 63110.

TABLE 1. Strains, plasmids, and phage

Strain, plasmid, or phage	Genotype or phenotype	Comments	Reference or source
<b>Strains</b>			
RP437	<i>thr-1</i> (Am) <i>leuB6 his-4 metF159</i> (Am) <i>eda-50 rpsL1356 thi-1 ara-14 mtl-1 xyl-5 tonA31 tsx-78 lacY1</i> F <sup>-</sup>		18
AJW536	RP437 <i>cheA</i> (M98L) <i>zig::Tn10, polA</i> (Ts)		14
RP1616	RP437 $\Delta$ <i>cheZ</i> 6725		J. S. Parkinson
RP9535	RP437 $\Delta$ <i>cheA</i> 1643 <i>eda</i> <sup>+</sup>		16
RP1078	RP437 $\Delta$ ( <i>cheW-tap</i> )2217		18
RP2867	RP437 $\Delta$ ( <i>tap-cheB</i> )224 <i>eda</i> <sup>+</sup>		18
RP5231	RP437 $\Delta$ ( <i>cheY-cheZ</i> )4213 <i>eda</i> <sup>+</sup>		J. S. Parkinson
RP1515	RP437 <i>cheA</i> 169(Am) $\Delta$ <i>lac-169</i> <i>eda</i> <sup>+</sup>		22
RP1516	RP437 <i>cheA</i> 157(Am) $\Delta$ <i>lac-169</i> <i>eda</i> <sup>+</sup>		22
UU1118	RP437 <i>cheA</i> ( $\Delta$ 7–247) <i>eda</i> <sup>+</sup>		7
UU1121	RP437 <i>cheA</i> ( $\Delta$ 150–247) <i>eda</i> <sup>+</sup>		J. S. Parkinson
VB13	RP437 $\Delta$ <i>tsr</i> 7021 $\Delta$ ( <i>tar-tap</i> )5201 <i>trg::Tn10 thr</i> <sup>+</sup> <i>eda</i> <sup>+</sup>		28
DHB6521	SM551 ( $\lambda$ InCh1 lysogen)		5
BC200	RP1616 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into RP1616 via $\lambda$ InCh1	This study
BC201	RP437 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into RP437 via $\lambda$ InCh1	This study
BC203	VB13 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into VB13 via $\lambda$ InCh1	This study
BC206	RP9535 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into RP9535 via $\lambda$ InCh1	This study
BC207	RP1078 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into RP1078 via $\lambda$ InCh1	This study
BC208	RP5231 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into RP5231 via $\lambda$ InCh1	This study
BC209	AJW536 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into AJW536 via $\lambda$ InCh1	This study
BC210	UU1121 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into UU1121 via $\lambda$ InCh1	This study
BC211	UU1118 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into UU1118 via $\lambda$ InCh1	This study
BC212	RP1515 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into RP1515 via $\lambda$ InCh1	This study
BC213	RP1516 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into RP1516 via $\lambda$ InCh1	This study
BC214	RP2867 $\Delta$ ( $\lambda$ <i>att-lom</i> ):: <i>bla lacI</i> <sup>q</sup> <i>ptac-cheZ-gfp</i>	pBJC104 into RP2867 via $\lambda$ InCh1	This study
<b>Plasmids and phage</b>			
pAG3	<i>ptac cheA</i> (1–149) Amp <sup>r</sup>	Expresses CheA P1 domain	7
pBAD18	<i>araC</i> <sup>+</sup> Amp <sup>r</sup>	<i>paraBAD</i> expression vector	8
pCJ30	<i>lacI</i> <sup>q</sup> Amp <sup>r</sup>	<i>ptac</i> expression vector	2
pPM2	<i>gfp mut-2</i>	Expresses GFP Mut2	6
pBJC100	<i>cheZ</i> Amp <sup>r</sup>	<i>cheZ</i> in pBAD18	This study
pBJC101	<i>paraBAD cheZ-gfp</i> Amp <sup>r</sup>	<i>cheZ-gfp</i> in pBAD18	This study
pBJC102	<i>gfp mut-2</i> Amp <sup>r</sup>	<i>gfp mut 2</i> in pBAD18	This study
pBJC104	<i>ptac cheZ-gfp</i> Amp <sup>r</sup>	<i>cheZ-gfp</i> in pCJ30	This study
$\lambda$ InCh1	Kan <sup>r</sup> CI857	$\lambda$ InCh for pBR-derived plasmids	5

poles but also laterally in 15 to 20% of the cells (Fig. 1A and B). This pattern was similar to that seen by immunofluorescence (12) with antibody to CheA or the serine chemoreceptor Tsr. Thus, these patches are likely to represent CheZ-GFP associated with clusters containing chemoreceptors, CheA and CheW. Similar patterns of fluorescence were observed in cells of strain RP1616 carrying plasmid pBJC101.

Essentially identical numbers and distributions of patches of fluorescence were seen cells of strain BC214 ( $\Delta$ *tap-cheB*) and BC208 ( $\Delta$ *che-cheZ*), but no patches were detectable in cells of strain BC206 ( $\Delta$ *cheA*) (Fig. 1C), strain BC207 ( $\Delta$ *cheW-tap*) or strain BC203 ( $\Delta$ *tar-tap*  $\Delta$ *tsr trg::Tn10*). These results mirror the strain dependence for formation of receptor clusters (11, 12). Immunogold labeling of thin sections of cells that was carried out with anti-CheZ antibody indicated that the wild-type CheZ protein also localizes in clusters near the cell poles (J. R. Maddock, personal communication).

**Localization of CheZ-GFP to receptor patches requires CheA<sub>S</sub> but not CheA<sub>L</sub>.** We next examined the dependence of CheZ-GFP patching on CheA<sub>S</sub> and CheA<sub>L</sub>. Strain BC209 produces only CheA<sub>L</sub> because ATG codon 98 of *cheA*, the CheA<sub>S</sub> start codon (9), has been changed to CTG. The M98L version

of CheA<sub>L</sub> has about 70% of the kinase activity of wild-type CheA, and a strain producing M98L CheA<sub>L</sub> makes swarms with 70% the diameter of wild-type swarms (21). Strains BC212 and BC213 contain amber mutations in the *cheA* sequence between the start codon of CheA<sub>L</sub> and codon 98, so that both strains produce only CheA<sub>S</sub>. A schematic of the CheA polypeptides produced by these strains is shown in Fig. 2A. The ratio of the intensities of bands of CheA<sub>L</sub> versus CheA<sub>S</sub> detected on immunoblots prepared with our CheA antiserum (22) is 2.3:1 in extracts from strain BC200 (Fig. 2B). The ratio of the intensities of the CheA<sub>L</sub> band in extracts of strain BC209 and of the CheA<sub>S</sub> band in extracts from strains BC212 and BC213 is about the same. Thus, each form of CheA exists in a normal amount in the absence of the other. We note that because most of the highly antigenic P1 domain is missing from CheA<sub>S</sub>, our immunoblots may underestimate its amount. Thus, the actual CheA<sub>L</sub>/CheA<sub>S</sub> ratio in strain BC200 may be closer to the 1:1 value previously reported for late exponential-phase, highly motile *Escherichia coli* cells (27).

Cells of strain BC209 contained no detectable bright patches and showed only the diffuse cytoplasmic fluorescence found in cells of strain BC206 ( $\Delta$ *cheA*) (Fig. 1C). Cells of the CheA<sub>S</sub>-

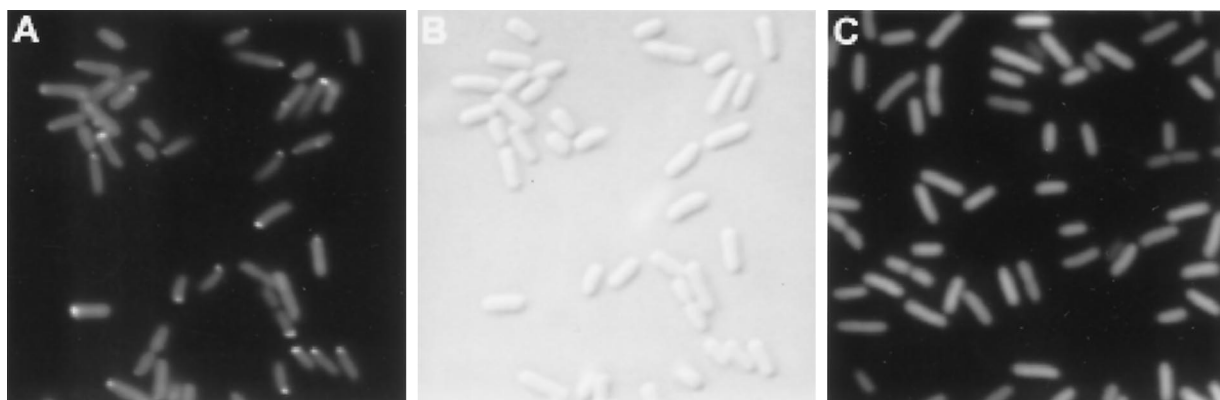


FIG. 1. CheZ-GFP localizes to patches. (A) Fluorescence micrograph of BC200 cells grown to late exponential phase at 32°C in tryptone broth (15) containing 1 mM IPTG. The cells exhibit uniform background fluorescence, presumably due to CheZ-GFP dimers in the cytoplasm. Localized bright patches of fluorescence are visible in most cells. (B) Differential interference contrast photomicrograph of the same cells as in panel A. Comparison of the two images reveals that nearly all cells fluoresce with the same intensity, indicating that their CheZ-GFP contents are similar. (C) Fluorescence micrograph of cells of strain BC206 ( $\Delta cheA$ ) grown as described for panel A. Note that the level of background fluorescence is the same but that no intense patches of fluorescence are visible. Cells were observed at a magnification of  $\times 1,575$  with a Zeiss Axioplan 2, and the images were captured with a Hamamatsu C5810 charge-coupled device camera. The peak excitation wavelength was 484 nm, and emitted light was detected at 510 to 530 nm.

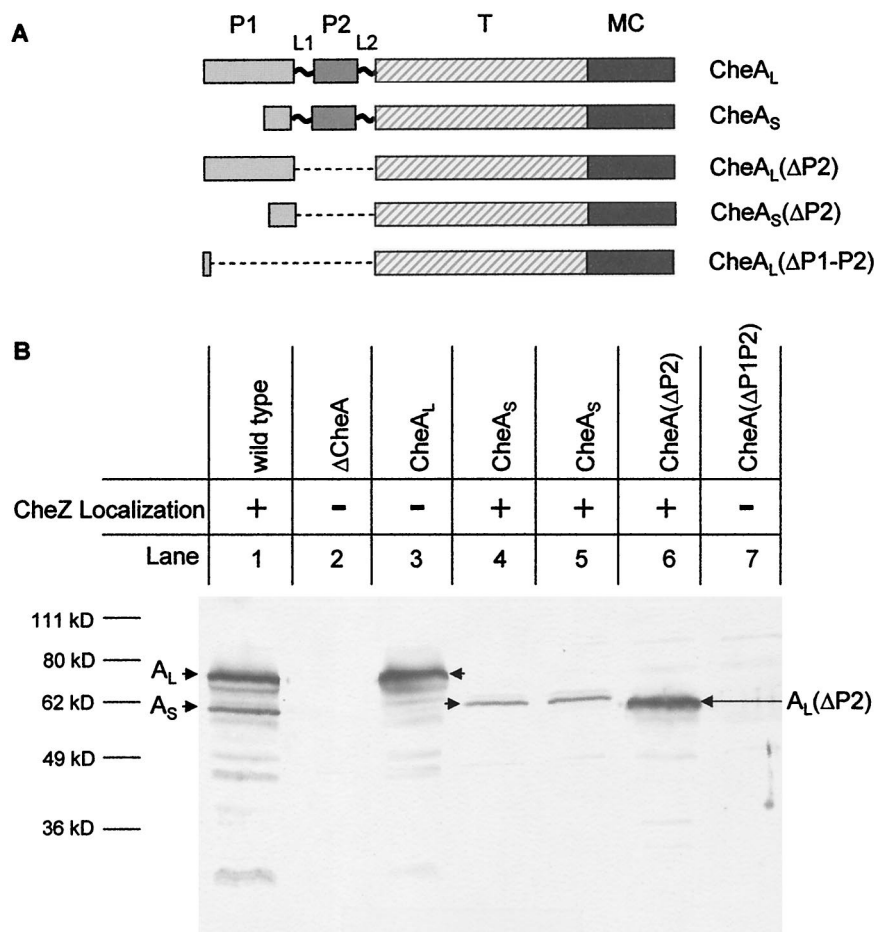


FIG. 2. Cellular levels of various forms of the *E. coli* CheA protein. (A) Schematic representation of CheA<sub>L</sub>, CheA<sub>S</sub>, CheA<sub>L</sub>( $\Delta P2$ ), CheA<sub>S</sub>( $\Delta P2$ ), and CheA( $\Delta P1-P2$ ). P1, P2, T, and MC represent the phosphorylation, CheY-binding, dimerization and catalytic, and CheW/receptor input domains, respectively. (B) Immunoblot with polyclonal CheA antiserum (24). Extracts were prepared from 1 ml of cells grown to an optical density at 590 nm of 0.8 and resuspended in 100  $\mu$ l of sodium dodecyl sulfate loading dye. Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (10% gel). Lanes: 1, BC201; 2, BC206; 3, BC209; 4, BC212; 5, BC213; 6, BC210; 7, BC211. The CheA protein species produced by each strain is indicated above the lane, and the corresponding band is indicated with an arrow.

only strains BC212 and BC213 localized CheZ-GFP to patches just like strain BC200 (Fig. 1A). These data suggested that CheZ interacts with CheA<sub>S</sub> at the patch *in vivo* and is required for the pattern of CheZ localization observed in wild-type cells. It has previously been shown that CheA<sub>L</sub>-only and CheA<sub>S</sub>-only cells form receptor patches with equal facility (21).

**The N-terminal region of CheA<sub>S</sub> may bind CheZ.** Strain BC210 (*cheAΔP2*) expresses a CheA protein missing the P2 domain, which binds CheY (16, 25). Its swarming ability is only slightly decreased from that of its wild-type parental strain RP437. Strain BC211 [*cheAΔ(P1-P2)*] expresses a CheA protein in which a *cheA*-internal deletion removes most of the P1 and P2 domains, and it does not form chemotactic swarms. Chromosomally encoded CheZ-GFP had a patchy distribution in cells of strain BC210 but not in cells of strain BC211. Expression of the P1 domain (25) from plasmid pAG3 enabled strain BC211 to form chemotactic swarms with about half the diameter of those made by strain RP437 but did not restore CheZ-GFP localization in cells of this strain. Although our polyclonal CheA antiserum did not visualize the CheAΔ(P1-P2) protein in immunoblots, it must be present at some level for complementation to occur. A protein of the size expected for CheA<sub>L</sub>ΔP2 was detected with this antiserum in an extract from strain BC210 (Fig. 2B), but no band was seen at the position expected for CheA<sub>S</sub>ΔP2.

The inability to detect CheAΔ(P1-P2) and CheA<sub>S</sub>ΔP2 with our polyclonal antiserum may represent the absence of epitopes recognized by the antiserum, since 23 of 27 monoclonal antibodies raised against full-length CheA target P1 and P2 (J. S. Parkinson, personal communication). Those data suggest that P1 and P2 are the most antigenic portions of CheA and that even polyclonal antisera against intact CheA are deficient in recognizing the remainder of the protein. We obtained monoclonal serum directed against the P4 (catalytic) domain of *E. coli* CheA (C. O'Connor and P. Matsumura, unpublished data), but it failed to visualize any form of CheA, including either full-length CheA<sub>L</sub> or CheA<sub>S</sub>, on immunoblots.

**Mutations within a specific region of *cheZ* eliminate polar localization of CheZ-GFP.** To identify which part(s) of CheZ is responsible for localization, we introduced a selection of previously identified *cheZ* missense mutations (3, 4, 20) scattered throughout the gene into plasmid pBJC101. Of these 17 mutations, the ones causing the L90S and F117S substitutions completely eliminated polar localization of CheZ-GFP, and cells expressing these proteins failed to form patches, looking just like cells of the *ΔcheA* strain BC206 (Fig. 1C). Error-prone PCR mutagenesis (31) generated a mutant CheZ-GFP protein containing the W94R substitution that also did not localize to receptor patches.

A summary of the swarming behavior and localization patterns supported by these mutant proteins is given in Table 2. Note that the T25P, L28P, A87V, and A87G mutant proteins showed an intermediate level of patch formation; many cells lacked visible patches, but a significant minority of cells showed essentially normal patterns of CheZ-GFP localization. Except for A87V, which completely eliminated swarming, these substitutions caused partial defects in chemotactic swarm formation. Immunoblot analyses with anti-CheZ antibody indicated that most of the mutant CheZ-GFP proteins were present in normal amounts, regardless of their ability to function in che-

TABLE 2. Effect of *cheZ* mutations on CheZ-GFP localization and chemotactic swarming

Mutation source	Mutation	Localization <sup>a</sup>	Swarm phenotype <sup>b</sup>	Reference
Random mutagenesis	W94R	–	–	This study
Previously published	T25P	+	+	20
	L28P	+	+	20
	D50G	++	+	20
	A65V	++	–	4
	M83T	++	–	4
	A87G	+	+	4
	A87V	+	–	4
	L90S	–	–	4
	F117S	–	–	4
	F141I	++	–	3
	D143G	++	–	4
	T145M	++	–	3
	I149T	++	–	20
	E158G	++	+	20
	N182Y	++	–	20
	G188E	++	–	4
	V205E	++	–	4
Site-directed mutagenesis of amphipathic helices	W94S	–	–	This study
	W97S	–	+	This study
	F98S	–	++	This study
	I102S	++	++	This study
	L104S	++	++	This study
	A107S	++	+	This study
	L110S	–	–	This study
	V111S	–	–	This study
	T114A	–	–	This study
	L118S	–	–	This study
	V121S	–	–	This study

<sup>a</sup> ++, wild-type localization of CheZ-GFP; +, weaker localization or localization in fewer cells; –, no localization.

<sup>b</sup> ++, wild-type swarm in aspartate-minimal motility agar; +, swarm with a smaller diameter and less-distinct chemotactic rings; –, no swarming, as with the *ΔcheZ* parent strain.

motaxis or localize to patches. The exceptions were the T25P and L28P proteins, which were found at ~50% of normal levels based on relative band intensities on immunoblots.

**Mutations that disrupt CheZ-GFP localization alter residues on the hydrophobic faces of two amphipathic helices.** A computer-generated secondary-structure prediction for CheZ indicated that the region of CheZ targeted by localization-defective (Loc<sup>–</sup>) mutations forms two α helices separated by a short loop. This prediction has been confirmed by the recently published structure of CheZ (30). Helical-wheel projections revealed that each helix should be amphipathic and that Ala-87, Leu-90, Trp-94, and Phe-117 are located on hydrophobic faces of the helices. Since Loc<sup>–</sup> mutations affecting these residues exchange polar residues for nonpolar ones, we hypothesized that hydrophobic interactions involving these residues could be important for CheZ function.

Almost all of the residues on the hydrophobic faces of the predicted helices were replaced with Ser by site-directed mutagenesis. The only exceptions were Met-83, which we had already tested as the previously identified M83T mutation (4), and Thr-114, which was replaced with Ala. Each mutant protein was expressed from plasmid pBJC104 in strain RP1616, and its chemotaxis and CheZ-GFP localization phenotype were analyzed (Table 2). Ser replacements in the region from

## A

## Enteric species

				****			
<i>Escherichia coli</i> CheZ (78)	PHQ	DQMEKSAKAL	TQRWDDWFAD	PIDLADAREL	VTDTRQFLPDV		
<i>Salmonella enterica</i> CheZ (78)	PHQ	DAMEKEAKAL	TQRWDEWFEN	PIELSDAREL	VTDTRQFLRDV		
<i>Enterobacter cloacae</i> CheZ (79)	PHQ	DAMEKGAKAL	SKRWDEWFEN	PIELADAREL	VTDTRQYLGDV		
<i>Yersinia pestis</i> CheZ (78)	PRQ	NELESSAKAL	KIRWDEWFAN	PIELSDARSL	VTDTREYLAIV		

## Non-enteric species

<i>Vibrio cholera</i> CheZ (99)	PIA	DKLHESLLLI	RPEWNGLMNG	RIELMHFKSL	CHRIDDLISQV
<i>Pseudomonas aeruginosa</i> CheZ (78)	PLV	NQLGDDSREL	HQEWQRFMR	EIDADGFREL	AKRIEQFLVRS
<i>Pseudomonas putida</i> CheZ (112)	PVL	NDLASEAKAL	STDWQRFMR	EVAAPEFRDL	VKRVDSFLTHS
<i>Ralstonia solanacearum</i> CheZ (86)	PVQ	DQLEGEAEEL	VTRWQSWMDR	QLGDDEIRSL	VEITNGFLRSV
<i>Xanthomonas campestris</i> CheZ (106)	LLA	NQLR-----	-----D--G-	-----	-----GLNQDQ

## B

## Enteric species

											*
<i>E. coli</i> CheA (82)	QLNTDIINL	FLETKDIMQE	QLDAYKQSQE	PDAASFYIC	QALRQLALEA	KGETPSAVTR	LSVVAKE				
<i>S. enterica</i> CheA (82)	QLNTDIINL	FLETKDIMQE	QLDAYKNSEE	PDAASFEYIC	NALRQLALEA	KGETTPAVVE	TAALSAAI				
<i>Y. pestis</i> CheA (82)	RLSTEIINL	FLETKDIMQE	QLDAYKTSQE	PHSESEFYIC	QALRQLALDA	LDQPTTEDQP	TTENQSAT				

## Non-enteric species

<i>V. cholera</i> CheA (83)	HVSPSILMDT	MLKALDITVNE	QFRAVQEREP	LQPADPELLD	ELHRLSKP--	-----ASED	EDEAAEAH
<i>P. aeruginosa</i> CheA (83)	RVSSSEIMDV	VIQALDITVNA	MFDQVREQE	PTPATPELLA	ALARLAEPEG	AEPAPVQAP	PAVPPAE
<i>P. putida</i> CheA (83)	RVDAELMDV	VLEALDITVNS	MFGQVRERAD	VTPATPQLLA	ALSRLAEP--	-GGAEAAEAP	APAAPVAE
<i>R. solanacearum</i> CheA (82)	QLRENMVDA	FLETKDVLKS	QLEAYRHEHP	IDTATLEYMV	AKLNLSLTAED	GAPATAAAPA	AAAPAAAP
<i>X. campestris</i> CheA (83)	ALQSELIAL	LLVCCDHIHA	LVETAADPSH	ADPAALAAEA	EPLLAQLQTY	LQGSVCGATV	SDAQRSTP

FIG. 3. Alignment of amino acid sequences of the CheZ apical loop region (30) and the C-terminal portion of the P1 domain of CheA (17). (A) Alignment of CheZ sequences from representatives of each bacterial genus for which a putative *cheZ* gene has been sequenced. Asterisks indicate the conserved D(D/E)WF motif at residues 95 to 98 of *E. coli* CheZ. The number in parentheses indicates the residue number for the first position in the sequence shown. Residues highlighted in dark gray are identical in all of the aligned sequences except that of *X. campestris*. Residues highlighted in light gray are identical within the enteric or the nonenteric bacteria. Boldface type indicates positions at which residues are chemically conserved within the enteric or nonenteric bacteria. The sequence listed as *Salmonella enterica* is for serovar Typhimurium. In addition to the species and serovars shown, comparisons were made with the *cheZ* sequences from *Salmonella enterica* serovar Typhi, *Yersinia enterocolitica*, *Vibrio parahaemolyticus*, *Pseudomonas syringae*, and *Xanthomonas axonopodis*. Those sequences did not differ significantly from those shown for their congeners. (B) Alignment of the available amino acid sequences for the C-terminal portion of the P1 domain of CheA from the species whose CheZ sequences are shown in panel A. All sequences were obtained from the Entrez database (<http://www.ncbi.nlm.nih.gov/PubMed/>) and aligned by using the AlignX program of the Vector NTI Suite molecular biology software package.

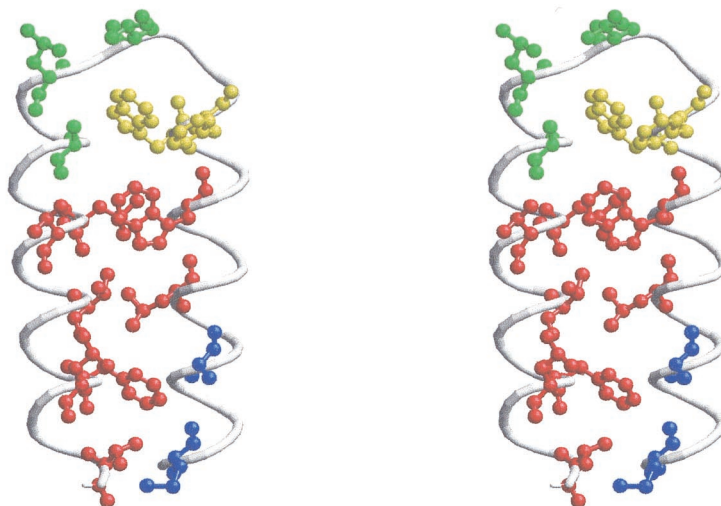
L110 to V121 and the T114A substitution conferred a Che<sup>-</sup>Loc<sup>-</sup> phenotype, whereas the I102S, L104S, and A107S substitutions had little effect on either chemotaxis or localization. The W94S replacement imposed the same Che<sup>-</sup>Loc<sup>-</sup> phenotype as W94R. Cells expressing the W97S or F98S version of CheZ-GFP had a Che<sup>+</sup>Loc<sup>-</sup> phenotype, implying that the overall conformation and function of these mutant proteins were not significantly compromised. In the absence of the inducer IPTG, each mutant protein was produced at physiologically normal levels in about the same amount as wild-type CheZ-GFP.

Proposed CheA<sub>S</sub>-CheZ interaction sites are conserved in

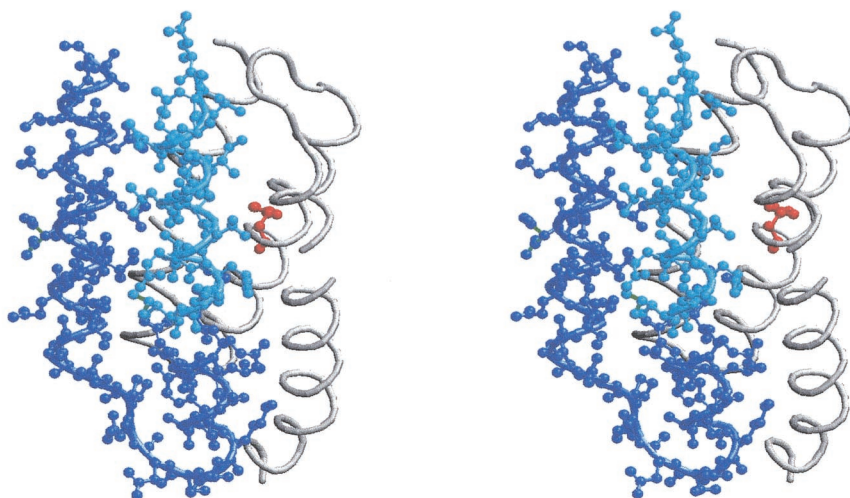
enteric bacteria. The deduced amino acid sequences corresponding to the predicted helix-turn-helix hairpin of *E. coli* CheZ were compared for the 14 gram-negative proteobacteria for which CheZ sequences were available in the Entrez database. Figure 3A presents the alignment of nine of these sequences, selected to avoid redundancy and to represent at least one member of each genus. Residues 95 to 98 are conserved as D(D/E)WF in all of the enteric species, which are the only bacteria known to express CheA<sub>S</sub> (14). Nonenteric  $\gamma$ -proteobacteria, including *Vibrio* and *Pseudomonas*, do not display this motif. However, there is substantial sequence conservation at other positions in this region, with the notable exception of

FIG. 4. Stereo views of the crystal structures for the apical loop region of CheZ (28) and the P1 domain of CheA (15). (A) Apical loop of *E. coli* CheZ. Residues 83 to 121 are shown. Substitutions at the residues in blue displayed a Che<sup>-</sup>Loc<sup>+</sup> phenotype. Substitutions at the residues in red produced a Che<sup>-</sup>Loc<sup>-</sup> phenotype. Substitutions at the residues in yellow resulted in a Che<sup>+</sup>Loc<sup>-</sup> phenotype. Residues shown in green were not affected in either chemotaxis or localization by substitution at these positions. (B) *Salmonella* CheA P1 domain. Residues in dark blue are in  $\alpha$ -helices 4 and 5 and are present CheA<sub>S</sub>. Residues in cyan are in the N-terminal portion of  $\alpha$ -helix 4 missing in CheA<sub>S</sub> and make hydrophobic contacts with residues in  $\alpha$ -helix 5 of full-length P1. The site of phosphorylation (His-48) is indicated in red. (C) Conserved residues in the remnant P1 domain of CheA<sub>S</sub>. Yellow residues are conserved among all enteric species. Those indicated in dark gray are chemically conserved among enteric species, whereas those shown in light gray are variable.

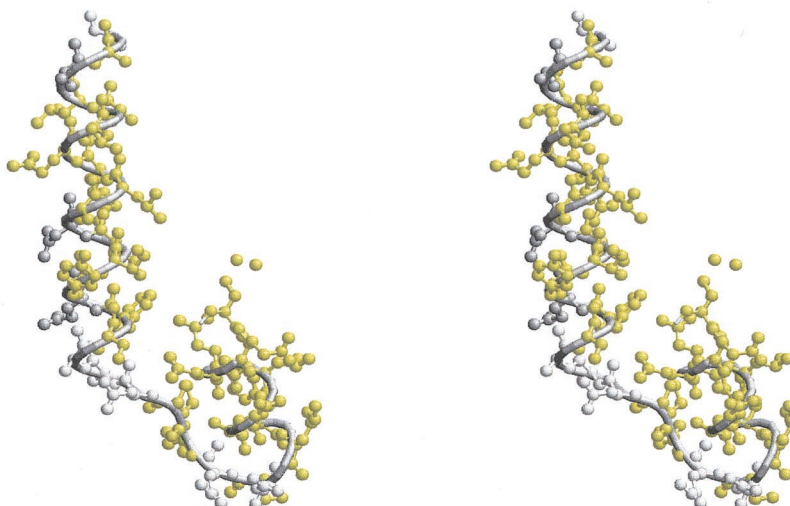
A



B



C



the CheZ proteins from the two *Xanthomonas* species, which lack the entire region encompassing the apical helix-turn-helix hairpin. When the sequences of CheA from these same species are compared (Fig. 3B), it is clear that sequences corresponding to the putative N terminus of CheA<sub>S</sub> are also conserved in the enteric bacteria.

**Conclusion.** The results presented here establish that CheZ localization to the receptor patch occurs only when CheA<sub>S</sub> is present. The two residues of CheZ most clearly implicated in interaction with CheA<sub>S</sub> are Phe-97 and Trp-98. These two aromatic residues, especially Phe-97, are solvent exposed in the apical hairpin loop of the CheZ crystal structure (Fig. 4A) and might be expected to have an energetically favorable interaction with a hydrophobic partner. Mutations causing a Che<sup>-</sup> Loc<sup>-</sup> phenotype replace hydrophobic residues at the interhelix packing surface of the hairpin (Fig. 4A) or residues located at the subunit interface of the CheZ homodimer (30). These substitutions may destabilize the hairpin or interfere with CheZ dimerization, respectively.

The N-terminal sequence of CheA<sub>S</sub> corresponds to the C-terminal portion of the P1 domain. The crystal structure of the *Salmonella* P1 domain (17) is shown in Fig. 4B. It contains five  $\alpha$ -helices, and Met-98, the first residue of CheA<sub>S</sub>, resides in the middle of the fourth helix, which extends to Lys-106. Residues Ala-113 through Ala-130 comprise the fifth helix. Based on this structure, the N terminus of CheA<sub>S</sub> is predicted to be an amphipathic helix of 8 residues followed by a turn and an amphipathic helix of 18 residues. In CheA<sub>S</sub>, the hydrophobic residues Leu-123 and Leu-126 should be exposed to solvent and available to interact with CheZ. Furthermore, purified CheZ binds to a 'P1-P2 fragment derived from CheA<sub>S</sub> but not to a P1-P2 fragment derived from CheA<sub>L</sub> (L. Kott and R. M. Weis, personal communication), and CheZ can be coprecipitated with an N-terminal fragment of CheA<sub>S</sub> containing its first 42 residues fused to GST (O'Connor and Matsumura, personal communication).

The absence of CheA<sub>S</sub> does not diminish chemotaxis in the assays that have been employed (19). However, within the enteric bacteria the selective conservation of amino acid sequences in the regions of CheA<sub>S</sub> and CheZ that we propose to interact (Fig. 3) constitutes a clear example of an evolutionary trace as defined by Lichtarge and colleagues (10, 13). When combined with knowledge of the concentrated distribution of a set of conserved residues on the surfaces of both of two potentially interacting proteins, such "evolutionarily privileged clusters" are strong predictors of protein-protein binding sites. The central premise of evolutionary trace analysis is that evolution has already done the experiment of mutation and functional analysis to identify important residues. The ability to disrupt CheZ localization through two otherwise benign substitutions of residues in a sequence that is conserved in the CheZ proteins of enteric bacteria (14) implies that the CheZ-CheA<sub>S</sub> interaction confers a selective advantage under some environmental conditions.

Larry Griffing mentored B.J.C in fluorescence microscopy. Michael Eisenbach, Ruth Silversmith, and Viktor Sourjik commented on an early draft of the manuscript. Sandy Parkinson and Alan Wolfe supplied strains, and Phil Matsumura supplied monoclonal antibody directed against the P4 domain of CheA. Janine Maddock graciously

communicated unpublished data on CheZ subcellular localization determined by immunogold labeling. Laila Kott and Bob Weis, and Phil Matsumura and Chris O'Connor, kindly shared unpublished information about binding of CheZ to the N terminus of CheA<sub>S</sub> in vitro.

Financial support was provided by Public Health Service grants GM37369 to M.D.M. and GM52583 to R.C.S.

## REFERENCES

- Baumann, C. T., C. S. Lim, and G. L. Hager. 1998. Simultaneous visualization of the yellow and green forms of the green fluorescent protein in living cells. *J. Histochem. Cytochem.* **46**:1073-1076.
- Bibikov, S. I., R. Biran, K. E. Rudd, and J. S. Parkinson. 1997. A signal transducer for aerotaxis in *Escherichia coli*. *J. Bacteriol.* **179**:4075-4079.
- Blat, Y., and M. Eisenbach. 1996. Mutants with defective phosphatase activity show no phosphorylation-dependent oligomerization of CheZ. *J. Biol. Chem.* **271**:1232-1236.
- Boesch, K. C., R. E. Silverman, and R. B. Bourret. 2000. Isolation and characterization of nonchemotactic CheZ mutants of *Escherichia coli*. *J. Bacteriol.* **182**:3544-3552.
- Boyd, D., D. S. Weiss, J. C. Chen, and J. Beckwith. 2000. Towards single-copy gene expression systems making gene cloning physiologically relevant: lambda InCh, a simple *Escherichia coli* plasmid-chromosome shuttle system. *J. Bacteriol.* **182**:842-847.
- Cormack, B. P., R. H. Valdivia, and S. Falkow. 1996. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**:33-38.
- Garzon, A., and J. S. Parkinson. 1996. Chemotactic signaling by the P1 phosphorylation domain liberated from the CheA histidine kinase of *Escherichia coli*. *J. Bacteriol.* **178**:6752-6758.
- Guzman, L., D. Belin, M. J. Carson, and J. Beckwith. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P<sub>BAD</sub> promoter. *J. Bacteriol.* **177**:4121-4130.
- Kofoed, E. C., and J. S. Parkinson. 1991. Tandem translation starts in the *cheA* locus of *Escherichia coli*. *J. Bacteriol.* **173**:2116-2119.
- Lichtarge, O., and M. E. Sowa. 2002. Evolutionary predictions of binding surfaces and interactions. *Curr. Opin. Struct. Biol.* **12**:21-27.
- Lybarger, S. R., and J. R. Maddock. 1999. Clustering of the chemoreceptor complex in *Escherichia coli* is independent of the methyltransferase CheR and the methyl-erasure CheB. *J. Bacteriol.* **181**:5527-5529.
- Maddock, J. R., and L. Shapiro. 1993. Polar location of the chemoreceptor complex in *Escherichia coli*. *Science* **259**:1717-1723.
- Madubushi, S., H. Yao, H., M. Marsh, M., D. M. Kristensen, A. Philippi, M. E. Sowa, and O. Lichtarge. 2002. Structural clusters of evolutionary trace residues are statistically significant and common in proteins. *J. Mol. Bio.* **316**:139-154.
- McNamara, B. P., and A. J. Wolfe. 1997. Coexpression of the long and short forms of CheA, the chemotaxis histidine kinase, by members of the family *Enterobacteriaceae*. *J. Bacteriol.* **179**:1813-1818.
- Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Morrison, T. B., and J. S. Parkinson. 1994. Liberation of an interaction domain from the phosphotransfer region of CheA, a signaling kinase of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **91**:5485-5489.
- Mourey, L., S. Da Re, J. D. Pedelacq, T. Tolstykh, C. Faurie, V. Guillet, J. B. Stock, and J.-P. Samama. 2001. Crystal structure of the CheA histidine phosphotransfer domain that mediates response regulator phosphorylation in bacterial chemotaxis. *J. Biol. Chem.* **276**:31074-31082.
- Parkinson, J. S., and S. E. Houts. 1982. Isolation and behavior of *Escherichia coli* deletion mutants lacking chemotaxis functions. *J. Bacteriol.* **151**:106-113.
- Sanatinia, H., E. C. Kofoed, T. B. Morrison, and J. S. Parkinson. 1995. The smaller of two overlapping *cheA* gene products is not essential for chemotaxis in *Escherichia coli*. *J. Bacteriol.* **177**:2713-2720.
- Sanna, M. G., and M. I. Simon. 1996. In vivo and in vitro characterization of *Escherichia coli* protein CheZ gain- and loss-of function mutants. *J. Bacteriol.* **178**:6275-6280.
- Skidmore, J. M., D. D. Ellefson, B. P. McNamara, M. M. P. Couto, A. J. Wolfe, and J. R. Maddock. 2000. Polar clustering of the chemoreceptor complex in *Escherichia coli* occurs in the absence of complete CheA function. *J. Bacteriol.* **182**:967-973.
- Smith, R. A., and J. S. Parkinson. 1980. Overlapping genes at the *cheA* locus of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **77**:5370-5374.
- Sourjik, V., and H. C. Berg. 2000. Localization of components of the chemotaxis machinery of *Escherichia coli* using fluorescent protein fusions. *Mol. Microbiol.* **37**:740-751.
- Stewart, R. C., R. VanBruggen, D. D. Ellefson, and A. J. Wolfe. 1998. TNP-ATP and TNP-ADP as probes of the nucleotide binding site of CheA, the histidine protein kinase in the chemotaxis signal transduction pathway of *Escherichia coli*. *Biochemistry* **37**:12269-12279.

25. **Swanson, R. V., S. C. Schuster, and M. I. Simon.** 1993. Expression of CheA fragments which define domains encoding kinase, phosphotransfer, and CheY binding activities. *Biochemistry* **32**:7623–7629.
26. **Wang, H., and P. Matsumura.** 1996. Characterization of the CheA<sub>S</sub>/CheZ complex: a specific interaction resulting in enhanced dephosphorylating activity on CheY-phosphate. *Mol. Microbiol.* **19**:695–703.
27. **Wang, H., and P. Matsumura.** 1997. Phosphorylating and dephosphorylating complexes in bacterial chemotaxis. *J. Bacteriol.* **179**:287–289.
28. **Ward, S. M., A. Delgado, R. P. Gunslaus, and M. D. Manson.** 2002. A NarX-Tar chimera mediates repellent taxis to nitrate and nitrite. *Mol. Microbiol.* **44**:709–719.
29. **Wolfe, A. J., B. P. McNamara, and R. C. Stewart.** 1994. The short form of CheA couples chemoreception to CheA phosphorylation. *J. Bacteriol.* **176**:4483–4491.
30. **Zhao, R., E. J. Collins, R. B. Bourret, and R. E. Silversmith.** 2002. Structure and catalytic mechanism of the *E. coli* chemotaxis phosphatase CheZ. *Nat. Struct. Biol.* **9**:570–575.
31. **Zhou, Y., X. Zhang, and R. H. Ebright.** 1991. Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. *Nucleic Acids Res.* **19**:6052.