Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 5898–5902, August 1990
Biochemistry

FrzE of *Myxococcus xanthus* is homologous to both CheA and CheY of *Salmonella typhimurium*

WILLIAM R. MCCLEARY AND DAVID R. ZUSMAN*

Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

Communicated by H. A. Barker, May 25, 1990 (received for review April 23, 1990)

ABSTRACT ***Myxococcus xanthus*** **exhibits multicellular development. The "frizzy" (*frz*) mutants are unable to complete the developmental pathway. Instead of forming fruiting bodies, these mutants form tangled filaments of cells. We have previously shown that four of the *frz* gene products are homologous to enteric chemotaxis proteins and have proposed that the *frz* genes constitute a signal-transduction pathway that controls the frequency at which cells reverse their gliding direction. We show here that *frzE* encodes a protein with a calculated molecular mass of 83 kDa. FrzE is homologous to both CheA and CheY of *Salmonella typhimurium*, which are members of a family of "two-component response regulators." It is thought that the modulator components autophosphorylate and transfer a phosphate group to their cognate effector components. FrzE contains an unusual (alanine plus proline)-rich region that might constitute a flexible hinge facilitating phosphate transfer between functional domains. We suggest that FrzE is a second messenger that relays information between the signaling protein FrzCD and the gliding motor.**

Myxococcus xanthus is a gliding, Gram-negative soil bacterium that exhibits a complex life cycle involving extensive cell–cell interactions (1, 2). Cell contacts are repeatedly broken and reestablished as cells glide, reverse direction, and are reoriented from moving on irregular surfaces. Myxobacteria exhibit directional motility by biasing the time interval between reversals. Vegetative cells move in "hunting packs" and secrete hydrolytic enzymes that degrade organic substances. Starvation triggers the developmental program of these organisms. Development is characterized by both the morphogenesis of a multicellular structure, the fruiting body, and the differentiation of rod-shaped cells into spherical, environmentally resistant and metabolically quiescent myxospores. During fruiting body morphogenesis, ≈100,000 cells aggregate to a focal center and form a mound of cells in which differentiation occurs.

Mutants defective in fruiting body formation have been useful in defining events along the developmental pathway (3). The "frizzy" (*frz*) mutants are unable to aggregate properly during mound formation but instead form tangled filaments (4). The six *frz* genes are clustered on a 7.5-kilobase-pair (kbp) segment of the chromosome and are designated *frzA–frzG* (5, 6). Individual *frz* cells show aberrant motility behavior (6, 7). When plated at low-cell density wild-type cells reverse their gliding direction approximately every 7–8 min; *frzA*, *-B*, *-C*, *-E*, and *-F* cells rarely reverse direction—approximately every 1–2 hr; *frzD* and *frzG* cells reverse more frequently than wild type—every 1–2 min and 4–5 min, respectively. Tn*5–lac* insertions in the *frz* genes permitted an analysis of their expression during development (8). Transcription is developmentally regulated and shows peak levels of expression during mound formation. Recent DNA sequence analysis revealed that many of the *frz* gene products are homologous to known bacterial enteric chemotaxis proteins (6, 9). The *frzA*, *frzCD*, *frzG*, and *frzF* gene products are homologous to CheW, the methyl-accepting chemotaxis proteins (MCPs), CheB and CheR, respectively.

The enteric chemotaxis proteins constitute a pathway that senses chemoeffector concentrations outside the cell and transduces that information to the flagella (for reviews, see refs. 10–12). The MCPs are transmembrane proteins that detect chemoeffectors in the periplasm. Their cytoplasmic signaling domains initiate an intracellular second-messenger cascade that modulates swimming behavior. CheW couples MCP output to the cytoplasmic signaling components CheA and CheY (13, 14). CheA autophosphorylates and transfers its phosphate to CheY (15). It is thought that phospho-CheY interacts directly with components of the flagellar basal body to generate a tumble (16). Sensory adaptation is mediated by the reversible methylation and demethylation of specific glutamate residues of the MCPs. CheR and CheB are the MCP-specific methyltransferase and methylesterase, respectively. The methylesterase activity of CheB is modulated by CheA-mediated phosphorylation (17). We recently showed that FrzCD is methylated *in vivo* and that this methylation depends on *frzF* (6).

The sequence similarities to chemotaxis proteins combined with the biochemical and behavioral analyses of *frz* mutants† prompted us to hypothesize that the *frz* genes are components of a sensory transduction pathway necessary for the proper aggregation of cells during fruiting body morphogenesis. The chemotaxis signal-transduction pathway controls the swimming behavior of individual cells in response to chemoeffectors. We propose that the *frz* gene products control the gliding behavior of individual cells in response to aggregation and/or feeding signals.

It is surprising that organisms that move by such different mechanisms (swimming with flagella versus gliding without flagella) use a similar signal-transduction pathway to control motility. How similar are the Che and Frz signaling pathways? Does FrzCD interact directly with the gliding motor or does it also, like the MCPs, use a second-messenger system? If so, what is the nature of this system? In this paper we address some of these questions. We demonstrate that FrzE is homologous to both CheA and CheY and discuss the implications of these findings.

MATERIALS AND METHODS

Bacterial Strains and Culture Conditions. *M. xanthus* DZF1 (18) was the fruiting-proficient strain used in this study. *M. xanthus* growth and developmental conditions were as described (6). *Escherichia coli* strain DG98 was used for all subcloning and sequencing.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MCP, methyl-accepting chemotaxis protein; E_{2p}, *Escherichia coli* dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M35192).

DNA Sequencing. The DNA corresponding to the *frzE* complementation group was subcloned from pBB12 (5) and sequenced on both strands as described (6).

RESULTS

Nucleotide Sequence of *frzE*. The *frzE* gene product is necessary for proper cellular aggregation during development. We have shown by complementation analysis (5) and by maxicell experiments (19) that *frzE* is the largest gene in the frizzy region. This gene can encode a protein with a molecular mass of 80–100 kDa. In this paper we have extended those initial studies of *frzE* through DNA sequence analysis. The 2.7-kbp region from *Sac* I to *Stu* I (Fig. 1), which contains *frzE*, was subcloned and sequenced on both strands. An open reading frame of the appropriate size was identified. This reading frame exhibits the characteristic nucleotide bias (20) of high G+C organisms, such as *M. xanthus* [68.5% G+C (21)]. *frzE* shows 76%, 45%, and 93% G+C at positions one, two, and three, respectively. An ATG translational start codon at nucleotide 64 of Fig. 2 was identified based on an appropriately spaced ribosome-binding site (GTACG) and on the homology data presented below. From this predicted start site, *frzE* encodes a 777-amino acid protein with a calculated molecular mass of 83,095 Da. Hydopathicity data (not shown) suggest that FrzE is a soluble protein.

There are only 33 noncoding nucleotides between the stop codon of *frzCD* and the start codon of *frzE* and 42 nucleotides between the stop codon of *frzE* and the initiation codon of the *frzG* gene. The compact spacing between *frzCD*, *frzE*, and *frzG* supports the hypothesis that these *frz* genes are transcribed as part of a polycistronic message (8).

Homology of FrzE to Both CheA and CheY. Computer searches of protein sequence data bases showed that FrzE was homologous to two enteric chemotaxis proteins: CheA and CheY. Fig. 3 compares the predicted FrzE amino acid sequence to that of *Salmonella typhimurium* CheA (22). In addition to the boxed amino acid identities shown in this alignment there are also numerous conservative substitutions. There is a short conserved stretch of 10 amino acids between residues 46 and 55 as well as a large conserved central region of 360 amino acids that extends from residue 294 to 653 and shows 33% amino acid identity to CheA.

Immediately next to the C terminus of the CheA homology region is a final domain of 124 amino acids extending from residues 654 to 777 that shows 31% amino acid identity to *S. typhimurium* CheY (23) (Fig. 4). This homology spans the entire CheY protein.

FrzE Contains an Unusual (Alanine Plus Proline)-Rich Domain. FrzE contains a 68-amino acid region from residue 130 to 197 that is 38% alanine and 34% proline. Similar (alanine plus proline)-rich regions are also seen in peptides of the pyruvate dehydrogenase multienzyme complex of *E. coli* (24). A comparison of the FrzE (alanine plus proline)-rich region to three of the (alanine plus proline)-rich stretches found in the *E. coli* dihydrolipoamide acetyltransferase (E_{2p}) component of pyruvate dehydrogenase is shown in Fig. 5. Although all four of these regions are ≈75% (alanine plus proline), there are some significant differences. The FrzE (alanine plus proline)-rich region is much longer and contains a higher percentage of proline (34%) than the E_{2p} segments (12–17% proline). Each of the E_{2p} proline residues is directly preceded by an alanine residue, whereas many of the FrzE prolines are not preceded by alanine. There is also a higher percentage of charged residues (10–20%) in the E_{2p} segments than in FrzE (3%). Inspection of the FrzE sequence shows no striking repeating patterns. However, an Ala-Pro-Xaa-Xaa-Ala-Pro motif is occasionally seen in FrzE and E_{2p} (underlined in Fig. 5).

DISCUSSION

In this paper we show that FrzE is homologous to both CheA and CheY, which are members of the "two-component regulatory" pathways (25). These two-component regulators of bacterial processes generally respond to changing environmental stimuli and are often members of more extensive signal-transduction pathways. CheA belongs to a large family of homologous proteins that receive input from the environment. Biochemical experiments involving several members of this family [CheA (15), NtrB (26), and EnvZ (27)] have demonstrated a common mechanism of action. These proteins are histidine protein kinases that first autophosphorylate on a histidine residue and then function as phosphotransferases (25). Their substrates are members of another family of homologous proteins (the response regulators) represented by CheY (15), NtrC (26), and OmpR (27). Phosphorylation of these regulator proteins on an aspartate residue (28) modifies their activity.

Genetic analysis of a large number of *cheA* alleles indicated that CheA is composed of three functional domains (29). It was proposed that the N-terminal domain of CheA interacts with CheY to mediate phosphotransfer. A central domain is

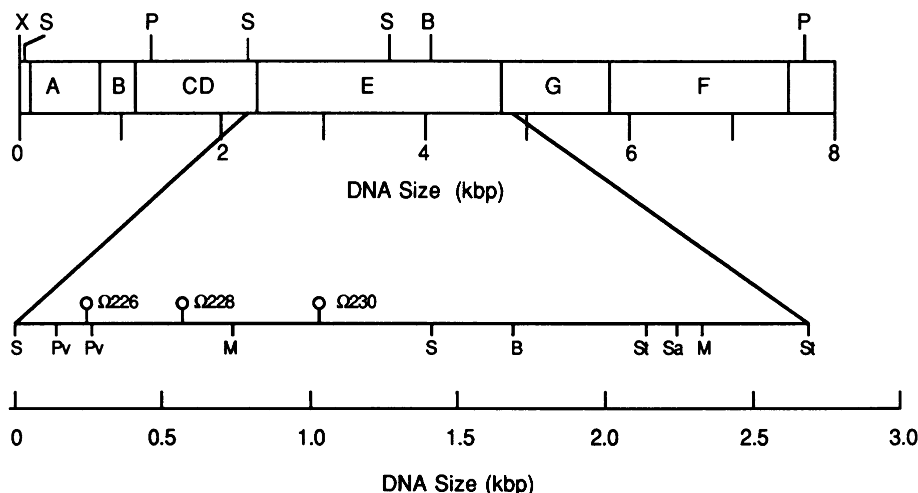

FIG. 1. Physical organization of the *M. xanthus frz* genes. ♀, Sites of Tn*5* insertion. Restriction sites: X, *Xho* I; S, *Sac* I; P, *Pst* I; Pv, *Pvu* II; M, *Mlu* I; B, *Bgl* II; St, *Stu* I; Sa, *Sal* I.

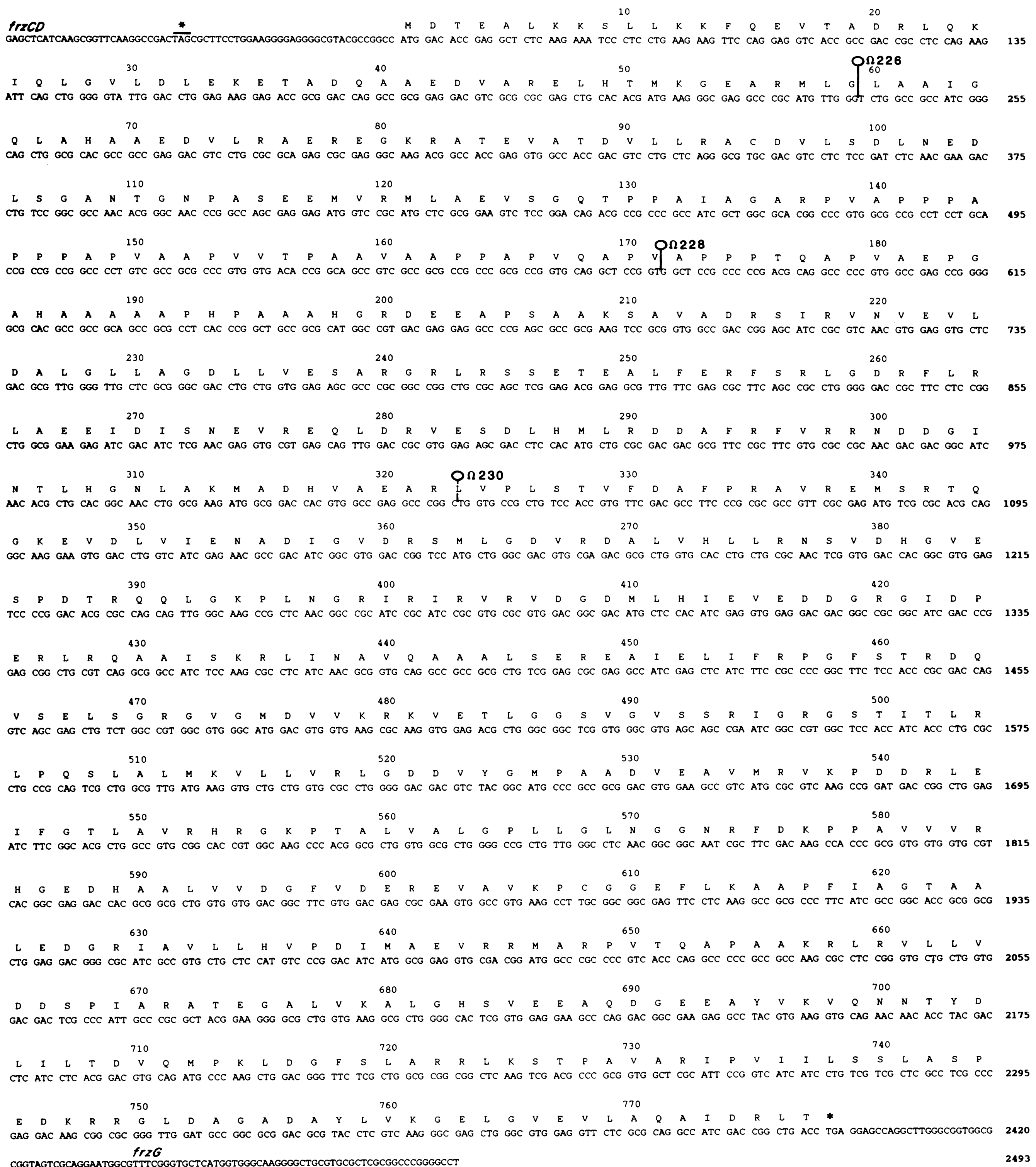

FIG. 2. Nucleotide sequence and predicted amino acid sequence of *frzE*. The left end of each Tn*5* insertion is identified by ⚲. Stop codon for *frzCD* and start codon of *frzG* are identified by * above the coding region.

thought to be involved with autophosphorylation and possibly multimerization. The C terminus controls phosphorylation in response to interactions with the MCPs and/or CheW.

The crystal structure of CheY was recently determined (30). CheY has a doubly wound five-stranded α/β structure. Phosphorylation occurs in a highly conserved acidic pocket at Asp-57 (28). Also conserved in the response-regulator protein family are multiple short stretches of hydrophobic residues thought important in maintaining secondary and tertiary structures. Phosphorylation is thought to cause a conformational change in CheY that allows it to interact directly with the switch mechanism of the flagellar basal body (16).

Except for a short stretch of amino acids that corresponds to the putative histidine phosphorylation site, there are no regions of sequence similarity between FrzE and CheA in the N-terminal 300 residues of FrzE. The central region of FrzE is conserved with much of the central and all of the C-terminal domains of CheA. The remaining 124 amino acid residues at

FIG. 3. Sequence alignment of *M. xanthus* FrzE to *S. typhimurium* CheA. Regions of amino acid identity are boxed. Numbering starts at the N terminus of each protein. The conserved histidine phosphorylation site is indicated by an arrow.

the C terminus of FrzE constitute a domain homologous to the entire CheY protein. All residues that are highly conserved in this family are maintained in FrzE. These data allow us to speculate that FrzE is a multidomain/multifunctional protein that functions as a second messenger linking FrzCD to the gliding motor. The conservation of functional domains between FrzE and CheA suggests common mechanisms of action. FrzE might interact with FrzA and/or FrzCD at its homologous input domain to receive a stimulus. Genetic data actually supports this finding (7). For example, *frzD* mutants are dominant to the wild-type allele and result in hyperreversal of cell movement. However, *frzD frzE* double mutants show the nonreversing Frz phenotype. This shows that *frzD* [which encodes a truncated FrzCD polypeptide (6)] depends on FrzE for its function. FrzE most likely autophosphorylates and subsequently transfers its phosphate to its C-terminal domain. Preliminary experiments indicate that FrzE is phosphorylated on an aspartate residue (data not shown). In analogy to phospho-CheY, phospho-FrzE might interact directly with the switch mechanism of the gliding motor and cause the organism to reverse direction.

The domain thought responsible for the interactions between CheA and CheY is not conserved between FrzE and CheA. How might the histidine kinase domain transfer phosphate to the response regulator domain of FrzE? The (alanine plus proline)-rich segment in the N-terminal region of FrzE might provide a clue to one possible mechanism of intramolecular phosphotransfer. The (alanine plus proline)-rich regions of pyruvate dehydrogenase were originally identified because of their striking ^{1}H NMR spectra (24), which indicated that these domains are exposed to solvent and are conformationally flexible. In the pyruvate dehydrogenase multienzyme complex, substrates are shuttled between physically separate active sites with swinging lipoyl arms (31). It has been proposed that the flexible (alanine plus proline)-rich region of the dihydrolipoamide acetyltransferase subunits facilitate this movement (24). A similar mechanism might be employed by FrzE. Its (alanine plus proline)-rich domain might constitute an interdomain hinge to allow phosphotransfer between spatially separate domains.

In summary, FrzE is unusual because it contains both components of the two-component regulatory proteins on

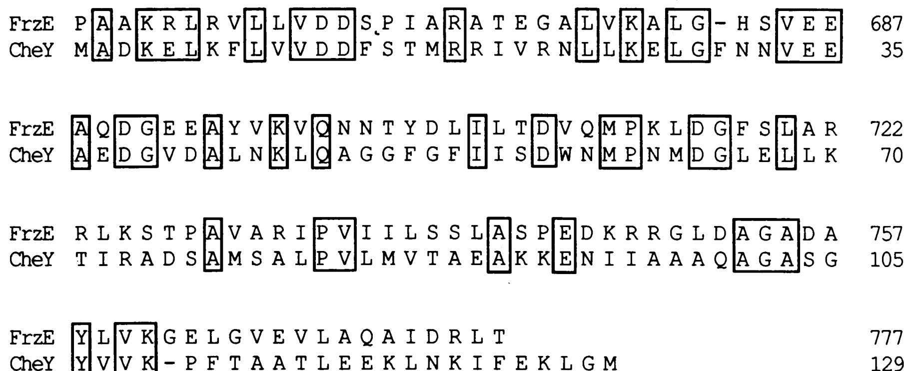

FIG. 4. Sequence alignment of *M. xanthus* FrzE to *S. typhimurium* CheY. Regions of amino acid identity are boxed. Numbering starts at the N terminus of each protein.

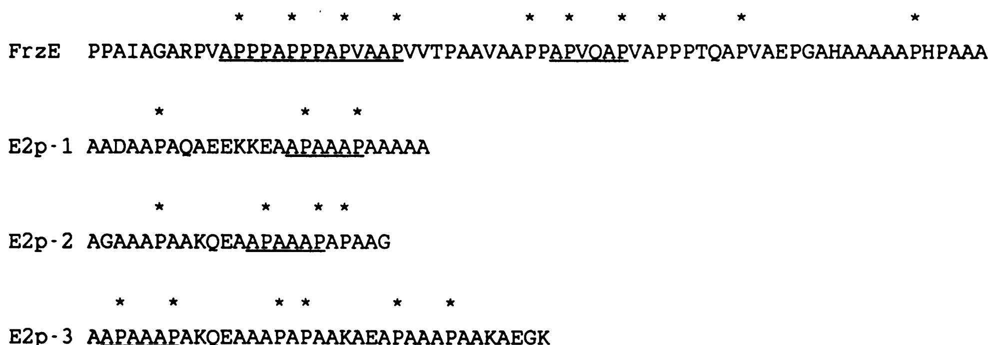

FIG. 5. Comparison of the (alanine plus proline)-rich segments from *M. xanthus* FrzE and the *E. coli* pyruvate dehydrogenase multienzyme complex. Proline residues directly preceded by an alanine are noted by *; Ala-Pro-Xaa-Xaa-Ala-Pro motifs are underlined.

one polypeptide. A similar motif is observed for the VirA protein of *Bordetella pertussis*, which functions as a transcriptional activator of virulence traits (32). We do not know whether there is any significance to this combined motif. The data presented in this paper add support to the hypothesis that the *frz* genes constitute a sensory transduction pathway that controls the frequency at which cells reverse their gliding direction. Immunologic blot analyses show that FrzE is present in both vegetative and developmental cells (unpublished work). We propose that FrzE is a second messenger that relays information between the signaling protein FrzCD and the gliding motor. Further investigation of FrzE may provide insight into the gliding apparatus.

We thank Mark McBride, Kathleen O'Connor, Kathy Trudeau, and Robin Weinberg for helpful discussions. We also thank the Animal Support Facility at the University of California at Berkeley for assistance with animal procedures. This research was supported by Public Health Service Grant GM 20509 from the National Institutes of Health and Grant DMB-8820799 from the National Science Foundation. W.R.M. was the recipient of National Research Service Award Traineeship GM 07232-12.

1. Rosenberg, E. (1984) *Myxobacteria: Development and Cell Interactions* (Springer, New York).
2. Zusman, D. R. (1984) *Q. Rev. Biol.* **59**, 119–138.
3. Morrison, C. E. & Zusman, D. R. (1979) *J. Bacteriol.* **140**, 1036–1042.
4. Zusman, D. R. (1982) *J. Bacteriol.* **150**, 1430–1437.
5. Blackhart, B. D. & Zusman, D. R. (1985) *Mol. Gen. Genet.* **198**, 243–254.
6. McCleary, W. R., McBride, M. J. & Zusman, D. R. (1990) *J. Bacteriol.*, in press.
7. Blackhart, B. D. & Zusman, D. R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8767–8770.
8. Weinberg, R. A. & Zusman, D. R. (1989) *J. Bacteriol.* **171**, 6174–6186.
9. McBride, M. J., Weinberg, R. A. & Zusman, D. R. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 424–428.
10. Koshland, D. E., Jr. (1981) *Annu. Rev. Biochem.* **50**, 765–782.
11. Macnab, R. M. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, eds. Neidhart, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E. (Am. Soc. Microbiol., Washington, DC), pp. 732–759.
12. Stewart, R. C. & Dahlquist, F. W. (1987) *Chem. Rev.* **87**, 997–1025.
13. Liu, J. & Parkinson, J. S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8703–8707.
14. Sanders, D. A., Mendez, B. & Koshland, D. E., Jr. (1989) *J. Bacteriol.* **171**, 6271–6278.
15. Hess, J. F., Oosawa, K., Matsumura, P. & Simon, M. I. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7609–7613.
16. Bourret, R. B., Hess, J. F. & Simon, M. I. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 41–45.
17. Lupas, A. & Stock, J. (1989) *J. Biol. Chem.* **264**, 17337–17342.
18. Dworkin, M. (1962) *J. Bacteriol.* **84**, 250–257.
19. Blackhart, B. D. & Zusman, D. R. (1986) *J. Bacteriol.* **166**, 673–678.
20. Bibb, M. J., Findlay, P. R. & Johnson, M. W. (1984) *Gene* **30**, 157–166.
21. Mesbah, M., Premachandran, U. & Whitman, W. B. (1989) *Int. J. Syst. Bacteriol.* **39**, 150–167.
22. Stock, A., Chen, T., Welsh, D. & Stock, J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1403–1407.
23. Stock, A., Koshland, D. E., Jr., & Stock, J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7989–7993.
24. Radford, S. E., Laue, E. D., Perham, R. N., Martin, S. R. & Appella, E. (1989) *J. Biol. Chem.* **264**, 767–775.
25. Stock, J. B., Ninfa, J. J. & Stock, A. M. (1989) *Microbiol. Rev.* **53**, 450–490.
26. Keener, J. & Kustu, S. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4976–4980.
27. Forst, S., Delgado, J. & Inouye, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6052–6056.
28. Sanders, S. A., Gillece-Castro, B. L., Stock, A. M., Burlingame, A. L. & Koshland, D. E., Jr. (1989) *J. Biol. Chem.* **264**, 21770–21778.
29. Oosawa, K., Hess, J. F. & Simon, M. I. (1988) *Cell* **53**, 89–96.
30. Stock, A. M., Mottonen, J. M., Stock, J. B. & Schutt, C. E. (1989) *Nature (London)* **337**, 745–749.
31. Miles, J. S., Guest, J. R., Radford, S. E. & Perham, R. N. (1988) *J. Mol. Biol.* **202**, 97–106.
32. Stibitz, S., Aaronson, W., Monack, D. & Falkow, S. (1989) *Nature (London)* **338**, 266–269.