

Thermostable Chemotaxis Proteins from the Hyperthermophilic Bacterium *Thermotoga maritima*

RONALD V. SWANSON,[†] M. GERMANA SANNA, AND MELVIN I. SIMON*

Division of Biology, California Institute of Technology, Pasadena, California 91125

Received 25 August 1995/Accepted 27 October 1995

An expressed sequence tag homologous to *cheA* was previously isolated by random sequencing of *Thermotoga maritima* cDNA clones (C. W. Kim, P. Markiewicz, J. J. Lee, C. F. Schierle, and J. H. Miller, J. Mol. Biol. 231: 960–981, 1993). Oligonucleotides complementary to this sequence tag were synthesized and used to identify a clone from a *T. maritima* λ library by using PCR. Two partially overlapping restriction fragments were subcloned from the λ clone and sequenced. The resulting 5,251-bp sequence contained five open reading frames, including *cheA*, *cheW*, and *cheY*. In addition to the chemotaxis genes, the fragment also encodes a putative protein isoaspartyl methyltransferase and an open reading frame of unknown function. Both the *cheW* and *cheY* genes were individually cloned into inducible *Escherichia coli* expression vectors. Upon induction, both proteins were synthesized at high levels. *T. maritima* CheW and CheY were both soluble and were easily purified from the bulk of the endogenous *E. coli* protein by heat treatment at 80°C for 10 min. CheY prepared in this way was shown to be active by the demonstration of Mg²⁺-dependent autophosphorylation with [³²P]acetyl phosphate. In *E. coli*, CheW mediates the physical coupling of the receptors to the kinase CheA. The availability of a thermostable homolog of CheW opens the possibility of structural characterization of this small coupling protein, which is among the least well characterized proteins in the bacterial chemotaxis signal transduction pathway.

Thermotoga maritima is a hyperthermophilic eubacterial species with an optimal growth temperature of 80°C and a maximal growth temperature of 90°C (18). A number of proteins have been purified directly from *T. maritima*, and all have proven to be thermostable (5, 10, 11, 14, 22, 37, 38). Although *T. maritima* has not yet been extensively characterized genetically, a large number of expressed sequence tags and genomic sequence tags which have been obtained from this organism (19) provide the means to rapidly isolate genes for thermostable homologs of many known proteins. These thermostable homologs may prove useful in the structural characterization of proteins where mesophilic counterparts have proven intractable. Furthermore, the opportunity to isolate a wide variety of *Thermotoga* proteins will advance the study of the structural basis of thermostability. In their initial random sequencing project, Kim et al. (19) identified 52 clones which exhibit significant similarity to previously identified proteins. Among those matches identified was a clone similar to CheA, the histidine kinase involved in chemotaxis.

The bacterial chemotaxis system has been most thoroughly characterized in the enteric bacteria *Escherichia coli* and *Salmonella typhimurium*. In these organisms, a small family of membrane-bound receptors modulates the activity of a kinase, CheA, in complexes with a small coupling protein, CheW. CheA transfers phosphoryl groups to CheY, a response regulator protein which functions as a soluble transducer and interacts with the flagellar switch. Two proteins, CheR and CheB, play antagonistic roles in the adaptation of the receptors. Between the receptors and the flagellar switch, six chemotaxis proteins comprise the signal transduction pathway.

Chemotaxis genes homologous to those found in *E. coli* and *S. typhimurium* have been isolated from several other bacteria (4, 16, 40). Although significant differences may exist in the functioning of chemotactic signal transduction pathways in different bacteria, the main components are conserved and probably perform analogous functions.

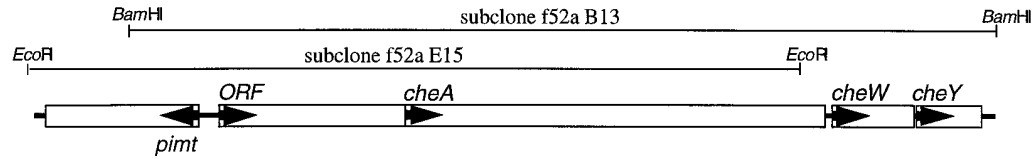
Of the six Che proteins, three-dimensional structural data are available for only CheY (26, 32). Several of the *E. coli* chemotaxis proteins have proven recalcitrant to structural characterization because of difficulty in crystallization and/or problems with solubility at high concentration. In many cases, it may be necessary to derive structures from other sources in order to augment our understanding of chemotaxis at a structural level. Although one cannot predict a priori that an organism will yield a homolog of a protein which will be amenable to structural characterization, thermophiles may have particular advantages. Thermostable proteins are easy to isolate when heterologously expressed in *E. coli* because thermal denaturation can be used as a purification step. For nuclear magnetic resonance (NMR) studies, the ability to acquire spectra at higher temperatures means that larger proteins are accessible to analysis. For crystallization, thermostable proteins may prove more rigid at conventionally used temperatures and therefore may yield higher quality diffraction patterns.

The availability of sequence tag information for a number of *Thermotoga* genes coupled with the likelihood that genes in a pathway are cotranscribed makes cloning and expression to obtain thermostable homologs an easily implemented strategy. In this study, we describe the use of sequence tag information to clone the complete *cheA* gene from *T. maritima*. During sequencing, open reading frames (ORFs) corresponding to *cheW* and *cheY* were found downstream of *cheA*. *Thermotoga* CheW and CheY were expressed in *E. coli*, and the recombinant proteins were found to be thermostable. This result validates the strategy and in the case of CheW provides a path to

* Corresponding author. Mailing address: Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125. Phone: (818) 395-3944. Fax: (818) 796-7066.

[†] Present address: Recombinant BioCatalysis Inc., Sharon Hill, PA 19079.

A.



B.

```

1  GAATTCACCATGCTGGTGGTCTGACTACAGTGAAGAATAGCAGAGAAACGGGACGGGTTATAGAAATAGAGAGCGGTCAATAGAACAGGATATATGTGAAAGAGCTGTGTAACCCAC 120
121 ATGAGCATGAGATACGAAAGCTCTGTAGCGCAGCTTTCCCACTGAGTTAAAGGTTCTCTATCTCTGGAGCGTCCCGTGTAGTTCATCTCTGTCATCGAGAACTCTACCACGCC 240
241 GTTTGGCCAGCGTAGTAGACGTTCCATCAATCTCTGTGAGTCTCTGTGAGGAGATCCACAGCTCCACGAGATATGAGAACGTACGAGTAAAGTCTCTGTGTAATGAAATTC 360
361 TCTCAAAAGCTTTCTGTCTCTCGAAGATTTCCGAGATTGCCCGCTGCCGTTATGAACCTGGTTTCCAGTTGTAGTTTCCACAGATACGGATCCTTTTGTGAACAGGAAAGC 480
481 GGGCTGTCTCTTGAAGTTTCAAAATGATAGGTACTATCACCTCCCACTCTCTGTAGTTGAGTGAACACGCTCTCCGCGACCTCATCCACACGACTGTACGAGATACATCGTA 600
601 CGGGGAAATCCCGGAAACCGGTAGTATCCATCGCCACAGACAAAATAACGTTCTCTATTTCCAGGCGTTCCACGTTCTCTTTCGCGATCTCGCAGATTTTCCGTAATTTCCACGGA 720
721 GACCACAGGCGCTTCTCACCCACACTCTGCTCATCAGCGACGGTTGTAACCGGTTCTCCGCTATTTTCGAGGACTCTCATTCCTTTGTCCAGACCCCACTCCATTAACACGCC 840
841 CATCAAGAGGGCTGACTGGAAGTCTGTATTTCTTCCATCATCGTAGGAACACAGACGATATCTCTGTAACGTAAGAGTGGGTAGGACTCTCTCAGGAACCTCTACGGGGTAT 960

1081 TTCCAGAAAGCCTTCGCGATGTGATCGTCCACCCGTACTTTTTAAGATCCAAAGAGCTTTTCTCTATAACATCCCGCTGTGTGTATCATATCACTGAGGAGGATCGTATGAAG 1080
1081 AAGCAGCGATCATCACTATCGGAAGTGAAGTCTGGAAGGTTTATGATCTCAACAAACACGACAGTTTCTCTGCCAGGAATTGAAGAACTCGGGATATAGAGTTGTCAAGGTATCAACG 1200
1201 GTAGGAGATGATTTGATCTCCATCTCAGAAAGAGTGAAGACCTCTCTCAAGGTGATCTACTCATCTCACAGGAGGACTCGGCCACACAGGACGATCTCACCCGGATGCACTC 1320
1321 CAAAAGTCTGAACAGGATCTCAAACTCAATGAAGAACTCTCTCGAAGATAAAGAGAGATAAAAAATACCATTTCTGAAATACCGCAGAACATTGAAGACAGGCTTTAGTAATA 1440
1441 GACGTCGCGAAGTTCTGCAACACCTGTGGGAAGTCTCTCAGGACAACTCTCTGATGTGGGAAAAATCGTGATTCTCTCGCGGACCTCCAGAGAGCTGATCCCATGTTCAAC 1560
1561 GCACGAAAGATCGATTGAGAACTCCAGATGCCCTGTATCAGGTGTGTTTGAATACTACAGCATTTCCGGAAGCGCTTCTCGAAGATCTTTTGAAGATATTTCTTACTCCCAAGACATA 1680

1681 GTGGAAGTAGCCACCATGCGCATCATGTAGAAGAGTACGACTCAGATTGACCACCATATGAAGAACAAAGATATCTGGACGAGATGGTGAAGAAAGATTTCTGATAGAGCTGGCGAA 1800
1801 CACCTCTACGGTGTAAACGACGAAAGATGGAAGAGTCTGGTTAGACTCTGAAAGACAGGAAAAACGCTCGCGCTTGCAATCTGCAGGCTGGCATGTCTTCATCTTTGGTG 1920
1921 GTGACGATGCGCGGGGCTTCAGAGGTTTTCATAGGAGGTGTGGTACGCTACAGCAACGATCTGAAAAACACATCTCGGTGTGTAGAGAGGACATCTCAAAAACACGCGGTGGTGAAG 2040
2041 GAAGATGTGTACAGGAGATGACGAGGGGCTAAAAAATCAGCGGTGAGATATCTGCGCTCTATCTCCGGAATAGCAGGCGCATCAGGTGGTACTCCGAGAAACCCCTAGGTACA 2160
2161 GTTTTCATAGATATCTTTGAACATGAACACATTAATGAGGTATATTTTACAGCGCATGAAACATGATAAGAACAGATCCGCTATGATGGCGCTGGAAACCTCAGAAAAATCTTG 2280
2281 AAGGGCGTGAAGATGATGGAAGAATATCTCGGAGTGTGTTGCGATGAGACAAAGATACCTTCAAAATCTGAACGATACCTTCTCGAATTGGAGAAAAATCCCGAAGATATGG 2400
2401 AACTCATAAAGAGCGTTTCAAGGCTCTTCAATACCTGAAAGGAATGGCAGGTACAATGGGTTTTCAGATGCGGAAGCTCTGCCATACCTTGAAGAACTCTCGATTAAGCAGAA 2520
2521 ACAGGAGATAAAGATACTTCGATCTCTCGATAAGATCTTCGCGGGGTCGATATGATACACAGAAATGTTGATAAGATCTCTCCGAGGGAAGTACGACGATCGGAGAAACATAG 2640
2641 ACCTGTCTCCGACCACTTAAAGGCTTGCATCTCGGAAAGAGAGCTTCAAGAAATCAAAATGAAACGAAACAAAGGTGAGGAAGAACACAAAGGAATCAACAGCAATG 2760
2761 AAGAGCTGTAGTCTTCTCGAAGAGTGTCCACGCTTCTCAGGAAGCGAGAAACAGGTTTCAAAACGTTTATATTAAGTGATTTCTCAAGAGGAACGCGATGTAATCCGCCA 2880
2881 GGATTTACCTCTTTTTCACAGCTTGAAGAACTGAAGTGTGAAGTGTGAGAACGATTTCTCTCGTTGAAGAGATAGAAGAGAGAAATCTGAAACGAACTGGAACCTCTCTGTAATCT 3000
3001 CCCCTGTGATCTGGAGAACTCTCTGAAGCTCTCTCAAGCATCGCGGATATAGAGAGGTTATATATAAAGAAATACCCGCTCACCAGAAATCGAGGGCTGAGAAAGAACCCGAGA 3120
3121 AAGAGAGAAATCTGAAAGAACTGAGGAAAGCGCAAGAAAGAAAGGTTATTTGCAACAGCTCAGGATAGATATAGAGAACTGGGCAATTTGATGATTTGATGGGAACTGTGTA 3240
3241 TCCGAGAGAGCAGAAATCTGGAAGCTCAAGAAATACAACTAAAGAACTGGATGAGATTTGTCTCATCTCAGCAGGATCACTTAGACCTTCAAGATTTGTGATGAGATCAGAA 3360
3361 TGGTTCCATCTCTCTTTTTCACAGATTTCCCTCGAATGGTGAAGACCTTCCCAAAAGAGATGAACAAAGAGTGAATTTCTCATGAGAGGAGAGACAGACGAGACGTT 3480
3481 TCGTTGAAGAAATTTGGCAACCCCTGCTCATCTCTGAGAAACGCAATGACACCGGTATAGAACCCAAAGAGAAAGCAATAGCCAAAGGAAACCCCCCATTTGGAACCTCATTTCTCT 3600
3601 CGGCACGTCCAGGAGAAACACGTTGGTAAATAGAGTCAAGATGACGGAAGGGGTATAGACAGGAAAGAAATCATCAGAAAGCCATAGAAAGGGACTCATAGATGAATCAAGGGCG 3720
3721 CTACCTCTCTGATCAGGAGATTTGAACTTCTCTCTGTTCCGGGATCTCCCAAAAGAGAAAGTCTCAGAAAGTCTCCGGAAGAGCGGTGGGAATGGATGTGTTGAAATATGGTGG 3840
3841 AATCTTTGAATGTTAGCATAGCATAGAGAGCGAGAGAAATGAAGAAACAAAGTTTACGATAGACTTACCGCTCACTCTGCGCATCATTCAGGCGCTCTCTGCTCAAGATCAACATCTCG 3960
3961 TCTACGCGATCCGATAGCGAACATAGACACAATCTCAGCCTTTCAAAAGAGGATATCAAGAGTTTCAGGACAGAGATGTGATGATCATAGAGAGAGAGTGAATCCCGTTTACCGTCT 4080
4081 TGTGGGAAGTGTCTCAATAGAGCACAAGAGAACTGGAGGAGATGGAAGCGTTATTGTGAGGGTAGGAAACAGGAAGTACGGTATCTCGGTAGACGATCTCTCGGTACGAGCAGATA 4200
4201 TCGTGATAAATCTCTTGAAGAGTGTCTCTGAGGTGAAGAAATCAGCGGACGAGTATCTCGGTGATGTAGTATAGCGCTGATATCAACGCTCCGCGATTGTATAACGGGGGT 4320
4321 GAAACATTTGGCGGATGCTTTGAAGAGTTCGAAGTTCTATCTTTGAAATAGATGAACAGGCACTGGCTTTGATGTCGACACATAGAAATGGTGTGAGAAATCCGATATACACC 4440
4441 TGTGCCAAATCAAGACACTTCGTAGAGGGTGTGATCAATTAAGAGGAAAGAAATCATTTCTGTGTTGAACCTTTCGGAAGATTTCTGGTATTTCAATTCGACGAGCAAAAGATGAAGATAT 4560
4561 AATGTGCGCAGAACGAAGGATGTAGAAGTGGGATTTCTTTGATAGAGTCTCGGAGGCTCTGAGAAATACCGAGAAATCAGCTGATTTGACAAATGTCTCCGATAAATTCGGTAAAAA 4680
4681 ATCAAGGGCTCTTGTAAAAACCGACGAGGCTCATCATCTATCTGGATATCGACAGATATCGAAGAAATACCGTTAAGGAGGGTGTGTAAGATGGGAAAGAGATTTTGATAGT 4800
4801 CGATGATGACGCTTCATGAGAGATGTTGTAAGATATCATCAAAAGCGGATGACGAGTTTGCAGGGGAGCAACAAACGCTGTGAAGCCGTGAAATAATCAAGGAATCAAAAC 4920
4921 GGACATCTGACGATGACATCACCATGCGCGAGATGAACGGTATCGATCGATCAAGAGATATCGAGATGACGACCAACGCAAGATCATCTGTCGACGTGCGACGAGCAGCAGG 5040
5041 AATGTGATAGACATTAAGAGCTTGAAGAGTGTGATTTGTAAGACCTTTCAGGCTTTCAGGCTGTAGAGGCTCTCAACAGGTTTTCGAAGTGAATCCCATGAGTGGGATTTCT 5160
5161 GCAGTTTCTTCTGGCTTTTCGAATAGTCTCTCTTCTTCTCTGTTTACTATTTTGTGAAGCGGGAACTTCATTTCAAAAGGATGCC 5251

```

FIG. 1. (A) Organization of the *cheAWY* region of the *T. maritima* chromosome. Sequences were obtained from subclones of the two restriction fragments shown. Five ORFs are present with the orientations indicated. (B) Nucleotide sequence of the *cheAWY* region. The 5,251-bp sequence was determined on both strands. Translations of four of the five ORFs present are displayed below the DNA sequence. Stop codons are indicated with asterisks. Translation of the fifth ORF encoding a putative protein isoaspartyl methyltransferase, *pimt*, which is transcribed from the opposite strand is not shown.

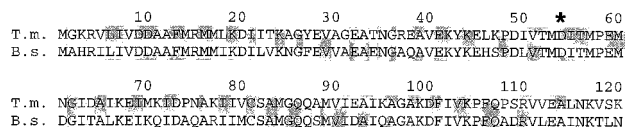


FIG. 2. Comparison of the predicted amino acid sequence of CheY from *T. maritima* (T.m.) with the CheY sequence from *B. subtilis* (B.s.) (3). CheY proteins from these two organisms are remarkably conserved with 85 identical residues, shown in gray, over a total length of 120 amino acids. The putative site of phosphorylation, Asp-54, is denoted with an asterisk. This site has been mutated in the *B. subtilis* CheY and has been shown to be essential for chemotaxis (2). In contrast to the high level of identity that these proteins have with each other, the *T. maritima* and *B. subtilis* CheY proteins show only 29 and 36% identity, respectively, to CheY from *E. coli*.

future structural understanding which bypasses reliance on the heretofore intractable *E. coli* protein.

MATERIALS AND METHODS

Cloning and sequencing. Two oligonucleotides, T102 (5'-AATGGTGAGAGACCTTGCCA3') and T103 (5'-GTTTCCCTCGTACGTGCCGAG3'), were synthesized on the basis of a sequence tag with similarity to CheA, ESTMX-138 (19). The sequences were chosen such that PCR amplification (28) using the oligonucleotides as primers would create an easily detectable product (220 bp). A *Sau*3A *T. maritima* λ library (19) was plated on *E. coli* LE392. One hundred twenty plaques were picked and individually eluted into 100 μ l of SM (0.1 M NaCl, 50 mM Tris (pH 7.5), 0.02% gelatin, 10 mM MgSO₄) at 37°C for 1 h, 10 μ l from each plaque was pooled into 12 groups of 10, and 7.5 μ l from each of these pools was amplified with primers T102 and T103, using 25 iterations of the following reaction cycle: 1 min at 94°C, 1 min at 54°C, and 1 min at 72°C. Each reaction mix consisted of 25 pmol of each primer, 200 μ M each deoxynucleoside triphosphate, 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.01% gelatin, and 1.25 U of AmpliTaq DNA polymerase in a total volume of 25 μ l overlaid with mineral oil. A single positive pool was identified by agarose gel electrophoresis of the reaction products, and an individual reaction was run on the each member of that pool. Phage from the positive isolate were replated at low density, and single plaques were isolated and retested by PCR to yield the final positive clone, λ F52a. DNA was isolated from a large-scale preparation of λ F52a and was digested with *Eco*RI and *Bam*HI. These fragments were subcloned into dephosphorylated pBluescript KS+ and screened for inserts, using 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal). White colonies were screened for subclones containing *cheA* by PCR with the T102 and T103 primers.

All sequencing reactions were performed by using either the Applied Biosystems dye primer or dye terminator cycle sequencing kit and a model 373A automated DNA sequencer. Sequences were aligned and edited by using the Sequencher program (Genecodes, Ann Arbor, Mich.). Database searches were performed with the program BLASTX (1), and final alignments were constructed by using FASTA and PILEUP from the Genetics Computer Group package of sequence analysis software.

Protein expression and purification. *E. coli* K0641/*recA* (Δ *cheY*) (8) was used as a host for expression of *Thermotoga* CheY. pTMY was constructed by amplifying the *cheY* gene from 100 ng of pF52aB13 with *Taq* polymerase in a 50- μ l reaction, using the conditions described above and the primers TMY301 (5'-CCGATCTTCACTTCGAAACCTTGTGAG3') and TMY501 (5'-CCGGAATTCATTAGAGGAGAAATTAACATATGGGAAAGAGAGTTTGATA3'). The PCR product was digested with *Eco*RI and *Bgl*II and cloned into similarly digested pQE12 vector (Qiagen).

K0641/*recA*(pREP4, pTMY) was cultured at 37°C in LB and induced for 5 h with 1 mM isopropylthiogalactoside (IPTG) upon reaching an optical density at 600 nm of 0.4. Cells were harvested by centrifugation and resuspended in 50 mM Tris (pH 8.0)–5 mM MgCl₂–2 mM 2-mercaptoethanol–0.1 mg of lysozyme per ml. After brief sonication, the lysate was cleared by centrifugation at 87,000 \times g for 1 h. Thermolabile proteins were removed by placing the protein solution in an 80°C water bath for 10 min and then centrifuging it at 16,000 \times g for 20 min. The protein solution was then passed over a DEAE column equilibrated in 50 mM Tris (pH 7.5)–5 mM MgCl₂–2 mM 2-mercaptoethanol, and the flowthrough fractions containing *Thermotoga* CheY were collected and pooled.

A *Thermotoga cheW* expression vector was constructed by amplifying *cheW* from 100 ng of pF52aB13, using the primers TMW501 (5'-CCGAATTCATATGAAACATTTGGCGGATGCTTTG-3') and TMW32 (5'-CCGGAATTCGGAATTCCTTACACACCTCCTTAACGGT-3'). The amplification product was cut with *Eco*RI and cloned into pBS to give pBSTMW213. The insert was sequenced, and both pBSTMW213 and pCW/CY (a gift from F. W. Dahlquist) were cut with *Nde*I and *Hind*III. Restriction fragments corresponding to the pCW/CY vector and the *Thermotoga cheW* were gel purified, combined, ligated, and cloned to give pCW/TMW.

E. coli M15 (39) containing pREP4 [*neo*, *lacI*] (Qiagen) was used as a host for

expression of CheW. *Thermotoga* CheW was purified by essentially the same procedure as described for *Thermotoga* CheY except that cells were induced with IPTG for 15 h. The buffer used for lysis was 20 mM Tris (pH 8.0)–50 mM KCl–1 mM EDTA–0.5 mM phenylmethylsulfonyl fluoride–0.1 mg of lysozyme per ml. Similarly, the buffer used for DEAE chromatography was 20 mM Tris (pH 8)–50 mM KCl.

Phosphorylation assays. Protein phosphorylation and dephosphorylation assays were performed essentially as described previously (29). One hundred fifty picomoles of *E. coli* or *Thermotoga* CheY was incubated with 4 mM [³²P]acetyl phosphate in a total volume of 10 μ l for 2 min at different temperatures (4, 25, 50, 80, and 100°C). Reactions were stopped with the addition of sodium dodecyl sulfate (SDS) loading buffer containing 10 mM EDTA. Reactions were quantitated by running the products on an SDS–15% polyacrylamide gel and then exposing the gel overnight to a phosphor storage screen.

Autodephosphorylation rates were determined by incubating 150 pmol of CheY with 2 mM [³²P]acetyl phosphate at either 25 or 50°C. After 3 min, 60 mM unlabeled acetyl phosphate was added to the reaction, and dephosphorylation was monitored by removing aliquots to SDS loading buffer.

Nucleotide sequence accession number. The nucleotide sequence reported in this paper has been submitted to GenBank and assigned accession number U30501.

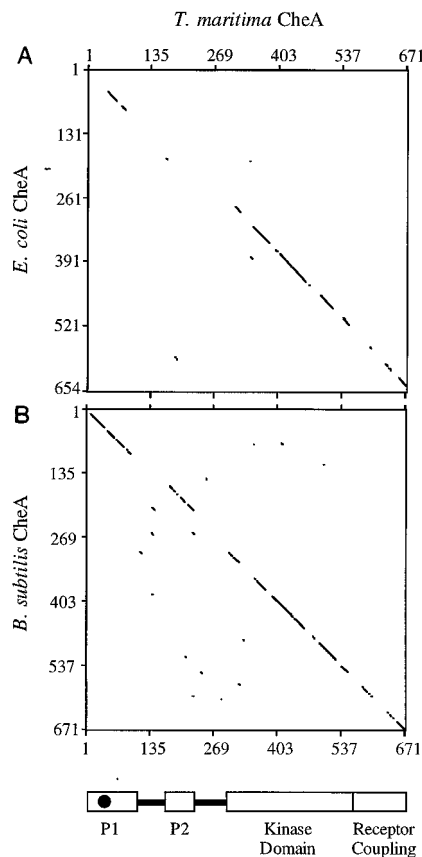


FIG. 3. Dot matrix homology plots of *T. maritima* CheA compared with *E. coli* CheA (20) (A) and *B. subtilis* CheA (13) (B). The diagonal lines in the plots represent regions of high similarity. A schematic diagram of CheA based on the *E. coli* protein is shown below the plots. The P1 domain contains the site of histidine autophosphorylation (black dot) and is capable of phosphotransfer to CheY (17). The P2 domain contains binding sites for CheY (34) and CheB (21). The kinase domain contains conserved sequences which define the two-component kinase family (27, 31). The receptor coupling domain contains sequences responsible for coupling receptor signalling to kinase activity (7). While the *Thermotoga* CheA is homologous to both *E. coli* and *B. subtilis* proteins in the kinase domain, homology to *E. coli* in the P1-P2 region is confined to the site of autophosphorylation. Homology to *B. subtilis* in the P1-P2 region is extensive and reflects the similarity of CheY from these organisms.

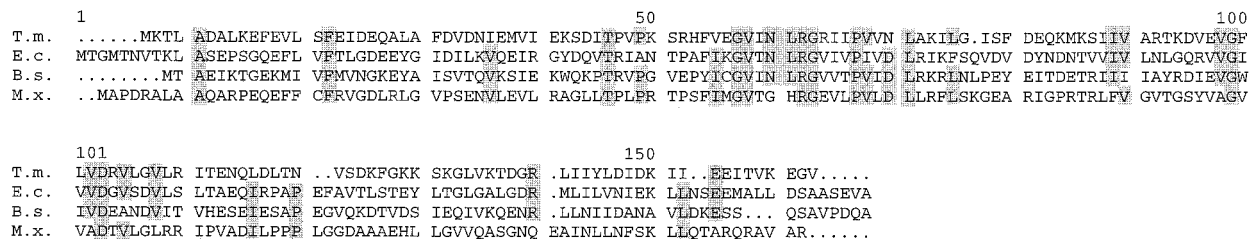


FIG. 4. Alignment of CheW sequences from four different organisms. Protein sequences from *T. maritima* (T.m.), *E. coli* (E.c.), *B. subtilis* (B.s.), and *M. xanthus* (M.x.) were aligned by using the program PILEUP. Residues conserved in at least three of the sequences are shown in gray.

RESULTS AND DISCUSSION

Two oligonucleotides were synthesized on the basis of a sequence tag from *T. maritima* which displayed homology to CheA (19). These oligonucleotides were used to identify a clone from a λ library containing the genomic copy of the *cheA* gene. Subclones were obtained from the λ clone, and sequencing of two overlapping fragments revealed the presence of five ORFs (Fig. 1A). Database searches performed with the BLASTX program (1) detected significant homologies to database sequences for all five ORFs. In addition to CheA, ORFs homologous to CheW and CheY were identified on the same coding strand (Fig. 1A). Immediately preceding and partially overlapping *cheA* (Fig. 1B) is an ORF which is similar to an ORF of unknown function previously identified in *E. coli* (12). Although the position of this ORF suggests that it is cotranscribed with *cheAWY*, the *E. coli* homolog has not been implicated in any chemotaxis or motility functions. A fifth ORF present upstream of these four genes and divergently transcribed displays strong homology to protein L-isoaspartyl methyltransferases from several sources. This type of enzyme has been shown to act in a repair pathway for damaged proteins (9). Because of the increase in rate of spontaneous deamidation associated with higher temperature, this type of activity would be even more important for hyperthermophilic organisms than for the mesophilic organisms from which these enzymes have been isolated previously. In *E. coli*, methylation of specific glutamyl residues in the chemotaxis receptors is responsible for adaptation to high levels of attractant. However, CheR the enzyme responsible for this activity bears no sequence similarity to the L-isoaspartyl transferases, and therefore it is likely that the proximity of this gene to the *cheAWY* operon in *T. maritima* is simply coincidental.

Relationships of the chemotaxis proteins from different sources. CheY displays the highest identity among the proteins identified in *T. maritima* to previously identified proteins from other organisms. *T. maritima* CheY is 70.8% identical to *Bacillus subtilis* CheY (Fig. 2), while it is only 29.1% identical to *E. coli* CheY. This difference may reflect a similarity in function to the *B. subtilis* chemotaxis system in contrast to the *E. coli* system. CheY apparently plays slightly different roles in these two model systems. In both systems, CheY is phosphorylated by CheA and acts as a soluble messenger to relay signals to the flagellar switch. In *E. coli*, interaction of phospho-CheY with the switch causes tumbling swimming by inducing clockwise flagellar rotation, whereas in *B. subtilis*, interaction of phospho-CheY with the switch apparently induces smooth swimming behavior (2, 3).

Although the *T. maritima* CheA is not overall significantly more similar to either of the *E. coli* or *B. subtilis* protein, the pattern of conserved regions reveals an interesting feature. While homology to the *E. coli* protein is localized to the central kinase domain, the *Thermotoga* protein displays significant

similarity to the *B. subtilis* CheA protein within two additional tracts near the amino terminus (Fig. 3). As originally observed by comparison of the closely related *E. coli* and *S. typhimurium* CheA sequences, these two regions constitute conserved domains connected to each other by an unconserved linker and in turn to the kinase domain by a second unconserved linker (27, 34). Construction and functional analysis of small fragments corresponding to these conserved regions revealed that the first, P1, maintains the ability to transfer phosphoryl groups to CheY (16, 33) and that the second, P2, confers the ability to stably bind to CheY (25, 33, 34) and to CheB (21). Thus, both P1 and P2 have been shown to interact with CheY. The conservation of the P1 and P2 regions between *B. subtilis* and *T. maritima* is consistent with the high degree of similarity exhibited by CheY and supports the role of these domains in protein-protein recognition.

CheW functions to couple the receptor signalling state to the kinase (6) and participates in the formation of receptor-kinase complexes (15, 30). CheW is not known to have any catalytic activity on its own. CheW is not homologous to any known group of proteins at the sequence level and therefore may represent a novel class of signal transduction coupling protein. The *T. maritima* CheW sequence is equally divergent from the other known CheW sequences: 25.3% identity to *E. coli*, 24.6% identity to *B. subtilis*, and 25.9% identity to *Myxococcus xanthus*. (Fig. 4), with conserved residues concentrated in the central region of the protein.

Autophosphorylation and thermostability of recombinant *T. maritima* CheY. The *Thermotoga cheY* sequence was amplified by PCR, the resulting fragment was cloned into an inducible expression vector, and the sequence of the insert was subsequently confirmed. Addition of IPTG to growing cultures resulted in the induction of a protein of the predicted size of *Thermotoga* CheY, approximately 13 kDa (Fig. 5). After lysis of the cells, *Thermotoga* CheY was found in the soluble fraction. Substantial purification was achieved simply with incubation at 80°C of the cleared lysate; remaining contaminants were removed by passing the heat-treated fraction over a DEAE column which does not bind *Thermotoga* CheY (Fig. 5).

Several representatives of the response regulator family of proteins have been shown to autophosphorylate with low-molecular-weight phosphodonors (24). Purified *Thermotoga* CheY was tested for the ability to autophosphorylate with [32 P] acetyl phosphate at four temperatures (Fig. 6). Like *E. coli* CheY, the *Thermotoga* protein is able to autophosphorylate. In addition, it displays significant activity at 80°C, whereas the *E. coli* protein is completely inactivated at this temperature. *Thermotoga* CheY may be stable and functional at even higher temperatures, but thermal destruction of the acetyl phosphate precludes the use of this assay as a functional test.

In addition to autophosphorylation, the response regulator proteins are capable of catalyzing their own dephosphoryla-

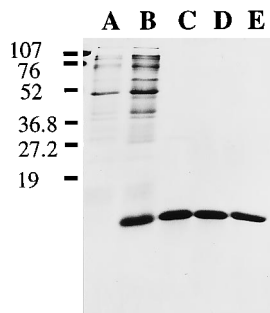


FIG. 5. Expression and purification of *T. maritima* CheY from *E. coli*. Samples represent uninduced *E. coli* M15(pREP4, pTMY) culture (lane A) and the culture after induction with IPTG (lane B), after treatment at 80°C and centrifugation (lane C), and after removal of contaminants with DEAE (lane D). The protein loaded in each lane represents the same proportion of the original culture volume. Purified *E. coli* CheY is shown for comparison (lane E). Sizes are indicated in kilodaltons.

tion. The autodephosphorylation rate of phosphorylated *Thermotoga* CheY was measured at 25 and 50°C by quenching an [32 P]acetyl phosphate labeling reaction with unlabeled acetyl phosphate and measuring the decay of the labeled protein (29) (Fig. 7). The autodephosphorylation of *Thermotoga* CheY is strongly temperature dependent; at 25°C the half-life of phospho-CheY is approximately 150 s, while at 50°C the half-life is approximately 25 s (Fig. 7). The ability to manipulate the half-life of CheY phosphate may prove useful in studying the phosphorylated state. An NMR study of the conformational changes which occur upon phosphorylation of *E. coli* CheY has recently appeared (23). The longer half-life of the phosphorylated form of *Thermotoga* CheY at 25°C may facilitate such studies in the future.

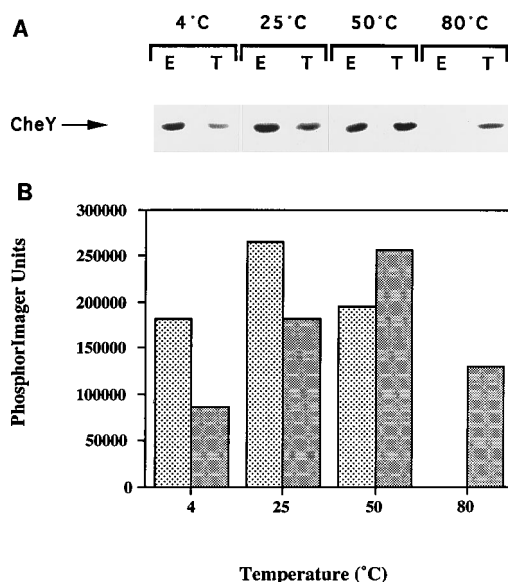


FIG. 6. *Thermotoga* CheY autophosphorylates by using acetyl phosphate as a phosphodonator. Shown is comparison of *Thermotoga* CheY and *E. coli* CheY autophosphorylation at four different temperatures. (A) Reaction products were electrophoresed after a 2-min incubation at the indicated temperatures. The labeled protein was quantitated with a PhosphorImager (B). There is no detectable phosphorylation of the *E. coli* CheY at 80°C, while the *Thermotoga* protein is still highly phosphorylated at this temperature. The use of this assay at elevated temperatures is not quantitative because of thermal destruction of acetyl phosphate.

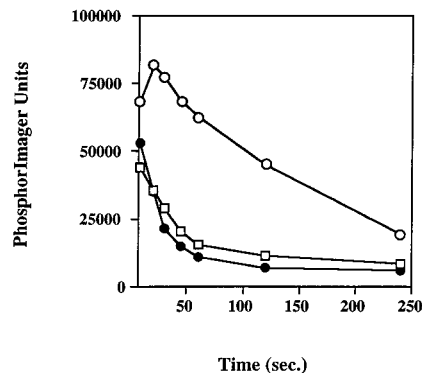


FIG. 7. Autodephosphorylation of *Thermotoga* CheY is temperature dependent. The autodephosphorylation rate of CheY was measured by allowing the protein to autophosphorylate with [32 P]acetyl phosphate and quenching the reaction with excess unlabeled acetyl phosphate. Aliquots were then removed, and the reaction was stopped by adding SDS-polyacrylamide gel electrophoresis loading buffer to measure the decay in labeled protein at the indicated times. ○, *Thermotoga* CheY at 25°C; □, *E. coli* CheY at 25°C; ●, *Thermotoga* CheY at 50°C.

Expression and purification of recombinant *T. maritima* CheW. *Thermotoga* CheW was cloned into an expression vector, and the protein was produced at high levels upon induction with IPTG (Fig. 8). CheW proved as easy to purify as CheY; again the major purification step consisted of the removal of thermolabile native *E. coli* proteins (Fig. 8). Although CheW does not possess enzymatic activity which would allow us to monitor its function at elevated temperatures, the fact that it is soluble during incubations at 80°C provides an indication of its thermostability.

Conclusion. Microbial genome sequence tagging projects of the type initiated by Kim et al. (19) provide rapid pathways to the isolation of homologs of interest to researchers in many fields. In particular, thermophilic bacteria and archaea can be expected to yield homologs of particular value to structural biologists. The *Thermotoga* CheY and CheW proteins have recently been crystallized (unpublished results). In addition, the structures of both proteins are being investigated by NMR spectroscopy. Of the large family of response regulator proteins, only two distinct response regulator domains have been characterized by NMR spectroscopy or crystallography (26, 32, 35). The most thoroughly characterized response regulator is

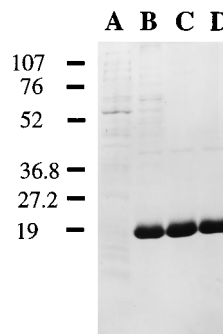


FIG. 8. Expression and purification of *T. maritima* CheW in *E. coli*. Uninduced *E. coli* M15(pREP4, pCW/TMW) culture (lane A) and the culture after induction with IPTG (lane B), after treatment at 80°C and centrifugation (lane C), and after removal of contaminants with DEAE (lane D). The protein loaded in each lane represents the same proportion of the original culture volume. Sizes are indicated in kilodaltons.

the *E. coli*/*S. typhimurium* CheY protein. *Thermotoga* CheY is 29% identical to the *E. coli* protein. Volz (36) has extensively analyzed the relationships of the response regulators to one another and has shown that on average, two response regulators are only 23% identical. Thus, by studying the *Thermotoga* CheY protein, we may not gain extensive insight into *E. coli* chemotaxis, but we will expand our comparative understanding of the structure and function of the response regulator family. In contrast, *Thermotoga* CheW promises to provide us with our first glimpse of the structure of a CheW protein from any source.

ACKNOWLEDGMENTS

We express our appreciation to Jeffrey Miller for sharing results prior to publication, to John Moore for supplying the *T. maritima* λ library, and to Nancy Chen, Steve Marsh, and Rick Colayco for DNA sequencing.

This work was supported by Public Health Service grant AI19296 (to M.I.S.). R.V.S. was supported by National Research Service Award Fellowship GM14767, and M.G.S. was supported by the Institute Pasteur-Cenci Bolognetti.

REFERENCES

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Bischoff, D. S., R. B. Bourret, M. L. Kirsch, and G. W. Ordal. 1993. Purification and characterization of *Bacillus subtilis* CheY. *Biochemistry* **32**:9256–9261.
- Bischoff, D. S., and G. W. Ordal. 1991. Sequence and characterization of *Bacillus subtilis* CheB, a homolog of *Escherichia coli* CheY, and its role in a different mechanism of chemotaxis. *J. Biol. Chem.* **266**:12301–12305.
- Bischoff, D. S., and G. W. Ordal. 1992. *Bacillus subtilis* chemotaxis: a deviation from the *Escherichia* paradigm. *Mol. Microbiol.* **6**:23–28.
- Blamey, J. M., and M. W. W. Adams. 1994. Characterization of an ancestral type of pyruvate ferredoxin oxidoreductase from the hyperthermophilic bacterium, *Thermotoga maritima*. *Biochemistry* **33**:1000–1007.
- Borkovich, K. A., N. Kaplan, J. F. Hess, and M. I. Simon. 1989. Transmembrane signal transduction in bacterial chemotaxis involves ligand-dependent activation of phosphate group transfer. *Proc. Natl. Acad. Sci. USA* **86**:1208–1212.
- Bourret, R. B., J. Davagnino, and M. I. Simon. 1993. The carboxy-terminal portion of the CheA kinase mediates regulation of autophosphorylation by transducer and CheW. *J. Bacteriol.* **175**:2097–2101.
- Bourret, R. B., J. F. Hess, and M. I. Simon. 1990. Conserved aspartate residues and phosphorylation in signal transduction by the chemotaxis protein CheY. *Proc. Natl. Acad. Sci. USA* **87**:41–45.
- Brennan, T. V., J. W. Anderson, Z. Jia, E. B. Waygood, and S. Clarke. 1994. Repair of spontaneously deamidated Hpr phosphocarrier protein catalyzed by the L-isoaspartyl-(D-aspartate) O-methyltransferase. *J. Biol. Chem.* **269**:24586–24595.
- Bronnenmeier, K., A. Kern, W. Liebl, and W. L. Staudenbauer. 1995. Purification of *Thermotoga maritima* enzymes for the degradation of cellulose materials. *Appl. Environ. Microbiol.* **61**:1399–1407.
- Brown, S. H., C. Sjöholm, and R. M. Kelly. 1993. Purification and characterization of a highly thermostable glucose isomerase produced by the extremely thermophilic eubacterium, *Thermotoga maritima*. *Biotechnol. Bioeng.* **41**:878–886.
- Cohen, S. P., H. Haechler, and S. B. Levy. 1993. Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. *J. Bacteriol.* **175**:1484–1492.
- Fuhrer, D. K., and G. W. Ordal. 1991. *Bacillus subtilis* CheN, a homolog of CheA, the central regulator of chemotaxis in *Escherichia coli*. *J. Bacteriol.* **173**:7443–7448.
- Gabelsberger, J., W. Liebl, and K. H. Schleifer. 1993. Purification and properties of recombinant β -glucosidase of the hyperthermophilic bacterium *Thermotoga maritima*. *Appl. Microbiol. Biotechnol.* **40**:44–52.
- Gegner, J. A., D. R. Graham, A. F. Roth, and F. W. Dahlquist. 1992. Assembly of an MCP receptor, CheW, and kinase CheA complex in the bacterial chemotaxis signal transduction pathway. *Cell* **70**:975–982.
- Greck, M., J. Platzter, V. Sourjik, and R. Schmidt. 1995. Analysis of a chemotaxis operon in *Rhizobium meliloti*. *Mol. Microbiol.* **15**:989–1000.
- Hess, J. F., R. B. Bourret, and M. I. Simon. 1988. Histidine phosphorylation and phosphoryl group transfer in bacterial chemotaxis. *Nature (London)* **336**:139–143.
- Huber, R., T. A. Langworthy, H. König, M. Thomm, C. R. Woese, U. B. Sleyter, and K. O. Stetter. 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Arch. Microbiol.* **144**:324–333.
- Kim, C. W., P. Markiewicz, J. J. Lee, C. F. Schierle, and J. H. Miller. 1993. Studies of the hyperthermophile *Thermotoga maritima* by random sequencing of cDNA and genomic libraries. *J. Mol. Biol.* **231**:960–981.
- Kofoid, E. C., and J. S. Parkinson. 1991. Tandem translation starts in the *cheA* locus of *Escherichia coli*. *J. Bacteriol.* **173**:2116–2119.
- Li, J., R. V. Swanson, M. I. Simon, and R. M. Weis. 1995. The response regulators CheB and CheY exhibit competitive binding to the kinase CheA. *Biochemistry* **34**:14626–14636.
- Liebel, W., R. Feil, J. Gabelsberger, J. Kellermann, and K. H. Schleifer. 1993. Purification and characterization of a novel thermostable 4- α -glucanotransferase of *Thermotoga maritima* cloned in *Escherichia coli*. *Eur. J. Biochem.* **207**:81–88.
- Lowry, D. L., A. F. Roth, A. P. B. Rupert, F. W. Dahlquist, F. J. Moy, P. J. Domaille, and P. Matsumura. 1994. Signal-transduction in chemotaxis, a propagating conformation change upon phosphorylation of CheY. *J. Biol. Chem.* **269**:26358–26362.
- Lukat, G. S., W. R. McCleary, A. M. Stock, and J. B. Stock. 1992. Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc. Natl. Acad. Sci. USA* **89**:718–722.
- 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.
- Moy, F. J., D. F. Lowry, P. Matsumura, F. W. Dahlquist, J. E. Krywko, and P. J. Domaille. 1994. Assignments secondary structure, global fold, and dynamics of chemotaxis Y protein using three- and four-dimensional heteronuclear (¹³C, ¹⁵N) NMR spectroscopy. *Biochemistry* **33**:10731–10742.
- Parkinson, J. S., and E. C. Kofoid. 1993. Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* **26**:71–112.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. *Science* **239**:487–489.
- Sanna, M. G., R. V. Swanson, R. B. Bourret, and M. I. Simon. 1995. Mutations in the chemotactic response regulator CheY that confer resistance to the phosphatase activity of CheZ. *Mol. Microbiol.* **15**:1069–1079.
- Schuster, S. C., R. V. Swanson, L. A. Alex, R. B. Bourret, and M. I. Simon. 1993. Assembly and function of a signal transduction complex monitored by surface plasmon resonance. *Nature (London)* **365**:343–347.
- Stock, A., D. E. Koshland, and J. Stock. 1985. Homologies between the *Salmonella typhimurium* CheY protein and proteins involved in the regulation of chemotaxis, membrane protein synthesis, and sporulation. *Proc. Natl. Acad. Sci. USA* **82**:7989–7993.
- Stock, A. M., E. Martinez-Hackert, B. F. Rasmussen, A. West, J. B. Stock, D. Ringe, and G. A. Petsko. 1993. Structure of the Mg²⁺ bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis. *Biochemistry* **32**:13376–13380.
- Swanson, R. V., D. F. Lowry, P. Matsumura, M. McEvoy, M. I. Simon, and F. W. Dahlquist. 1995. Localized perturbations in CheY structure monitored by NMR identify a CheA binding interface. *Nature Struct. Biol.* **2**:906–910.
- 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.
- Volkman, B. F., M. J. Nohaile, N. K. Amy, S. Kustu, and D. E. Wemmer. 1995. 3-dimensional solution structure of the N-terminal receiver domain of NtrC. *Biochemistry* **34**:1413–1424.
- Volz, K. 1993. Structural conservation in the CheY superfamily. *Biochemistry* **32**:11741–11753.
- Wrba, A., R. Jaenicke, R. Huber, and K. O. Stetter. 1990. Lactate dehydrogenase from the extreme thermophile *Thermotoga maritima*. *Eur. J. Biochem.* **188**:195–201.
- Wrba, A., A. Schweiger, V. Schultes, and R. Jaenicke. 1990. Extremely thermostable D-glyceraldehyde-3-phosphate dehydrogenase from the eubacterium *Thermotoga maritima*. *Biochemistry* **29**:7584–7592.
- Zamenhof, P. J., and M. Villarejo. 1972. Construction and properties of *Escherichia coli* strains exhibiting α -complementation of β -galactosidase fragments in vivo. *J. Bacteriol.* **110**:171–178.
- Zusman, D. R., and M. J. McBride. 1991. Sensory transduction in the gliding bacterium *Myxococcus xanthus*. *Mol. Microbiol.* **5**:2323–2329.