Thermostable chemotaxis proteins from the hyperthermophilic bacterium Thermotoga maritima.

R V Swanson, M G Sanna and M I Simon

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 proteins 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
FIG. 1. (A) Organization of the cheAUY 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 cheAUY 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.
Future structural understanding which bypasses reliance on the heretofore intractable *E. coli* protein.

**MATERIALS AND METHODS**

**Cloning and sequencing.** Two oligonucleotides, T102 (5′ ATTGTTGAGAGACA GCATTCAAGCCGATTTGCGAG3′) and T103 (5′ GTTCTCCCTGATACGGTGGCGAG3′), were synthesized on the basis of a sequence tag with similarity to CheA. ESTXM-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 466-bp DNA fragment from *T. maritima* T.m. library (19) was plated on *E. coli* LE392. One hundred twenty plaques were picked and individually eluted into 100 μL 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 μL from each plaque was pooled into 12 groups of 10, and 7.5 μL 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 μM each deoxynucleoside triphosphate, 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM MgCl2, 0.01% gelatin, and 1 U of AmpliTag DNA polymerase in a total volume of 25 μL 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 each member of that pool. Phage from the positive isolate were replated to yield the final positive clone, Ap52a. DNA was isolated from a large-scale preparation of single plaques were isolated and retested by PCR to yield the final positive clone, which was confirmed to be CheY from *T. maritima*.

**Protein expression and purification.** *E. coli* K0641::recA (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-μl reaction, using the conditions described above and the primers TMY301 (5′-CGAGATCTCATCAGAAACTTCGCGAG3′) and TMY501 (5′-CGCGGAATTCGGAGGAATATTGCGTGTATGAA3′). The PCR product was digested with EcoRI and BamHI. These fragments were subcloned into dephosphorylated pBluescript K+ 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.

**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]acetate phosphate in a total volume of 10 μl for 2 min at different temperatures (4, 25, 50, 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]acetate 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 U80501.
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 cheA/WY, 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 cheA/WY 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 tumbly 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 regions 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 phosphodonor (24). Purified *Thermotoga* CheY was tested for the ability to autophosphorylate with [32P]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-
The autodephosphorylation rate of phosphorylated Thermotoga CheY was measured at 25 and 50°C by quenching an [32P]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.

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
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 Colacoy 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