Protein Science (1998), 7:403–412. Cambridge University Press. Printed in the USA.

Crystal structures of CheY from *Thermotoga maritima* do not support conventional explanations for the structural basis of enhanced thermostability

KEN C. USHER,[1] AIDA FLOR A. DE LA CRUZ,[1,3] FREDERICK W. DAHLQUIST,[1] RONALD V. SWANSON,[2,4] MELVIN I. SIMON,[2] AND S. JAMES REMINGTON[1]

[1]Institute of Molecular Biology and Departments of Chemistry and Physics, University of Oregon, Eugene, Oregon 97403
[2]Division of Biology, California Institute of Technology, Pasadena, California 91125

(RECEIVED July 23, 1997; ACCEPTED October 30, 1997)

Abstract

The crystal structure of CheY protein from *Thermotoga maritima* has been determined in four crystal forms with and without Mg^{++} bound, at up to 1.9 Å resolution. Structural comparisons with CheY from *Escherichia coli* shows substantial similarity in their folds, with some concerted changes propagating away from the active site that suggest how phosphorylated CheY, a signal transduction protein in bacterial chemotaxis, is recognized by its targets. A highly conserved segment of the protein (the "γ-turn loop," residues 55–61), previously suggested to be a rigid recognition determinant, is for the first time seen in two alternative conformations in the different crystal structures. Although CheY from *Thermotoga* has much higher thermal stability than its mesophilic counterparts, comparison of structural features previously proposed to enhance thermostability such as hydrogen bonds, ion pairs, compactness, and hydrophobic surface burial would not suggest it to be so.

Keywords: chemotaxis, protein crystallography, signal transduction, thermostability

CheY, a signal transduction protein found in many bacteria, determines the direction of rotation of the flagellar motors, depending on its state of phosphorylation. It participates in several different protein–protein interactions that are critical to the relay of a sensory signal from cell-surface receptors to the flagellar motor. CheY accepts a phosphate group from phospho-CheA; its dephosphorylation rate is modulated by CheZ, and it only interacts with the flagellar motor proteins FliG and FliM when phosphorylated. In *Escherichia coli* CheY (ECY), data from NMR experiments (Lowry et al., 1994) and from mutational studies of CheY (Bourret et al., 1993) and its flagellar motor targets (Roman et al., 1992) suggest that the structural change recognized by other proteins is not the presence or absence of the phosphate group itself, but rather a long-range conformational change induced in the protein upon phosphorylation. The NMR studies have not yet established the three-dimensional structure of the phosphorylated state of CheY; however, and the changes recognized by targets of the switch protein remain obscure.

TMY is much more thermostable than ECY, with over 35 °C difference between their respective T_ms. Magnesium-free TMY reversibly (≥70% reversible) undergoes an unfolding transition centered at 95°, with an estimated ΔH^0 of 78 kcal/mol (at T_m, assuming two-state unfolding), as well as a small thermally induced transition between folded states at about 40° (W. Deutschman, unpubl. obs.). Magnesium-free ECY has been shown to unfold reversibly at 59° with an apparent ΔH^0 of 55 kcal/mol (Filimonov et al., 1993; W. Deutschman, unpubl. obs.).

Comparisons of extremely thermostable proteins such as rubredoxin (Day et al., 1992) and glutamate dehydrogenase from *Pyrococcus furiosus* (Yip et al., 1995) with mesophilic counterparts have relied on calorimetric observation of irreversible thermal transitions (Klump et al., 1994), which may considerably over-estimate the melting temperature if the transition is not two-state, or may under-estimate it if another process besides global unfolding is being observed. Studies on citrate synthase (Russell et al., 1994) used loss of enzyme activity after a short period of heating to a given temperature as a standard, which encompasses all the drawbacks of irreversible calorimetry as well as the possibility that the protein is still folded but enzymatically inactive.

Reprint requests to: S. James Remington, Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403; e-mail: jim@uoxray.uoregon.edu.

[3]Present address: Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, Washington 98109.

[4]Present address: Recombinant BioCatalysis Inc., Sharon Hill, Pennsylvania 19079.

Abbreviations: ECY, *E. coli* CheY; TMY, *Thermotoga maritima* CheY; STY, *Salmonella typhimurium* CheY; I-apo, TMY crystal form I (apo-enzyme, space group $P2_1$); II-apo, TMY crystal form II (apo-enzyme, space group $P42_12$); III-Mn, TMY crystal form III (Mn^{++} bound, space group $P2_12_12_1$); IV-Mg, TMY crystal form IV (Mg^{++} bound, space group $P4_32_12$)

Table 1. *Sequence alignment of CheY from* Thermotoga maritima *(TMY) and from* E. coli *(ECY)*

```
ECY 1   MADKELKFLVVDDFSTMRRIVRNLLKELGFNNVEEAEDGVDALNKLQAGG 50
           .: :.|:||| . ||.:::::::. |:: .:||.:| :|::|...
TMY 1   ...MGKRVLIVDDAAFMRMMLKDIITKAGYEVAGEATNGREAVEKYKELK 47
ECY 51  YGFVISDWNMPNMDGLELLKTIRADGAMSALPVLMVTAEAKKENIIAAAQ 100
         ::|. | .||:|:|:: :|.|.  ::  . .::::.| :... :|.| .
TMY 48  PDIVTMDITMPEMNGIDAIKEIMKIDP..NAKIIVCSAMGQQAMVIEAIK 95
ECY 101 AGASGYVVKPFTAATLEEKLNKIFEKLGM 129
        |||.:::|||| :. : | |||: .
TMY 96  AGAKDFIVKPFQPSRVVEALNKVSK.... 120
```

These experiments place broad limits on the stability of such proteins, which is undoubtedly high, but they are not well suited for detailed energetic analysis of the unfolding process. Multidomain proteins, oligomeric proteins, and proteins that require covalent crosslink formation for folding are very likely not to unfold reversibly, due to the increased complexity of their folding pathways, whereas CheY is a small monomeric single-domain protein with no disulfides, iron-sulfur clusters, or other crosslinks, making it an excellent candidate for comparative reversible unfolding studies.

Several recent reports of comparisons of structures of proteins from extreme thermophiles with their mesophilic counterparts (see above; Blake et al., 1992; Hennig et al., 1995; Tanner et al., 1996) have tended to emphasize whatever unusual features are found (e.g., additional salt bridges, fewer internal cavities) without experimentally establishing that these features are, in fact, responsible for increased stability toward thermal denaturation. CheY is an excellent candidate for testing hypotheses about the structural basis of thermostability, because there are existing high resolution structures from *E. coli* (Volz & Matsumura, 1991; Bellsolell et al., 1994) and *Salmonella typhimurium* (Stock et al., 1989, 1993), both of which are mesophilic and nearly identical in sequence and structure, while *Thermotoga maritima* is an extreme thermophile and its CheY protein has fairly low sequence identity (28%, Table 1) with either of the others.

Results and discussion

The structure of *Thermotoga maritima* CheY (TMY) was solved in four different crystal forms with good stereochemistry (Table 2). Its overall structure is a five-stranded β-sheet flanked by five α-helices (Fig. 1). An overlay of the model of *E. coli* CheY (ECY; Volz & Matsumura, 1991) shows a RMSD of 2.0 Å for all 120 corresponding Cα atoms of TMY (crystal form I-apo), aligned as shown in Table 1. Although TMY is nine residues shorter than ECY, the deletions are readily accommodated in the structure because they occur at the N- and C-termini and in one surface loop. The Cα atoms of the models of the different crystal forms of TMY reported here are superimposable to within 0.3 to 0.5 Å RMSD (when residues 55–62 from loop 5 are excluded).

Alternative conformations of loop 5

A large conformational change is seen between different crystal forms of TMY in loop 5 (residues 55–61), which contributes a ligand to the metal-binding site of CheY and is directly adjacent to the active site Asp-54. Crystal forms III-Mn and IV-Mg have divalent cations bound at the active site and place loop 5 in the so-called "γ-turn loop" (Volz, 1993) conformation identical to that seen in the *E. coli* and *S. typhimurium* crystal structures of CheY (Fig. 2b). The two crystal forms (I-apo and II-apo) that do not contain bound divalent cations place loop 5 in a much different conformation (Fig. 2A), with dissimilar main chain phi-psi angles throughout (e.g., residues 59–61 form an ordinary β-turn as opposed to the 60–61 γ-turn in the "open" conformation). This "closed" loop conformation buries Asp-54 more completely than does the open loop conformation, and effectively precludes phosphorylation because one carboxyl oxygen is buried and the other would form part of the essential divalent cation binding site.

Several factors may affect the adoption of one loop conformation over the other. The most obvious of these is the presence of a bound divalent cation, which is described in detail in the next section, but there are other changes in hydrogen bonding patterns that may be important. In the closed loop conformation, the car-

Table 2. *TMY data collection and refinement statistics*

Crystal form	I-apo	II-apo	III-Mn	IV-Mg
Space group	$P2_1$	$P42_12$	$P2_12_12_1$	$P4_32_12$
Molecules per asymmetric unit	1	1	2	2
a	32.04	98.66	109.66	50.07
b	53.95	98.66	64.72	50.07
c	34.16	32.25	33.81	186.29
β	95.56	—	—	—
Resolution (Å)	1.9	2.3	2.2	2.8
R-merge[a] (%)	5.3	6.7	9.0	3.7
Completeness[b] (%)	83	82	97	95
R-factor[c] (%)	18.6	17.8	18.2	17.3
Deviations from ideality				
Bond lengths (Å)	0.017	0.013	0.016	0.013
Bond angles (°)	2.6	2.2	2.4	2.0
B-factor restraints (Å^2)	8.6	9.8	7.5	6.3
NCS restraints (Å)	n/a	n/a	0.62	0.55

[a] R-merge $= \Sigma|I_{hkl} - \langle I\rangle|/\Sigma I_{hkl}$, where $\langle I\rangle$ = mean of individual measurements of I_{hkl}.

[b] Ratio of observed reflections to those theoretically possible.

[c] R-factor $= \Sigma||F_o| - |F_c||/\Sigma|F_o|$ where F_o is an observed structure amplitude, and F_c is the calculated.

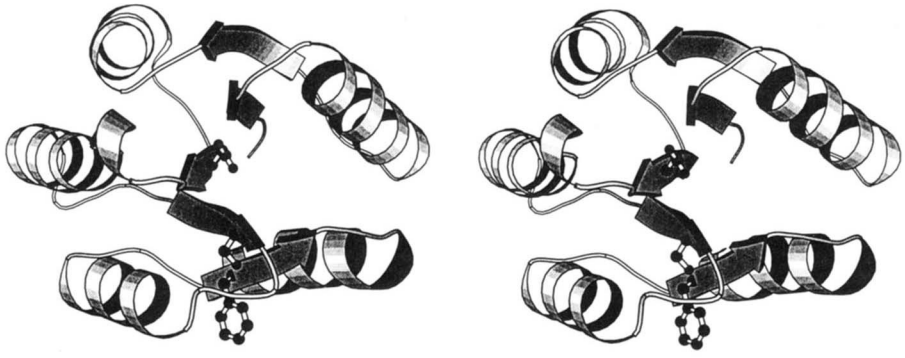

Fig. 1. Stereo ribbon diagram showing the fold and secondary structure of CheY from *Thermotoga maritima*, generated using Molscript (Kraulis, 1991). Side chains of Asp-54, Ser-82, and Phe-101 are shown.

bonyl oxygen of Thr-56 hydrogen-bonds to the side chain of Gln-86, which cannot occur with the open conformation. Gln-86 is near the probable site of phosphorylation, and changes in its interactions may effect signaling in TMY.

Both crystal forms (I-apo and II-apo) having the closed loop 5 conformation make an intermolecular crystal contact with three hydrogen bonds between Asn-61 of the loop and a neighboring molecule (Fig. 3), suggesting that this conformation is influenced

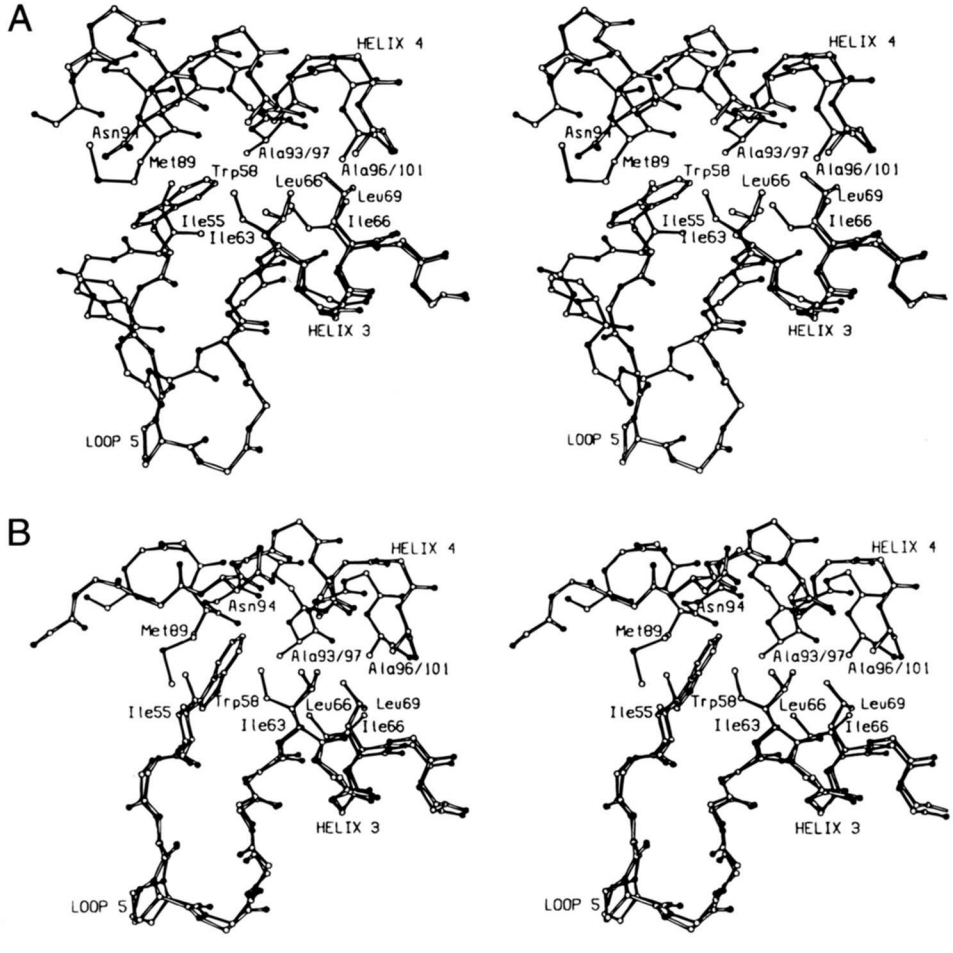

Fig. 2. Stereo view of the backbone of TMY overlaid on that of ECY, showing the shift in helix 4, and rearrangement of loop 5. The alignment was based on strand 3 and helix 3. Substitutions in the side chains of the core that allow the collapse of helix 4 in TMY are shown. **A:** Metal-free TMY structure I-apo (filled bonds), metal-free ECY structure (Volz & Matsumura, 1991, open bonds). **B:** TMY structure III-Mn (filled bonds), ECY/Mg structure (Bellsolell et al., 1994, open bonds). Carbon atoms are represented by open circles, nitrogen, and oxygen filled.

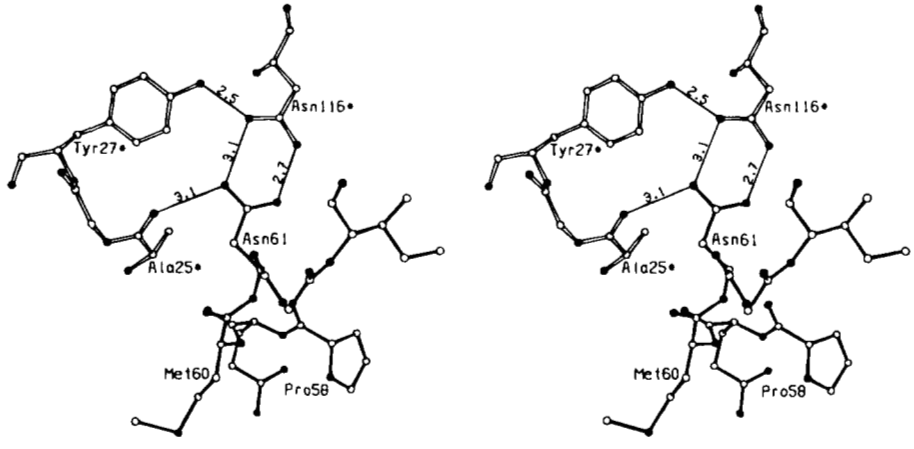

Fig. 3. Intermolecular crystal contact between Asn-61 and Asn-116 from a neighboring molecule in the metal-free TMY structures. This crystal contact occurs in the I-apo and II-apo crystal forms, but does not occur in the III-Mn or IV-Mg metal-bound crystal forms.

by crystal packing, which is otherwise dissimilar between each crystal form. The closed conformation buries about 130 $Å^2$ less non-polar surface area, suggesting that it is less stable in some respects, but this is probably offset by the new hydrogen bond. Also, crystal packing forces are typically small relative to the intramolecular forces that determine protein folding (Blake et al., 1992; Zhang et al., 1995), so that even if the Asn-61 crystal contact stabilizes the loop in a position that does not predominate in solution, it does indicate an unusual degree of flexibility in this region. NMR studies of ECY have suggested that the loop is quite mobile (Moy et al., 1994), and that it undergoes a conformational change upon phosphorylation (Lowry et al., 1994). The sequence of loop 5 is highly conserved among CheY and its homologues, and the open conformation has been suggested to form a rigid, universal recognition element across the CheY superfamily (Volz, 1993). However, we see two different, apparently isoenergetic, conformations of this loop in the crystal structures, which supports the flexibility seen in the NMR experiments, regardless of whether those specific conformations are functionally important in vivo.

Magnesium binding site

Two crystal forms of TMY (III-Mn and IV-Mg) were solved that contain divalent cations octahedrally coordinated in the active site (Fig. 4A–C). Metal binding is essentially identical to that described for *Salmonella typhimurium* CheY (STY) and ECY, with three solvent and three protein ligands (Stock et al., 1993; Bellsolell et al., 1994); however, in contrast to the latter report, no large-scale conformational changes in helix 4 and the preceding loop take place, in agreement with the results of Stock et al. (1993). The salt bridge from the conserved Lys 104 to the active site aspartates is not disrupted upon metal binding in our structures, as reported for STY by Stock et al. (1993). Crystal form III-Mn diffracts to highest resolution (Table 2) and the density of the Mn^{++} is higher (Fig. 4A) because it has more scattering electrons than does Mg^{++}, but both structures show the same features.

The most obvious distinction between the TMY structures with different loop conformations is that the crystal forms with divalent cations bound (III-Mn, IV-Mg) have loop 5 in the open form, whereas the structures with the closed-loop conformation (I-apo, II-apo) do not have a divalent cation in the active site. The closed-loop conformation apparently does not permit divalent metal ions to bind, despite the presence of three possible protein ligands in appropriate positions to be Mg^{++} ligands (Fig. 4C). One potential ligand does change, with the hydroxyl oxygen of Thr-56 being approximately in the position of a solvent molecule in the open-loop conformation. Crystal form II-apo was grown in the presence of 20 mM Mg^{++}, but no electron density is seen in the metal binding site. Therefore, the presence of metal is not causal in determining the loop 5 configuration, or in determining which crystal form results.

This result was confirmed with data collected from a II-apo crystal grown in the presence of 100 mM Ca^{++} (97% complete to 3.1 Å, *R*-merge = 9.5%). An $F_{obs,Ca} - F_{obs,Mg}$ electron density map calculated with these data showed no peaks in the active site above 2.5 σ (data not shown). Calcium has been shown in ECY to bind to the protein with a similar K_d (0.4 mM) as magnesium (1 mM, Needham et al., 1993), so the increase from 10 to 18 electrons should have been immediately obvious if there had been metal bound in the active site. The presence of the Thr-56 hydroxyl as a potential ligand should if anything have favored calcium binding, as hydroxyls are commonly seen to form calcium ligands (e.g., Reid & Hodges, 1980).

CheY requires that magnesium be bound for phosphoryl transfer, whereas the closed loop 5 conformation lacks a metal, suggesting that it is enzymatically inactive. The K_d of ECY for Mg^{++} is in the same range as physiological Mg^{++} concentrations in *E. coli*. The role of such weak Mg^{++} binding is not clear, but this conformation of magnesium-free CheY may have a functional role.

Comparison of stability determinants

The availability of thermodynamic and structural data on TMY and ECY allows one to compare various aspects of the overall structures of TMY and ECY and to relate these to the additional thermal stability of TMY. Some differences are evident from a sequence comparison alone, but these seem unlikely to account for a large increase. TMY is nine residues shorter than ECY, with shortened N-and C-termini and a two-residue deletion in loop 6 (Table 1). This may help stabilize TMY to some extent, but the CheY se-

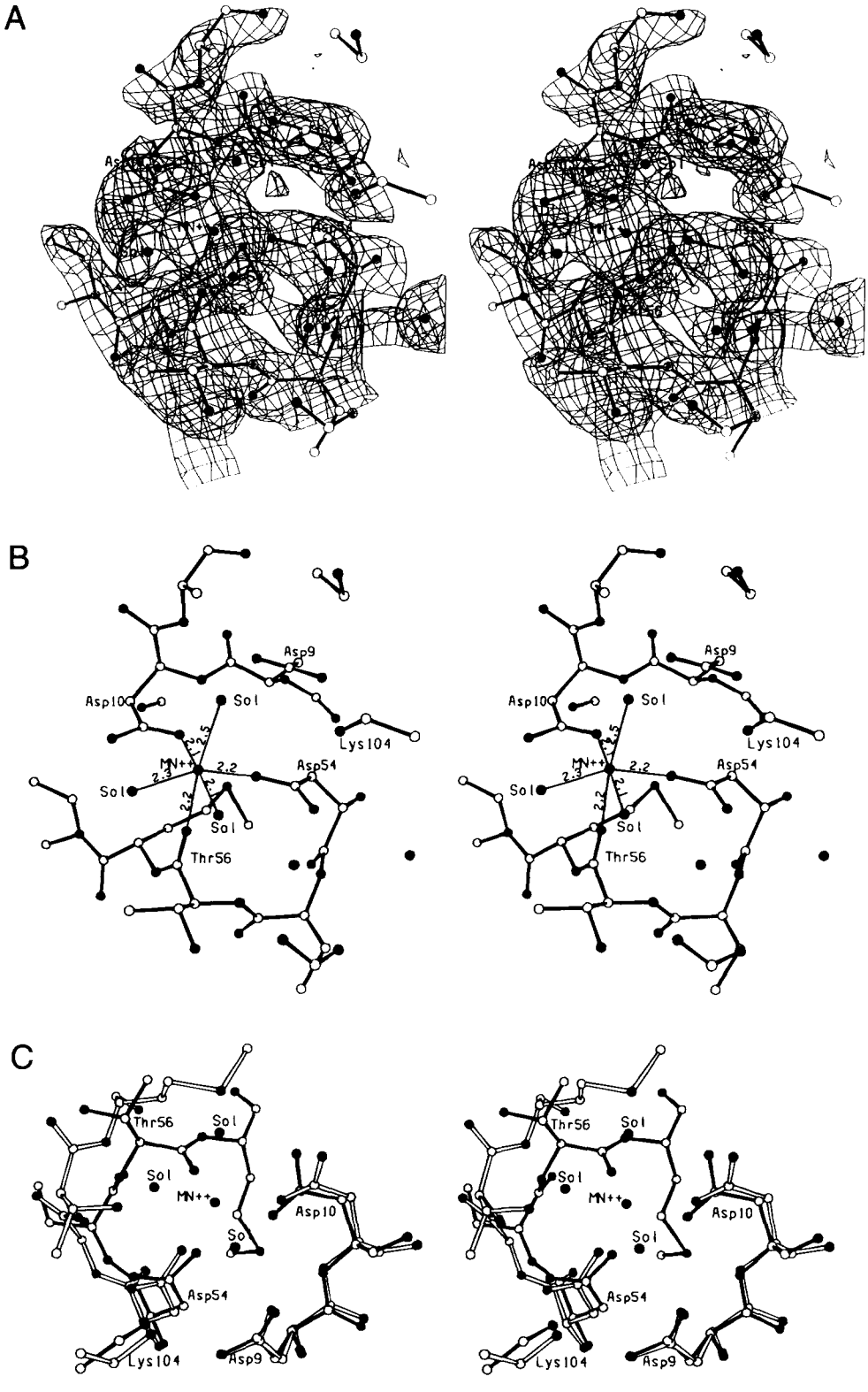

Fig. 4. Stereo views of the metal binding site of TMY, as seen in the III-Mn co-crystals. **A:** Representative portion of the final refined $2F_o - F_c$ 2.2 Å resolution electron density map contoured at 1σ, showing three protein ligands and three bound water molecules. **B:** Ligand identification and bond lengths to the bound Mn^{++}. Isolated filled circles are solvent molecules. **C:** Overlay of metal-free I-apo structure (open bonds) on the III-Mn metal binding site (filled bonds). The overlay was based on the superposition of the core β-strand C atoms. Two potential protein ligands (Asp-10 and Asp-54) remain approximately the same, while a third is apparently swapped, with the sidechain hydroxyl of Thr-56 from the metal-free form overlaying within 0.8 Å of a bound water molecule seen in the III-Mn form.

Table 3. *Sequence alignment of CheY from* Thermotoga maritima *(TMY),* Bacillus subtilis *(BSY), and* Rhizobium meliloti *(RMY)*

```
BSY 1   MAHRILIVDDAAFMRMMIKDILVKNGFEVVAEAENGAQAVEKYKEHKPDL 50
        |:.|:|||||||||||||:|||:.|.|:||.:||.|| :|||||| |||:
TMY 1   MGKRVLIVDDAAFMRMMLKDIITKAGYEVAGEATNGREAVEKYKELKPDI 50
        | ||||.|||.  :| ||  .:..||:|.  :|.:| |::||.....||:
RMY 1   MKKRVLTVDDSRTIRNMLLVTLNNAGFETI.QAEDGVEGLEKLDTANPDV 49

BSY 51  VTMDITMPEMDGITALKEIKQID AQARIIMCSAMGQQSMVIDAIQAGA  98
        ||||||||||:||.|:|||..|| :.|:||:|||||||.|||:||.|||
TMY 51  VTMDITMPEMNGIDAIKEIMKID..PNAKIIVCSAMGQQAMVIEAIKAGA 98
        :. ||.|| ::|:: |.::.| |  ....|:| .. ::...   | .|||
RMY 50  IVTDINMPRLDGFGFIEGVRKNDRYRAVPILVLTTESDAEKKNRARQAGA 99

BSY 99  KDFIVKPFQADRVLEAINKTLN 121
        |||||||||:.||:||:||.
TMY 99  KDFIVKPFQPSRVVEALNKVS  119
        .::|||||:|.::::|:::|.
RMY 100 TGWIVKPFDPTKLIDAIERVTA 120
```

quences from *Bacillus subtilis* (Bischoff & Ordal, 1991), a moderate thermophile, and *Rhizobium meliloti* (Greck et al., 1995), a mesophilic soil bacterium, show similar deletions (Table 3), suggesting that a major gain in stability is not made by the shorter sequence.

Another minor advantage that TMY may gain over ECY in its sequence is that it has two more proline residues than does ECY, one in an exposed loop and one at the N-terminal cap of a helix, both of which are locations where proline substitutions have been shown to increase protein stability in other proteins due to reduction of the entropy of the unfolded state (Matthews et al., 1987). However, Thr-112 of ECY compensates by making an N-capping hydrogen bond to the same helix that Pro-108 caps in TMY.

Comparing the three-dimensional structures, one can obtain statistics about several properties of the molecules that have been used to account for the thermostability of other proteins (Table 4). None of these properties indicate a particular advantage for enhanced stability of TMY. ECY actually buries somewhat more

Table 4. *Comparison of structural features in TMY versus ECY*

	TMY I-apo	TMY III-Mn	ECY[a] apo	ECY Mg
Buried hydrophobic surface (Å^2)[b]	7,333	7,477	8,307	8,116
Surface per carbon atom (Å^2)[c]	12.8	13.0	13.3	13.2
Total cavity volume (Å^3)[d]	172	183	193	162
Packing density[e]	0.760	0.754	0.757	0.755
Number of hydrogen bonds[f]	232	219	246	227
Number of charged hydrogen bonds[g]	25	38	40	36
Number of ion pairs[h]	5	7	5	6
Helix formation energy (kcal/mol)[i]	34.2	34.2	35.4	35.4

[a]The ECY structure used (Volz & Matsumura, 1991) contains seven amino acid side chains modeled at partial occupancy in two positions. The statistics are for an average of two models, one with all the side chains in their "A" positions, the other all in "B."

[b]Difference between the exposed carbon atom surface area with the protein in an extended conformation modeled by FRODO and that in the crystal structure, calculated using EDPDB (Zhang, 1992).

[c]Buried hydrophobic surface normalized to the number of carbon atoms in the structure.

[d]Cavity volume accessible to a 1.2 Å probe, calculated using PQMS (Connolly, 1993).

[e]Ratio of the Van der Waals volume of the protein calculated with PQMS and a 0 Å radius probe to the net solvent accessible volume calculated using PQMS and a 1.2 Å probe radius.

[f]Non-bonded contacts between potential hydrogen bond donors and acceptors within a distance of 3.5 Å, with no angle limitation, counted using EDPDB.

[g]Hydrogen bond contacts where either (but not both) of the donor or acceptor are charged (i.e., Arg or Lys donors, Glu or Asp acceptors, neither ECY nor TMY contain a His residue), calculated using EDPDB.

[h]Contacts between positively and negatively charged atoms within a distance of 4.0 Å, calculated using EDPDB. Multiple contacts between two residues are only counted once. Some structures include one or two interactions where the occupancies of the atoms involved are zero, indicating arbitrary positioning of the model: TMY I-apo (2), TMY III-Mn (1), and ECY/Mg (1).

[i]Sum of the energy-favoring helix formation from each of the residues classified as helical in TMY, excluding the first and last residues of each helix, calculated relative to glycine using a scale measured from amino acid substitutions at a site in a helix in T4 lysozyme (Blaber et al., 1994).

hydrophobic surface than does TMY, even when the measurement is normalized to the number of non-polar atoms in each molecule to account for the somewhat larger size of ECY. This should disfavor stability of TMY, although it has been calculated that the hydrophobic effect becomes less and less effective at stabilizing folded proteins as the temperature increases (Privalov, 1979).

The internal cavity volumes calculated for TMY and for ECY are very similar, as are their packing densities. One case where TMY does appear to have improved packing involves several amino acid substitutions: Trp-58/Ile-55, and Leu-66/Ile-63 Leu-69/Ile-66 substitutions going from ECY to TMY allow helix 4 of TMY to pack more closely with the rest of the protein (Fig. 2A, B), although the tryptophan substitution also reduces the size of the hydrophobic core.

The numbers of different categories of hydrogen bonds are all very similar between the structures. Unlike the case described in a comparison of D-glyceraldehyde dehydrogenases from thermophilic and mesophilic sources (Tanner et al., 1996), no trend increasing the number of hydrogen bonds, either between neutral or charged donors and acceptors, is seen in the TMY structure. Indeed, there is some variation in numbers of hydrogen bonds counted in different crystal forms of the same molecule that is similar to the difference seen between the two different proteins.

Peptide sequences differ in their tendency to form helical secondary structure (O'Neil & Degrado, 1990), and this propensity of those sequences in helical regions of a protein may affect the protein's overall stability. Calculations on the sequences of the helical regions of TMY and ECY show a similar degree of stabilization in both, based on the scale reported by Blaber et al. (1994), although ECY, the less stable of the two proteins, actually shows about 1 kcal/mol more stabilization by that summation. Sequences in β-strand secondary structure have not been seen to have a similar inherent stabilizing effect: their stability seems much more context dependent within the tertiary structure of the protein (Minor & Kim, 1994). A secondary structure prediction algorithm (Rost & Sander, 1993) run on the two sequences was somewhat better at predicting β-strand structure in TMY than in ECY (data not shown), but this difference is likely more reflective of the difficulty of accurately predicting β-strands in proteins than of a real difference in stability.

Some proteins from extreme thermophiles, such as glutamate dehydrogenase from *Pyrococcus furiosus* (Yip et al., 1995) and indole-3-glycerol phosphate synthase (Hennig et al., 1995), have been observed with extensive networks of surface ion pairs that are absent from their mesophilic counterparts. This is not observed in TMY. Each of the CheY structures has between a similar number of ion pairs counted on the surface of the protein (Table 4). The only such feature is seen in ECY, where Arg-18 and Arg-22 make a total of five bridges with Glu-35 and Glu-37, but *E. coli* is mesophilic.

The sources of the stability of proteins from extreme thermophiles are of interest due to their impact on our ability to understand and predict protein stability and folding in general, the utility of making thermostable enzymes for industrial purposes, and interest in the ecological niche occupied by extreme thermophiles. The crystal structure of TMY appears to be very similar to that of its mesophilic counterpart, and the observed trends between ECY and TMY seem to contradict several recent proposals about the sources of stability of proteins from extreme thermophiles.

Our results are consistent with the conclusions drawn from the first structural study of a thermostable enzyme, thermolysin (Matthews et al., 1974). Small changes observed, such as shortening of the sequence, increased numbers of proline residues, and readjustments of core packing, are likely to aid stability, but no single factor predominates, as has been suggested previously for other extremely thermophilic proteins. We conclude instead that the thermostability of TMY is due to many small stabilizing (and presumably additive) differences with respect to mesophilic counterparts. Considering that the high-temperature growth conditions of *Thermotoga* and other extreme thermophiles may resemble that of the first life forms, it seems reasonable that the proteins of most modern organisms have been destabilized evolutionarily to function in colder environments, a process that one could readily imagine to result from an accumulation of many small changes in the details of side-chain interactions. This hypothesis can and will be tested. The *B. subtilis* CheY is 70% identical in sequence to TMY (Table 3), with only 18 differences in internal side chains. We are in the process of replacing each of these singly and in combination in order to determine if proteins of intermediate stability can be constructed and the extent to which these changes are additive.

Possible role of Ser-82/Phe-101 in signaling

NMR studies on ECY (Lowry et al., 1994) suggest that structural changes due to phosphorylation are widespread, with large chemical shift changes occurring for about 22 backbone amides in several parts of the protein. Important differences appear between ECY and TMY in these regions. Thr-87 of ECY (Ser-82 in TMY) is one such residue, and while the backbone atoms overlay well between the two structures, the side-chain hydroxyl orients differently in the two structures (Fig. 5). In TMY the hydroxyl points into the bottom of a largely hydrophobic cleft, forming a long hydrogen bond to a bound solvent molecule. In ECY, the methylene group of Thr-87 occupies this position, and no bound solvent molecule is seen. The sidechain hydroxyl group points away to a different surface of the molecule and is about 6 Å from OD2 of the phosphorylatable active site aspartate. The side chain of Tyr-106 has been modeled in two conformations in ECY (Volz & Matsumura, 1991). In one, it fills the hydrophobic pit, while the second is solvent exposed (Fig. 5). It has previously been suggested that burial or exposure of Tyr-106 in ECY may effect signaling, based on a crystal structure of a mutant of ECY, threonine 87, to isoleucine (Ganguli et al., 1995). Also in ECY, Tyr-106 is adjacent to a major surface region (Ala-90, Val-108, Phe-111, Thr-117) shown by suppressor mutations to be critical for binding to the flagellar motor proteins FliG and FliM (Roman et al., 1992). In TMY, the equivalent residue (Phe-101) is in a single, solvent-exposed conformation. Modeling studies show that it could occupy the hydrophobic cleft in a similar conformation as Tyr-106, except that there would be a severe steric clash with the serine hydroxyl and the bound water molecule.

These results suggest that Ser-82 and Phe-101 may act together as a conformational switch, sensitive to the phosphorylation state of Asp-54. In this scheme, the unphosphorylated state is represented as described in the current structure, and phosphorylation causes Ser-82 to re-orient its hydroxyl towards the active site, releasing the bound solvent molecule and allowing Phe-101 to become buried (Fig. 6), making the surface of CheY, remote from the active site, complementary to its target on the flagellar motor. A similar mechanism is possible in ECY, and residue substitutions such as Ser/Thr and Phe/Tyr may be part of a subtle compensation

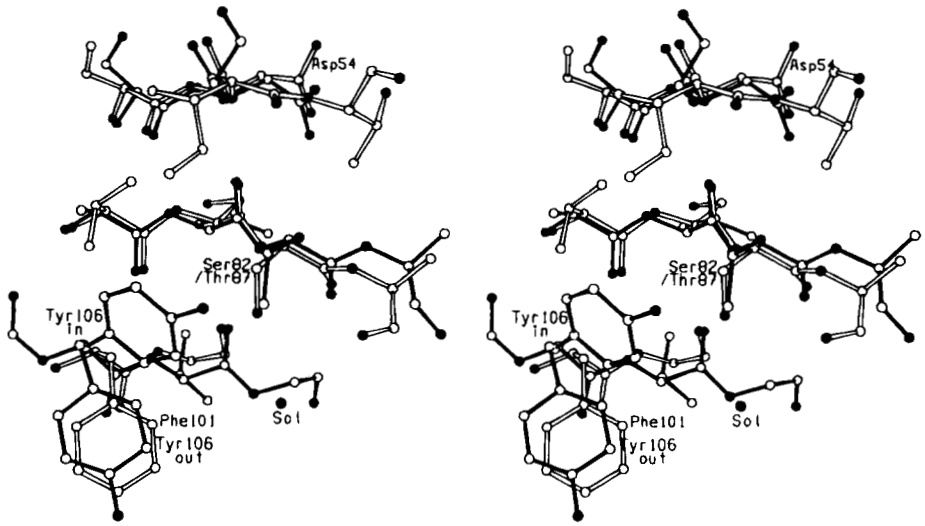

Fig. 5. Overlay of segments of the core β-sheet near Ser-82 from TMY (open bonds) versus Thr-87 of ECY (filled bonds), showing that their χ_1 side-chain rotamers differ by about 110°. The figure also shows the relationship of Ser-82 to other important residues, including Phe-101 and the active site Asp-54.

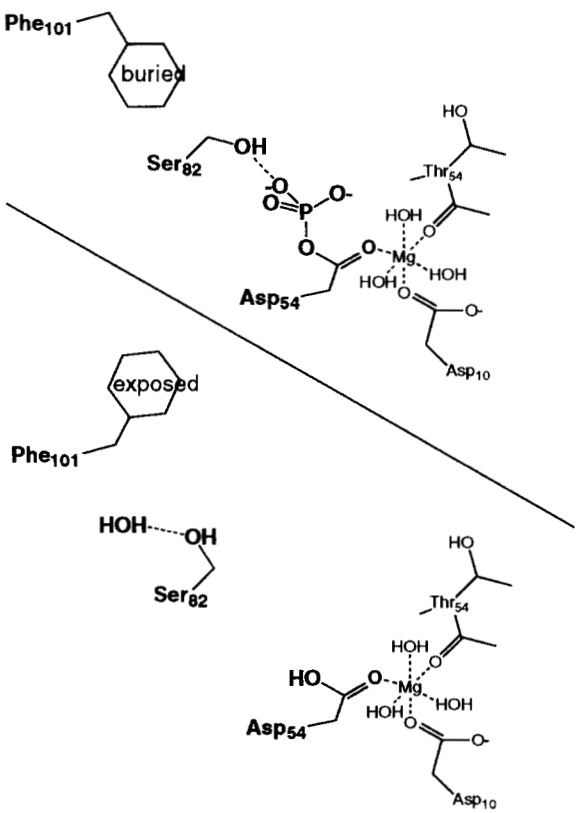

Fig. 6. Diagram of concerted side-chain movements suggested to take place upon phosphorylation: a change in the rotamer of Ser-82 permits burial of the solvent-exposed side chain of Phe-101, in turn making that portion of the surface of CheY complementary to its target on the flagellar motor.

adapting the operation of the switch to function at the very different growth temperatures of the two organisms.

Coordinates and structure factors have been deposited with the Brookhaven Protein Data Bank, ID codes 1TMY, R1TMYSF for crystal form I-apo; 2TMY, R2TMYSF for II-apo; 3TMY, R3TMYSF for III-Mn; 4TMY, R4TMYSF for IV-Mg.

Methods

CheY protein from *Thermotoga maritima* was expressed using the pQE12 expression vector (Qiagen) in *E. coli* K0641/*recA* (D*cheY*) and purified as described previously (Swanson et al., 1996). A 10-min incubation in an 80 °C water bath provided the primary means of purification, because it caused denaturation and precipitation of almost all proteins except the thermostable TMY. Purified TMY was concentrated to 5 mg/mL and crystallized by vapor diffusion from 0.1 M Tris buffer (pH 8.5) and 1.8 M ammonium phosphate in the absence of divalent cations. Single crystals (of crystal form I-apo) grew to a maximum size of about 0.2 × 0.3 × 0.1 mm^3. Crystal form II-apo in space group $P42_12$ was grown from 0.2 M ammonium sulfate, 0.1 M acetate buffer (pH 4.6), 30% PEG 4000, and 15 mM $MgCl_2$. Crystal form III-Mn grew in space group $P2_12_12_1$ from manganese chloride (15 mM), 0.2 M ammonium sulfate, 0.1 M Hepes (pH 7.0), and 25% PEG 4000, while IV-Mg crystals in space group $P4_32_12$ were grown under similar conditions but containing a different divalent cation: magnesium chloride (15 mM), 0.2 M ammonium sulfate, 0.1 M Hepes (pH 7.0) and 30% PEG 4000. Identical IV-Mg crystals also grew using Tris (pH 8.5) buffer or Pipes (pH 6.5) buffer.

X-ray data were collected and processed using an *R*-Axis IIc image plate system and supplied software from Molecular Struc-

ture Corporation. Initially, a model based on the known structure of the *E. coli* apo-enzyme (Volz & Matsumura, 1991), with non-conserved loops and side chains removed, was used in an attempt to solve the TMY structure (crystal form I-apo) by molecular replacement using the programs ROTFUN and FASTRAN (Zhang, 1992), but these attempts failed to give a satisfactory solution. Further X-ray data were collected on a crystal (form I-apo) soaked in 1 mM $HgCl_2$ mother liquor, using a Xuong-Hamlin multi-wire detector and processed using the supplied software (Howard et al., 1985) having an *R*-merge of 6.6% and data completeness of 84% (including Bijvoet pairs) to 3.2 Å resolution, and an *R*-iso of 16.8% against the native data. A single 8σ peak in the ($V = 0.5$) section of the difference Patterson function identified the heavy atom position. Solvent channels and some fragments of secondary structure were visible in an electron density map based on the SIR phases. The modified *E. coli* model was translated so that the position corresponding to SG of Cys-84, the only cysteine in TMY, was coincident with the location of the mercury atom, then rotated about all possible orientations (as well as translated in a 1 Å grid from -5 to $+5$ Å in x, y, and z) about the heavy atom site using the program RSS (S.J. Remington, unpubl. obs.). Correlation coefficients calculated in real space between the coordinates and the SIR map revealed the correct orientation. Rigid-body and initial *xyz* coordinate refinement with the TNT program package (Tronrud et al., 1987) of the clipped and correctly oriented *E. coli* model lowered the crystallographic *R*-factor from 50.1 to 42.3% at 3.5 Å resolution against the native data. A phase-combined electron density map with the SIR data was calculated with SIGMAA (Read, 1986), and model-building of misplaced and missing side chains and residues using O (Jones et al., 1991) ensued, alternating with further refinement with TNT. In later cycles, the phase-combined map was abandoned in favor of one solely phased on the improved model, correlated temperature factors were refined, and 11 water molecules were added.

Data were collected for the three other crystal forms and molecular replacement solutions were found using the AMoRe program (Navaza, 1994). Non-crystallographic symmetry restraints were applied during TNT refinement of the crystal forms that had two molecules per asymmetric unit. Late in the refinement, side-chain atoms with a temperature factors over 90 $Å^2$ and weak electron density were assigned an occupancy of zero.

Coordinate overlays of the structures were aligned using Cα coordinates and the least squares minimization routine in the program EDPDB (Zhang, 1992). Methods used for the calculation of comparative structural statistics are listed in the footnotes of Table 4.

Acknowledgment

This work was supported by National Science Foundation Grants MCB9418479 (S.J.R.), and MCB9604213 (F.W.D.).

References

Babu YS, Bugg CE, Cook WJ. 1988. Structure of calmodulin refined at 2.2 Å resolution. *J Mol Biol 204*:191–204.

Bellsolell L, Prieto J, Serrano L, Coll M. 1994. Magnesium binding to the bacterial chemotaxis protein CheY results in large conformational changes involving its functional surface. *J Mol Biol 238*:489–495.

Bischoff DS, Ordal GW 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.

Blaber M, Zhang X-J, Lindstrom JD, Pepiot SD, Baase WA, Matthews BW. 1994. Determination of α-helix propensity within the context of a folded protein. *J Mol Biol 235*:600–624.

Blake PR, Park JB, Zhou ZH, Hare DR, Adams MW, Summers MF. 1992. Solution-state structure by NMR of zinc-substituted rubredoxin from the marine hyperthermophilic archaebacterium *Pyrococcus furiosus*. *Protein Sci 1*:1508–1521.

Bourret RB, Drake SK, Chervitz SA, Simon MI, Falke JJ. 1993. Activation of the phosphosignaling protein CheY. II. Analysis of activated mutants by 19F NMR and protein engineering. *J Biol Chem 268*:13089–13096.

Connolly ML. 1993. The molecular surface package. *J Mol Graphics 11*:139–141.

Day MW, Hsu BT, Joshua TL, Park JB, Zhou ZH, Adams MW, Rees DC. 1992. X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaebacterium *Pyrococcus furiosus*. *Protein Sci 1*:1494–1507.

Filimonov VV, Prieto J, Martinez JC, Bruix M, Mateo PL, Serrano L. 1993. Thermodynamic analysis of the chemotactic protein from *Escherichia coli*, CheY. *Biochemistry 32*:12906–12921.

Ganguli S, Wang H, Matsumura P, Volz K. 1995. Uncoupled phosphorylation and activation in bacterial chemotaxis, the 2.1-Å structure of a threonine to isoleucine mutant at position 87 of CheY. *J Biol Chem 270*:17386–17393.

Greck M, Platzer J, Sourjik V, Schmitt R. 1995. Analysis of a chemotaxis operon in *Rhizobium meliloti*. *Mol Microbiol 15*:989–1000.

Hennig M, Darimont B, Sterner R, Kirschner K, Jansonius JN. 1995. 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: Possible determinants of protein stability. *Structure 3*:1295–1306

Howard AJ, Nielsen C, Xuong NH. 1985. Software for a diffractometer with multiwire area detector. *Methods Enzymol 114*:452–471.

Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A 47*:110–119.

Klump HH, Adams MW, Robb FT. 1994. Life in the pressure cooker: The thermal unfolding of proteins from hyperthermophiles. *Pure Appl Chem 66*:485–489.

Kraulis P. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Lowry DF, Roth AF, Rupert PB, Dahlquist FW, Domaille PJ, Matsumura P. 1994. Signal transduction in chemotaxis, a propagating conformation change upon phosphorylation of CheY. *J Biol Chem 269*:26358–26362.

Matthews BW, Weaver LH, Kester C. 1974. The conformation of thermolysin. *J Biol Chem 249*:8030–8044.

Matthews BW, Nicholson H, Becktel WA. 1987. Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc Natl Acad Sci USA 84*:6663–6667.

Minor DL Jr, Kim PS. 1994. Context is a major determinant of β-sheet propensity. *Nature 371*:264–267.

Moy FJ, Lowry DF, Matsumura P, Dahlquist FW, Krywko JE, Domaille PJ. 1994. Assignments, secondary structure, global fold, and dynamics of chemotaxis Y protein using three- and four-dimensional heteronuclear (^{13}C, ^{15}N) NMR spectroscopy. *Biochemistry 33*:10731–10742.

Navaza J. 1994. AMoRe: An automated package for molecular replacement. *Acta Crystallogr A 50*:157–163.

Needham JV, Chen TY, Falke JJ. 1993. Novel ion specificity of a carboxylate cluster Mg(II) binding site: Strong charge selectivity and weak size selectivity. *Biochemistry 32*:3363–3367.

O'Neil KT, DeGrado WF. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science 250*:646–651.

Privalov PL. 1979. Stability of proteins: Small globular proteins. *Adv Protein Chem 33*:167–241.

Read RJ. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr A 42*:140–149.

Reid RE, Hodges RS. 1980. Cooperativity and calcium/magnesium binding to troponin C and muscle calcium binding parvalbumin: An hypothesis. *J Theor Biol 84*:401–444.

Roman SJ, Meyers M, Volz K, Matsumura P. 1992. A chemotactic signaling surface on CheY defined by suppressors of flagellar switch mutations. *J Bacteriol 174*:6247–6255.

Rost B, Sander C. 1993. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol 232*:584–599

Russell RJM, Hough DW, Danson MJ, Taylor GI. 1994. The crystal structure of citrate synthase from the thermophilic Archaeon, *Thermoplasma acidophilum*. *Structure 2*:1157–1167.

Stock AM, Mottonen JM, Stock JB, Schutt CE. 1989. Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis. *Nature 337*:745–749.

Stock AM, Martinez-Hackert E, Rasmussen BF, Stock JB, Ringe D, Petsko GA. 1993. Structure of the Mg^{2+}-bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis. *Biochemistry 32*:13375–13380.

Swanson RV, Sanna MG, Simon MI. 1996. Thermostable chemotaxis proteins from the hyperthermophilic bacterium *Thermotoga maritima*. *J Bacteriol 178*:484–489.

Tanner JJ, Hecht RM, Krause KL. 1996. Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution. *Biochemistry 35*:2597–2609.

Tronrud DE, Ten Eyck LF, Matthews BW. 1987. An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr A 43*:489–501.

Volz K, Matsumura P. 1991. Crystal structure of *Escherichia coli* CheY refined at 1.7-Å resolution. *J Biol Chem 266*:15511–15519.

Volz K. 1993. Structural conservation in the CheY superfamily. *Biochemistry 32*:11741–11753.

Yip KS, Stillman TJ, Britton KL, Artymiuk PJ, Baker PJ, Sedelnikova SE, Engel PC, Pasquo, AR, Consalvi, V. 1995. The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure 3*:1147–1158.

Zhang X-J. 1992. Use of poly-alanine mutagenesis to probe the structure of T4 lysozyme [Ph.D. Dissertation]. Oregon: University of Oregon.

Zhang X-J, Wozniak JA, Matthews BW. 1995. Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme. *J Mol Biol 250*:527–552.