Acta Cryst. (1995). D**51**, 145–154

The Three-Dimensional Structure of the Aspartate Receptor from *Escherichia coli*

BY JAMES U. BOWIE*

Department of Chemistry and Biochemistry, and DOE Laboratory of Structural Biology and Molecular Medicine, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90024, USA

ANDREW A. PAKULA

Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA, and Scriptech Pharmaceuticals, Cambridge, MA, 200 Boston Avenue, Suite 3000, Medford, MA 02155, USA

AND MELVIN I. SIMON

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

(*Received* 1 *July* 1994; *accepted* 13 *September* 1994)

Abstract

The crystal structure of the periplasmic domain of the aspartate receptor from *Escherichia coli* has been solved and refined to an *R* factor of 0.203 at 2.3 Å resolution. The dimeric protein is largely helical, with four helices from each monomer forming a four-helix bundle. The dimer interface is constructed from four helices, two from each subunit, also packed together in a four-helix bundle arrangement. A sulfate ion occupies the aspartate-binding site. All hydrogen bonds made to aspartate are substituted by direct or water-mediated hydrogen bonds to the sulfate. Comparison of the *Escherichia coli* aspartate-receptor structure with that of *Salmonella typhimurium* [Milburn, Prive, Milligan, Scott, Yeh, Jancarik, Koshland & Kim (1991). *Science*, **254**, 1342–1347; Scott, Milligan, Milburn, Prive, Yeh, Koshland & Kim (1993). *J. Mol. Biol.* **232**, 555–573] reveals strong conservation in the structure of the monomer, but more divergence in the orientation of the subunits with respect to one another. Mutations that render the *Escherichia coli* receptor incapable of responding to maltose are either located in spatially conserved sites or in regions of the structures that have high temperature factors and are therefore likely to be quite flexible. The inability of the receptor from *Salmonella typhimurium* to respond to maltose may, therefore, be because of differences in amino acids located on the binding surface rather than structural differences.

Introduction

Cells can monitor and respond to changes in their environment by means of specific transmembrane receptor proteins. In the case of various enteric bacteria, methyl-accepting chemotaxis receptors are responsible for detecting the presence of specific chemicals in the media and altering swimming behaviour appropriately (Macnab, 1987).

The *E. coli* aspartate receptor is a 553-amino-acid protein that spans the membrane twice *via* putative transmembrane helices (Krikos, Mutoh, Boyd & Simon, 1983). The protein mediates the movement of *E. coli* cells toward the attractants aspartate and maltose and away from the repellents nickel and cobalt (Reader, Tsu, Springer, Goy & Adler, 1979; Wang & Koshland, 1980). Aspartate binds directly to the periplasmic domain of the protein (Mowbray, Foster & Koshland, 1985), while the maltose response is mediated by binding to maltose-binding protein (Hazelbauer, 1975). It is not known whether the response to nickel and cobalt requires ancillary factors as is the case with maltose. In order to alter swimming behaviour upon binding of the various effectors, a signal must be transmitted from the periplasmic domain to the cytoplasmic domain *via* the transmembrane segments.

It is likely that transmembrane signalling involves some sort of conformational change initiated in the periplasmic domain of the protein, yet the nature of this structural alteration remains controversial (Milburn *et al.*, 1991; Milligan & Koshland, 1991; Pakula & Simon, 1992; Scott *et al.*, 1993; Yeh, Biemann, Pandit, Koshland & Kim, 1993). Milburn *et al.* have reported the structure of an active disulfide crosslinked mutant of the periplasmic domain of the *S. typhimurium* aspartate receptor both in the absence (AR-St-apo) and presence of aspartate (AR-St-asp) (Milburn *et al.*, 1991). They observed only minor changes in the structure of the individual subunits upon aspartate binding, but did note a slight rotation of the subunits in the bound form relative to the free form. The authors point out that while the observed rotation of the subunits is subtle, a small rotation could produce a large translation of the cytoplasmic domains since they are at a great distance

* To whom correspondence should be addressed.

from the pivot point. Additional support for this model was provided by the subsequent crystal structure determination of the periplasmic domain of the wild-type protein in the aspartate bound and unbound forms (Yeh *et al.*, 1993). Although the crystallization conditions and space groups were different than for the previously solved structures, essentially the same conformational change was found to occur upon aspartate binding. Thus, the observed conformational change is unlikely to be due to crystal packing forces. The model, however, contradicts biochemical evidence that signalling can occur *via* a single subunit and does not involve large relative movements of the first transmembrane helices (Milligan & Koshland, 1991; Stoddard, Bui & Koshland, 1992).

Materials and methods

Construction of a plasmid expressing the periplasmic fragment of the E. coli aspartate receptor

A plasmid, pAP202, was constructed that expresses the ligand-binding domain of the *E. coli* aspartate receptor, residues 32–188 (AR-Ec). Site-directed mutagenesis (Maniatis, Fritsch & Sambrook, 1982) was used to construct mutations in a recombinant m13 bacteriophage that contains the small *Kpn*I/*Bam*HI restriction fragment of the aspartate receptor from pNT201 (Borkavich, Kaplan, Hess & Simon, 1989). This bacteriophage was constructed by ligating the appropriate fragment of pNT201 into the unique *Kpn*I and *Bam*HI sites of m13mp19 (Yanisch-Perron, Vieira & Messing, 1985) using standard procedures (Maniatis *et al.*, 1982). Two oligonucleotides were used. The first of these (GCTTTCTGGCTATGGTGAACCATG-GAGAAAAACAGACTTCC) created a new ATG codon at wild-type residue position 32, changed the residue at wild-type position 33 from Leu to Val (CTT→GTT codon) and introduced an *Nco*I restriction site (CCATGG) overlapping the new ATG codon. The second oligonucleotide (GTTGCCACTGGGCT-CATCGGTAATCATCTGC) introduced a TGA stop codon at wild-type residue position 189. The unique *Nco*I/*Kpn*I fragment of the resulting recombinant m13 phage was introduced between the unique *Nco*I and *Kpn*I restriction sites of pAP119 (Pakula & Sauer, 1989) to create plasmid pAP202, which places the new gene expressing AR-Ec under the transcriptional control of the *tac* promoter.

Purification of the aspartate receptor periplasmic fragment

An overnight culture of pAP202/GW5100 was used to inoculate LB media (Maniatis *et al.*, 1982) containing 100 μg ml^{-1} ampicillin. This culture was grown at 310 K until the absorbance at 600 nm reached 1.0, whereupon isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 0.1 m*M* to induce expression of the AR-Ec protein. Growth was continued for an additional 3 h and cells were harvested by centrifugation. 125 g of wet cell paste was resuspended in 125 ml of SB [50 m*M* Tris–HCl (pH 7.5), 0.1 m*M* EDTA, 1.4 m*M* β-mercaptoethanol, 10% glycerol] with the additon of PMSF to a final concentration of 2.0 m*M* and the cells lysed by sonication. Cell debris was pelleted by centrifugation in a Ti45 rotor for 1 h at 35 000 rev min^{-1}. The volume of the supernatant was adjusted to 200 ml with SB and the AR-Ec protein was precipitated by the addition of 59 g ammonium sulfate. After stirring overnight at 277 K the pellet was recovered by centrifugation. The ammonium sulfate pellet was resuspended in SB and the remaining ammonium sulfate removed by dialysis against SB. The protein solution was then applied to a 300 ml column of DEAE sephacel and AR-Ec eluted in the flow-through fraction. The AR-Ec-containing flow-through fraction was then applied to a 200 ml blue sepharose column in SB and eluted with a linear gradient of 0 to 1.0 *M* NaCl in SB, 2 l per side. Fractions containing AR-Ec were pooled and, when necessary, additional purification was performed by gel filtration on a superose 12 column (Pharmacia).

Crystal growth

Crystals were grown by mixing 5 μl of 25 mg ml^{-1} AR-Ec in 200 m*M* ammonium sulfate adjusted to pH 7.0, with 5 μl of crystallization buffer (1.1 *M* ammonium sulfate, 20 m*M* ammonium formate and 30 m*M* formic acid, pH 3.6) on a silanized glass coverslip. The coverslip was then inverted over a well containing 1 ml of crystallization buffer. Diamond-shaped crystals up to 1 mm in the shortest dimension grew after 2 d incubation at room temperature. The crystals grow in space group $P4_12_12$ with cell dimensions $a = b = 87.1$ and $c = 106.0$ Å. The asymmetric unit contains one monomer and the crystals have a solvent content of 70%.

Data collection

Diffraction intensity data were either collected on a Xuong–Hamlin Mark II multiwire area detector (Hamlin, 1985) or a Rigaku R-AXIS II image-plate detector. For room temperature data collection, crystals were first transferred to synthetic mother liquor (1.5 *M* ammonium sulfate, 20 m*M* ammonoum formate and 30 m*M* formic acid) before mounting in a glass or quartz capillary. Diffraction intensity was found to decay fairly rapidly. Data beyond 3.5 Å resolution was generally too weak to collect after 8 h, and often much sooner. Consequently, it usually required many attempts to obtain reasonably complete data for the heavy-atom derivatives from a single crystal. Subsequent to collecting heavy-atom derivative data, we found that higher resolution data could be obtained by flash freezing the crystals under a stream of liquid nitrogen. For low-temperature data collection, crystals were transferred into synthetic mother liquor containing 100 m*M* cobalt chloride, with

30% glycerol as a cryoprotectant and soaked for 12 h or more. The crystals were scooped into a loop made of human hair attached to a goniometer head and flash-frozen under a stream of nitrogen at 123 K. The frozen crystals initially diffract to a maximum resolution of 2.0 Å, but were still found to decay, albeit at a much slower rate. Consequently, it was possible to obtain a data set to 2.3 Å at the low temperature. The cell dimensions changed to $a = b = 85.5$, $c = 103.8$ Å. The cobalt chloride was added in the hope of finding a cobalt-binding site. Because no cobalt binding was observed and this data set was of somewhat better quality than a previously obtained native data set, we decided to use this data set for refinement.

Heavy-atom derivatives were prepared by transferring the crystals to synthetic mother liquor containing the heavy-atom compounds. Most heavy-atom compounds were relatively insoluble in the synthetic mother liquor, but in several cases, the concentrations were high enough to produce useful derivatives. The barium tungstate (BaW) derivative was obtained by soaking a crystal for 12 h in 40 µl of a 30% saturated solution of BaW. The $Na_2U_2O_7$ (U) derivative was obtained after an 18 h soak in 40 µl of 33% saturated solution of U. The platinum ethylenediamine chloride (PtE) derivative was prepared by soaking a crystal for 18 h in 40 µl of a saturated solution of PtE.

Structure determination

Except where noted, crystallographic calculations were performed using the *CCP*4 package of programs (Collaborative Computational Project, Number 4, 1994). A single major site for the U derivative was easily identified from the difference Patterson and anomalous difference Patterson maps. The single U site was refined using the UCLA version of the program *HEAVY* (Terwilliger & Eisenberg, 1983) and a second minor U site was then identified from a difference Fourier map. The BaW and PtE sites were identified with cross difference Fourier maps using protein phases obtained from the U derivative. These sites were subsequently verified in Patterson maps and protein phases determined from the BaW and PtE derivatives could be used to return the U sites in cross difference Fourier maps. Because derivative data was collected on two different devices, difference Fourier maps and heavy-atom refinements were performed using native data sets collected using the same device. The mean factorial isomorphous difference between the two native data sets is 0.10. Heavy-atom positions and occupancies were refined separately and protein phases calculated for each derivative were then merged in various combinations. Protein phases calculated using the uranate derivative alone produced an interpretable electron-density map and combining phases with those obtained from both the PtE and BaW derivatives improved the maps somewhat. The overall figure of merit prior to phase refinement was 0.52 to 3.0 Å resolution. Phases were greatly improved using the program *SQUASH* of Zhang & Main (Zhang, 1993; Zhang & Main, 1990) which applies the constraints of solvent flatness, a reasonable density histogram and Sayre's equation. Multiple isomorphous replacement (MIR) phases were initially refined at 3.5 Å resolution and then slowly extended to 2.8 Å resolution in ten cycles, combining the refined phases with the MIR phases at each step. The phase-extension procedure is described by Zhang (Zhang, 1993). The mean phase error with respect to the partially refined room-temperature model improved from 67 to 57° to 2.8 Å resolution.

The final experimental electron-density map was extremely clear and could be fitted unambiguously. A portion of the electron-density map is shown in Fig. 1. Density was only missing from some solvent-exposed side chains, the amino and carboxy termini, and density was weak for a portion of the loop connecting helix *A* and helix *B*. While the region of the aspartate receptor expressed from the gene construct comprises amino acids 32–188, electron density was not observed outside residues 39–181. This either reflects disorder in the crystal or proteolysis during purification. The DNA sequence was also checked (data not shown).

Refinement

All refinement steps were carried out using *X-PLOR* Version 2.1 (Brünger, 1990) including all reflections with intensity greater than 1σ. A starting model comprising residues 39–181 was initially refined against native data collected at room temperature (native I) since low-temperature data had not yet been obtained. The *R* factor of the starting model dropped from 48.2 to 31.2 after a single round of positional refinement. A sulfate ion was built into the aspartate-binding pocket during refinement to account for a large electron-density peak found in

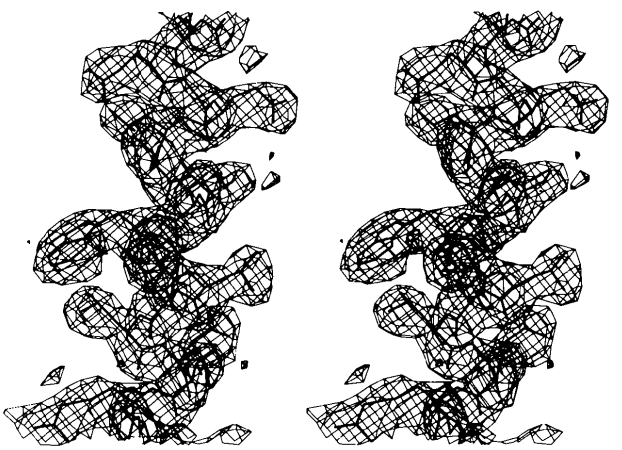


Fig. 1. Stereoview of a portion of the electron-density map. The map is contoured at 1σ of the map r.m.s. deviation and was calculated using data from 20 to 2.8 Å resolution with MIR phases after refinement with *SQUASH* (Zhang & Main, 1990).

Table 1. *Crystallographic statistics*

Data collection

Crystal	Detector	Resolution (Å)	No. of reflections measured	No. of unique reflections	Percentage complete	R_{merge}*	Final shell† Percentage complete	Final shell† $\langle I/\sigma \rangle$
Native I	R-AXIS	2.8	67447	10130	97	6.7	94	3.2
Native II	Hamlin	3.2	24163	6927	97	6.3	70	5.6
Native (cold)	R-AXIS	2.3	46244	15309	86	7.6	65	4.1
Se	R-AXIS	2.6	29621	10636	87	7.3	65	3.8
U	R-AXIS	2.8	40350	9657	93	6.9	89	3.3
W	Hamlin	3.3	23504	5571	86	9.3	61	6.7
PtE	Hamlin	4.0	8299	3538	95	8.7	69	9.8

Heavy-atom statistics

Derivative	R_{iso}‡	No. of sites	Resolution range (Å)	R_c§	fh/E¶ (acentrics)	fh/E¶ (centrics)	Anomalous data used?
U	0.19	2	20.0–3.0	0.55	1.55	1.10	Yes
W	0.24	4	10.0–3.3	0.67	1.09	0.69	No
PtE	0.18	3	20.0–4.0	0.59	1.30	0.85	No

Refinement statistics

R factor**	0.203
Resolution (Å)	8.0–2.3
No. of reflections included	14840
No. of protein atoms	1140
No. of sulfates	2
No. of waters	108

R.m.s. deviation from ideality

Bond lengths (Å)	0.016
Bond angles (°)	2.95
Dihedrals (°)	19.7
Impropers (°)	1.29

*$R_{merge} = 100 \times \sum_h \sum_i (|F_{hi}^2 - \langle F_h^2 \rangle|) / \sum_h \sum_i F_{hi}^2$, where F_{hi}^2 is the square of the ith intensity measurement of reflection h and $\langle F_h^2 \rangle$ is the mean squared intensity of the reflection.

† The final shell is defined as the resolution range from HR to HR + 0.1 Å, where HR is the highest resolution in Å.

‡ R_{iso} = Mean fractional isomorphous difference $= \sum |F_{PH} - F_P| / \sum F_P$, where F_{PH} and F_P are the derivative and native structure factors, respectively.

§ $R_c = \sum |F_{obs}^H - F_{calc}^H| / \sum F_{obs}^H$, where F_{obs}^H and F_{calc}^H are the observed and calculated heavy-atom structure factors, respectively.

¶ Phasing power $= f_H / E$, where f_H is the root-mean-square calculated heavy-atom structure factor and E is the root-mean-square lack-of-closure error.

** *R* factor $= \sum F_{obs} - F_{calc} / \sum F_{obs}$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

$F_o - F_c$ maps as well as the experimental map. This interpretation was subsequently confirmed by obtaining diffraction data from crystals grown using ammonium selenate as the precipitant (see below). The model was eventually refined to an *R* factor of 20.0 using the room-temperature data from 8 to 2.8 Å resolution. This model had good geometry and contained no water molecules (not shown).

At this point, further refinement against the room-temperature data was suspended, since we were able to obtain 2.3 Å data at low temperature. The refined room-temperature model was used as the starting point for refinement against the low-temperature data. The identity of the sulfate ion was confirmed and additional sites added by collecting data on crystals grown using ammonium selenate as the precipitant rather than ammonium sulfate. The crystals grown in ammonium selenate had the same space group as those grown in ammonium sulfate and had only slightly altered cell dimensions of $a = b = 84.9$ and $c = 102.9$ Å at 123 K. Three peaks greater than 5 standard deviations above the mean density were found in an $F_{Se} - F_S$ difference Fourier map. Sulfate ions were placed at these positions and included in the refinement. The *B* factors for one of the sulfate ions became quite large, probably reflecting a low occupancy and consequently was deleted from the model. In addition, the placement of residues 180 and 181 was deemed unreliable and these residues were deleted from the model. The final refinement statistics against the low-temperature data are shown in Table 1. The final model includes 140 protein residues, 108 water molecules and two sulfate ions. The density is weak and broken between helices *A* and *B* so that coordinates for residues 77–81 are not well defined and these residues are merely included in the final model for completeness. The large number of waters seen is probably due to increased solvent order at the low temperature used for data collection. While the final model has a relatively high average *B* factor of 36 for all protein atoms, the high overall *B* factor is consistent with a Wilson plot (not shown). This may reflect disorder in the crystal, rather than large thermal fluctuations. All residues in

the final refined structure are in allowed regions of the Ramachandran plot (Fig. 2) (Morris, MacArthur, Hutchinson & Thornton, 1992) and a profile window plot remains above 0.24 for the entire structure (Fig. 3) (Luthy, Bowie & Eisenberg, 1992).

Structures and structure alignments

Coordinates for the disulfide cross-linked *S. typhimurium* receptors in the unliganded and ligand-bound forms were kindly provided by Sung-Ho Kim. Structural alignments were performed using the program *ALIGN* (Satow, Cohen, Padlan & Davies, 1986). During the alignment optimization equivalent pairs greater than three times the current root-mean-square (r.m.s.) deviation are rejected.

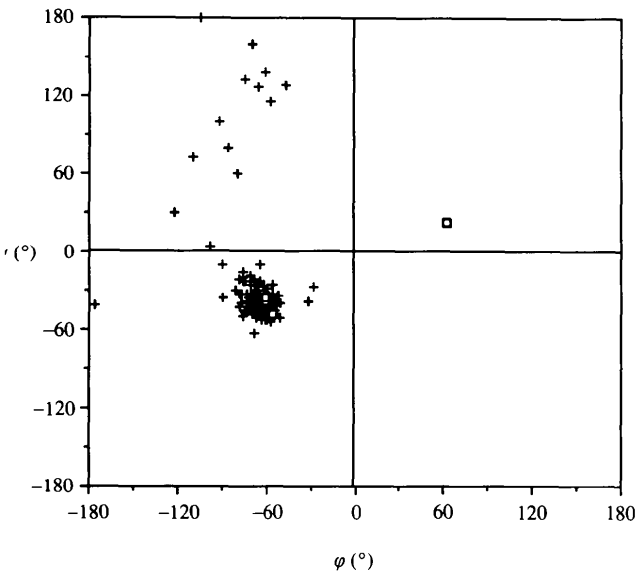


Fig. 2. The Ramachandran plot. Glycine residues are indicated by □ and non-glycines are indicated by +.

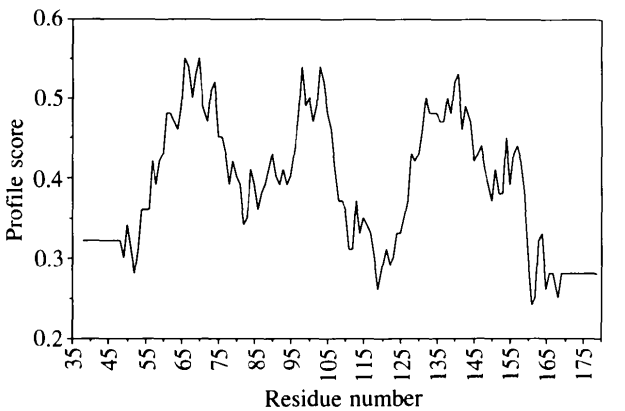


Fig. 3. The profile window plot. The average three-dimensional profile score in a 21-residue window is plotted using environments calculated from the aspartate receptor dimer. The three-dimensional profile score is a measure of the compatibility of a sequence with a structure (Luthy *et al.*, 1992).

Results and discussion

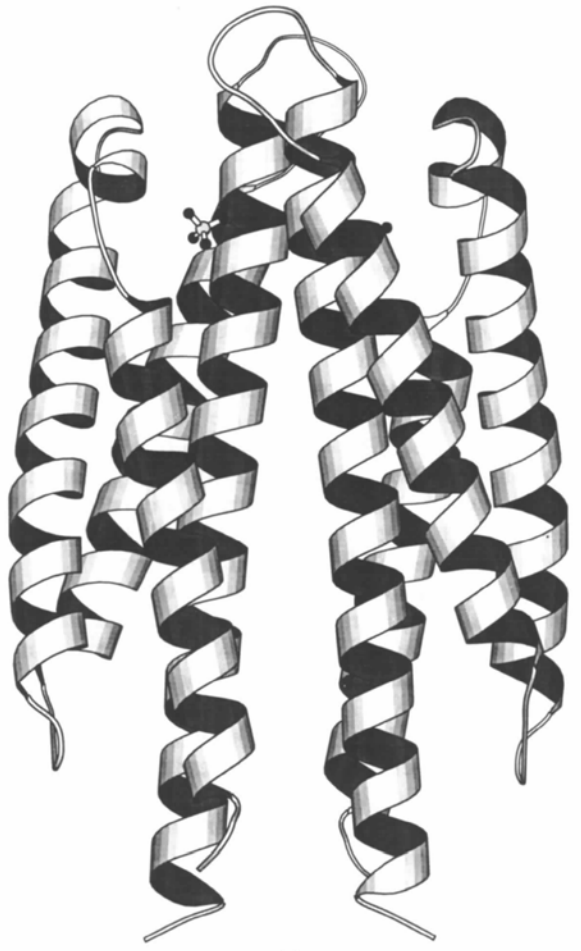
The structure

The structure of the aspartate receptor dimer is shown in Fig. 4. The monomer consists of four main helices, *A* (residues 40–75), *B* (residues 86–109), *C* (residues 117–141) and *E* (residues 154–176), and a short helix, *D* (residues 146–151), is found in the connection between helices *C* and *E*. The helical secondary structure of helices *A* and *E* breaks down at their termini. This is probably an artifact of truncation since there is strong evidence from disulfide cross-linking and mutation studies that the helices continue through the membrane (Lynch & Koshland, 1991; Pakula & Simon, 1992). The four main helices of the monomer are packed together in a commonly occuring structural motif called a four-helix bundle (Weber & Salemme, 1980). The interhelix packing angles within a single subunit are listed in Table 2 are reasonably close to expected values for four-helix bundle structures (Weber & Salemme, 1980).

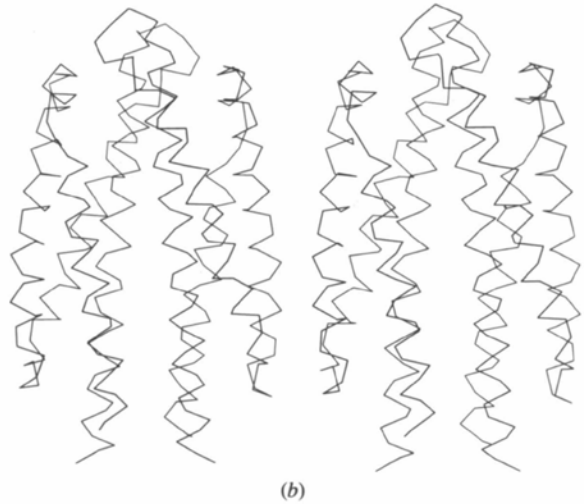
The dimer is made from two monomers that are related by a crystallographic twofold axis. A total of area of 1091 Å^2 is buried in the dimer interface and almost 90% of this surface is contributed by residues in helix *A* (716 Å^2) and helix *E* (254 Å^2). Thus, the main dimer contacts are made between the long helices *A* and *E* and the four central helices of the dimer are also organized in a four-helix bundle-like arrangement. With the exception of the *A–E'* packing interaction, the interhelix packing angles are not unusual (Table 2).

While our crystals were grown in the absence of known effector molecules, significant electron density was observed in the aspartate-binding pocket and was interpreted as a sulfate ion derived from the crystallization solution. This interpretation was later confirmed by collecting data from crystals grown in ammonium selenate rather than ammonium sulfate. A large peak of electron density was seen in the aspartate-binding pocket in an $F_{Se} - F_S$ difference Fourier map. As shown in Fig. 5, sulfate binding involves a large network of direct or water-mediated hydrogen bonds and salt bridges to side chains from both subunits of the dimer. The residues making direct or water-mediated hydrogen bonds to the sulfate ion are Arg64, Thr68, Tyr149, Gly152, Thr154 and Gln155 from one subunit and Arg69 and Arg70 from the other subunit. Indeed all of the significant hydrogen bonds between protein and aspartate seen in the AR-St-asp structure are substituted in the AR-Ec structure by hydrogen bonds to sulfate (compare Fig. 4 of Milburn *et al.*, 1991). It is possible then, that sulfate may mimic the effects of a bound aspartate, although the precise geometry of the active site is altered (see below).

We were unable to observe binding of known effectors of the AR-Ec. Crystals were grown under the same conditions used to grow the unliganded crystals, but in the presence of either 1 m*M* Asp or 100 m*M* $CoCl_2$. Crystals grown in the presence of either effector were identical



(*a*)



(*b*)

Fig. 4. The AR-Ec dimer structure. (*a*) A ribbon drawing showing the sulfate bound in the aspartate-binding pocket in a ball-and-stick representation. (*b*) A stereoview of the Cα trace. The figure was produced using the program *MOLSCRIPT* (Kraulis, 1991).

Table 2. *Interhelix packing angles in the AR-Ec structure*

The interhelix packing angle Ω is defined by Chothia, Levitt & Richardson (1981). Helices are labeled as described in the text. Deviations from ideality are given in standard deviation units based on mean and standard deviations for four-helix bundle structures reported by Weber & Salemme (Weber & Salemme, 1980). Converting to the Chothia *et al.* (Chothia *et al.*, 1981) definition of Ω, adjacent helices were found to have a mean interhelix packing angle of 18° and a standard deviation of 5.8°. For diagonally related helices a mean of 22° and a standard deviation of 7.0° was reported.

Helix pair	Ω(°)	Deviation from average
A–B	18	0
A–C	22	0
A–E	17	0.17
B–C	20	0.34
B–E	33	1.57
C–E	29	1.00
A–A′	13	0.86
A–E′	6	2.06
E–E′	22	0

to those grown in their absence and no significant peaks were observed in difference-density maps. The lack of aspartate binding probably reflects decreased affinity of the receptor for aspartate at the low pH of 3.6 necessary for crystal formation and competition with the high concentration of sulfate. The lack of a detectable cobalt-binding site may be due to the crystallization conditions or the requirement for some unidentified ancillary factor since direct binding of cobalt to the aspartate receptor has not yet been demonstrated.

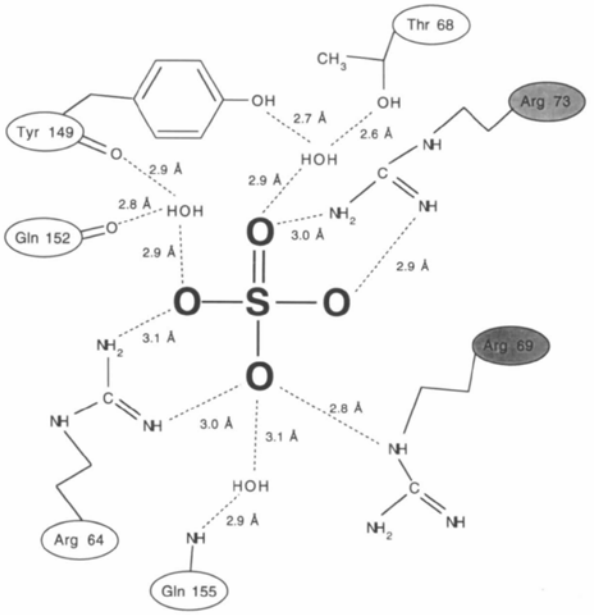


Fig. 5. Hydrogen bonding between the protein and sulfate. Residue labels from one subunit are indicated by the shaded ovals and residues from the other subunit are indicated by the unshaded ovals.

Comparison with the Salmonella aspartate receptor structure

Because the periplasmic domains of the *S. typhimurium* and *E. coli* aspartate receptors share 66% residue identity (amino acids 33–189), it is expected that the structures should also be quite similar (Chothia & Lesk, 1986). Indeed, the coordinates of one subunit of AR-Ec can be aligned to AR-St-apo coordinates with an r.m.s. deviation of only 0.54 Å on main-chain atoms (excluding residues 38-43, 77-87 and 177-179). The r.m.s. deviations of the AR-Ec main-chain coordinates to the two non-equivalent subunits in AR-St-asp are only 0.47 Å (excluding residues 38–43, 76–82 and 176–179) and 0.73 Å (excluding residues 38–42, 76–86 and 179). Fig. 6(*a*) shows the Cα distance separations of the aligned *S. typhimurium* structures and the *E. coli* structure. The only significant deviations occur in the long loop region between helices *A* and *B*, for which there is only weak density in our maps, and at the amino and carboxy termini where the helices break down in our structure.

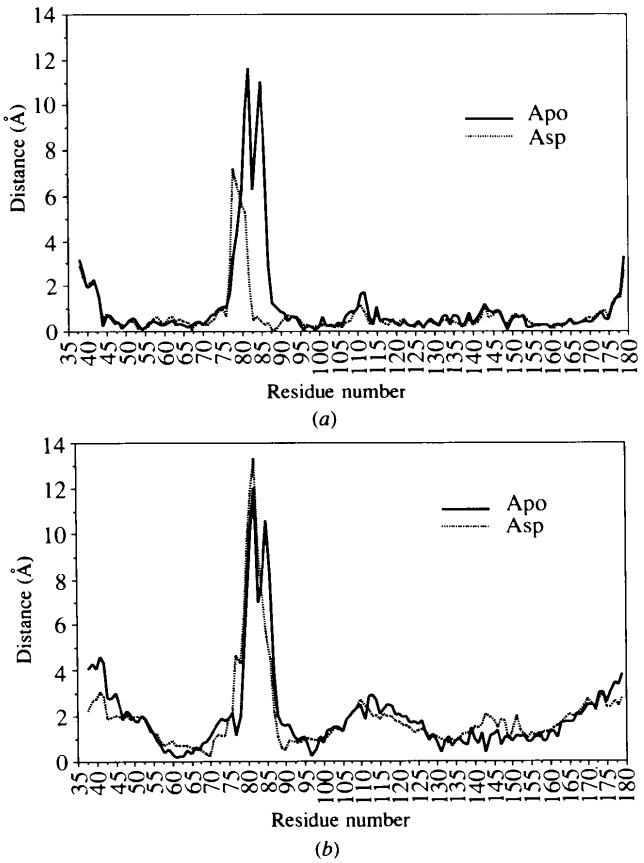


Fig. 6. Comparison of AR-Ec structure to the AR-St-apo and AR-St-asp structure. (*a*) Distances between equivalent Cα atoms in the AR-Ec structure and the AR-St-apo (solid line) and the AR-St-asp (dashed line) structures for the subunits that were optimally aligned. (*b*) Distances between equivalent Cα atoms in the AR-Ec structure and the AR-St-apo (solid line) and AR-St-asp (dashed line) structures for one subunit, when the other subunit is optimally aligned.

The only significant structural difference seen between AR-St-apo and AR-St-asp was a slight rotation (4°) of the subunits with respect to one another (Milburn *et al.*, 1991; Scott *et al.*, 1993; Yeh *et al.*, 1993). If a single subunit of the AR-Ec structure is aligned to a single subunit of the AR-St-apo and AR-St-asp structures and the coordinates of the second subunit are compared, more significant deviations are apparent (see Fig. 6*b*). Thus, while the structure of the individual subunits are quite similar, there appears to be less structural conservation in the relationship of the subunits with respect to one another.

The differences between the structures were examined more closely by looking for alterations in the relationships of individual helices. Each helix from the AR-Ec structure was aligned to the corresponding helix from the AR-St-apo and AR-St-asp structures and the r.m.s. deviation of all the corresponding helices in the dimer structures determined (Table 3). In general, deviations between helices within each subunit were smaller than deviations between helices in separate subunits. The largest deviations seen within a subunit occur between AR-Ec and subunit 1 of the AR-St-asp structure, particularly in the relationship of helices *B* and *E* and the relationship of helices *A* and *E*. The alteration in the relationship of the *A* and *E* helices is noteworthy in that it alters the orientation of the helices connected to the transmembrane domains. The r.m.s. deviation of the *E* helices when the *A* helices are optimally aligned is 1.46 Å and the r.m.s. deviation between the *A* helices when the *E* helices are optimally aligned is 1.58 Å. The difference in the relationship of the helices in the two structures can be approximated as a slight upward movement away from the membrane of helix *E* in the AR-St-asp structure relative to helix *E* in the AR-Ec structure.

The difference in orientation of the *A* and *E* helices noted above lead us to perform a similar analysis of the relationship of individual helices between the AR-St-apo and AR-St-asp structures (Table 4). As described by Milburn *et al.*, the largest deviations seen are between helices of different subunits (Milburn *et al.*, 1991). The deviations within subunits were again somewhat larger for subunit 1 of the AR-St-asp structure although the changes are quite small. The relative motion of helices *A* and *E* can be described as a small movement of helix *E* in the AR-St-asp structure downward toward the membrane relative to helix *E* in the AR-St-apo structure. It must be emphasized that this movement is only of the order of 1 Å, however. It is not striking.

Comparison of the aspartate-binding pockets

As described above, the aspartate-binding pocket in the AR-Ec structure is occupied by a sulfate ion which makes qualitatively similar contacts to the protein as aspartate in the AR-St-asp structure. A significant dif-

Table 3. *Comparison of AR-Eco and AR-St structures*

Numbers in the table are the root-mean-square deviations of Cα positions for the helix comparisons listed at the top of each column, after optimal alignment of individual helices listed at the left of each row. The subscripts eco, apo and asp refer to helices from the AR-Ec, AR-St-apo and AR-St-asp structures, respectively. Different subunits are distinguished by the numbers 1 and 2. Secondary structures were defined by the program *DSSP* and the longest possible helical regions from both structures were used in the alignments. For helix *A*, residues 40–75 were used. For helix *B* residues 88–108, for helix *C*, residues 117–141 and for helix *E*, residues 154–176 were used.

	Helices compared							
Helices aligned	*A*–*A*1	*B*–*B*1	*C*–*C*1	*E*–*E*1	*A*–*A*2	*B*–*B*2	*C*–*C*2	*E*–*E*2
A_{eco}–$A1_{apo}$	0.84	0.70	1.05	1.10	1.78	1.36	1.72	1.86
B_{eco}–$B1_{apo}$	1.15	0.42	0.86	1.05	2.09	2.23	1.59	2.10
C_{eco}–$C1_{apo}$	1.12	1.30	0.43	0.78	2.14	2.59	2.22	2.53
E_{eco}–$E1_{apo}$	0.94	1.10	0.64	0.34	1.97	1.90	1.57	2.18
A_{eco}–$A1_{asp}$	0.82	0.64	0.98	1.46	1.54	0.80	1.17	1.73
B_{eco}–$B1_{asp}$	1.11	0.43	1.01	1.55	1.98	1.83	1.66	2.01
C_{eco}–$C1_{asp}$	1.35	1.01	0.29	0.73	2.11	2.24	2.30	2.14
E_{eco}–$E1_{asp}$	1.58	1.83	1.03	0.37	2.59	2.61	2.85	2.62
A_{eco}–$A2_{asp}$	1.41	0.65	0.96	1.74	0.83	0.57	0.86	1.00
B_{eco}–$B2_{asp}$	1.88	1.76	1.64	2.21	1.04	0.28	0.61	0.79
C_{eco}–$C2_{asp}$	2.16	2.14	2.20	2.45	0.94	0.79	0.34	0.64
E_{eco}–$E2_{asp}$	2.02	1.60	1.77	2.14	1.02	1.09	0.75	0.30

Table 4. *Comparison of AR-St-apo and the AR-St-asp structures*

Numbers in the table are the root-mean-square deviations of Cα positions for the helix comparisons listed at the top of each column, after optimal alignment of individual helices listed at the left of each row. The subscripts apo and asp refer to helices from the AR-St-apo and AR-St-asp structures, respectively. Different subunits are distinguished by the numbers 1 and 2. Secondary structures were defined by the program *DSSP* and the helical regions from the AR-ST-apo structure were used for alignment. For helix *A*, residues 33–75 were used. For helix *B* residues 88–108, for helix *C*, residues 117–142 and for helix *E*, residues 154–177 were used.

	Helices compared							
Helices aligned	A_{apo}–$A1_{asp}$	B_{apo}–$B1_{asp}$	C_{apo}–$C1_{asp}$	E_{apo}–$E1_{asp}$	A_{apo}–$A2_{asp}$	B_{apo}–$B2_{asp}$	C_{apo}–$C2_{asp}$	E_{apo}–$E2_{asp}$
A_{apo}–$A1_{asp}$	0.45	0.51	0.72	0.91	1.02	1.11	0.75	0.53
B_{apo}–$B1_{asp}$	0.64	0.30	0.51	0.84	1.36	1.31	1.06	0.84
C_{apo}–$C1_{asp}$	1.28	0.65	0.30	0.78	1.92	2.15	1.57	1.47
E_{apo}–$E1_{asp}$	1.18	0.99	0.77	0.18	1.74	2.37	2.10	1.23
A_{apo}–$A2_{asp}$	1.18	1.96	2.22	1.77	0.35	0.73	1.00	0.95
B_{apo}–$B2_{asp}$	1.36	1.30	1.58	1.76	0.66	0.45	0.38	0.60
C_{apo}–$C2_{asp}$	1.31	1.08	1.34	1.73	0.60	0.67	0.22	0.49
E_{apo}–$E2_{asp}$	0.91	1.07	1.11	1.21	0.61	0.73	0.45	0.21

ference between the sulfate-bound and aspartate-bound crystals, is that sulfate binds to both binding pockets with perfect symmetry. In contrast, aspartate was found to bind to only one of the two binding sites in the *S. typhimurium* receptor (Yeh *et al.*, 1993). Given the apparently favorable binding pocket for sulfate ions, it is interesting that no sulfate binding is observed for any of the unliganded *S. typhimurium* structures even though sulfate was present in the crystallization conditions. We looked for differences between the aspartate-binding pockets of AR-Ec and the AR-St-apo and AR-St-asp structures by aligning the structures using only residues that surround the binding pockets. The deviations of atomic positions of key residues in the aligned structures are given in Table 5. The only large structural changes occur for the side chains, while only small shifts occur in backbone atoms. Why the side chains cannot then simply rearrange to accommodate a sulfate in the case of the AR-St-apo structure is not at all apparent.

Maltose-binding protein

A major difference between the aspartate receptors from *S. typhimurium* and *E. coli* is that the *E. coli* receptor responds to maltose by binding to a maltose–maltose-binding protein complex. Gardina, Conway, Kossman & Manson, have identified a number of mutants of the *E. coli* receptor that do not respond to maltose, though they are still able to respond to aspartate (Gardina, Conway, Kossman & Manson, 1992). Fig. 7 shows where these mutations lie on the AR-Ec structure along with the superimposed AR-St-apo structure. Many of these mutations occur in the most divergent region of the structure, the loop between helix *A* and helix *B*. The residues in this loop have high temperature factors, however, suggesting that this part of the structure is quite flexible and could easily adopt a different structure in the complex with maltose-binding protein. Indeed the conformation of this loop is substantially altered in the different subunits of the AR-St-asp and AR-St-apo structures. The remainder of the mutations occur in a relatively well ordered loop region between helix *C* and helix *D* where there is only small divergence between the AR-Ec receptor and the different AR-St receptor subunits. The largest deviation on Cα positions in this region is 1.5 Å and occurs at position 143 when the AR-Ec structure and subunit 1 of the AR-St-asp structure are compared. Thus, the difference in

Table 5. *Comparison of aspartate-binding sites in the E. coli and S. typhimurium aspartate receptor structures*

The table lists deviations in atomic positions on residues comprising the aspartate-binding site after superposition of the structures in the region of the binding site. The residues listed contact aspartate in the AR-St-asp structure or sulfate in the AR-Ec structure either directly or by water-mediated hydrogen bonds. Residues 64–68 and 149–155 from one subunit and residues 69–73 in the other subunit were included in the alignments. Site 1 refers to the occupied site and site 2 refers to the unoccupied site in the Ar-St-asp structure.

Residue	Atom name	AR-Ec to AR-St-apo comparison Deviation (Å)	AR-Ec to AR-St-asp site 1 comparison Deviation (Å)	AR-Ec to AR-St-asp site 2 comparison Deviation (Å)
Arg64	N	0.14	0.33	0.23
	CA	0.27	0.44	0.30
	C	0.32	0.34	0.36
	O	0.44	0.38	0.47
	CB	0.33	0.72	0.28
	CG	0.55	1.23	0.37
	CD	0.41	0.92	0.43
	NE	1.22	0.44	0.46
	CZ	0.60	2.11	0.34
	NH	2.21	3.63	2.45
	NH	1.99	3.17	2.21
Thr68	N	0.37	0.46	0.36
	CA	0.55	0.46	0.35
	C	0.61	0.42	0.45
	O	0.65	0.50	0.37
	CB	0.65	0.51	0.47
	OG	1.12	0.93	1.87
Tyr149	N	0.14	0.35	0.36
	CA	0.32	0.38	0.40
	C	0.34	0.59	0.29
	O	0.59	0.74	0.23
	CB	0.40	0.40	0.51
	CG	0.42	0.56	0.47
	CD	0.53	1.07	0.44
	CE	0.49	1.17	0.36
	CD	0.31	0.38	0.45
	CE	0.28	0.40	0.41
	CZ	0.37	0.76	0.35
	OH	0.43	0.83	0.27
Gln152	N	0.22	0.69	0.40
	CA	0.18	0.43	0.38
	C	0.22	0.42	0.42
	O	0.31	0.65	0.55
	CB	0.11	0.45	0.59
	CG	0.39	0.22	1.16
	CD	0.32	0.36	0.56
	OE	1.91	0.51	1.23
	NE	2.50	0.63	2.89
Thr154	N	0.35	0.45	0.20
	CA	0.46	0.39	0.26
	C	0.50	0.53	0.26
	O	0.54	0.49	0.35
	CB	0.44	0.56	0.33
	OG	0.43	0.62	0.41
	CG	0.39	0.53	0.29
Gln155	N	0.49	0.67	0.23
	CA	0.52	0.77	0.32
	C	0.53	0.79	0.40
	O	0.44	0.73	0.54
	CB	0.61	0.93	0.49
	CG	0.76	1.19	0.66
	CD	0.54	1.44	0.30
	OE	0.85	1.41	0.32
	NE	0.23	2.22	0.34
Arg69′	N	0.25	0.54	0.34
	CA	0.33	0.65	0.37
	C	0.33	0.60	0.35
	O	0.33	0.67	0.35
	CB	0.49	0.75	0.66
	CG	0.55	0.82	0.21
	CD	0.64	0.94	0.56
	NE	1.43	1.41	1.56
	CZ	1.00	1.78	0.76
	NH	1.66	1.83	1.11
	NH	0.81	2.20	1.71
Arg73′	N	0.70	0.73	0.60
	CA	0.80	0.82	0.63
	C	0.86	0.95	0.63
	O	1.11	1.13	0.69
	CB	0.64	0.59	0.58
	CG	1.20	0.72	0.87
	CD	1.27	0.74	0.96
	NE	1.42	0.67	1.21
	CZ	2.43	0.76	1.06
	NH	3.03	2.93	2.68
	NH	3.73	1.81	2.30

the ability of the two receptors to respond to maltose may simply be due to the identity of residues in the binding surface of the protein, rather than differences in the structures. Therefore, it may be possible to convert the non-responding *S. typhimurium* receptor to one that responds to maltose by changing a limited number of residues on the surface of the protein. It is somewhat surprising that many of the critical binding residues are located in flexible regions of the structure since binding would entail a large loss of configurational entropy. This may be a mechanism whereby maltose-binding protein can recognize a large number of residues on the aspartate receptor surface, thereby increasing specificity, without increasing affinity to unreasonable levels.

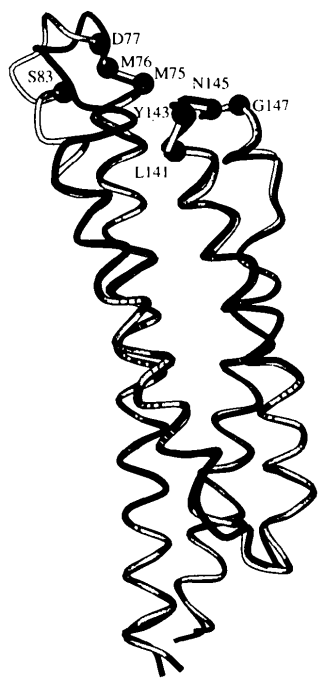


Fig. 7. Location of mal⁻ mutations. The structure of AR-Ec (light grey) is shown superimposed on the structure of AR-St-apo (dark grey). Position of the mal⁻ mutations are indicated on the AR-Ec structure by black balls at their Cα positions. The figure was produced using the program *MOLSCRIPT* (Kraulis, 1991)

Concluding remarks

In the five different aspartate receptor crystal forms grown in different conditions, the *E. coli* receptor described here, two wild-type and two mutant structures of the *S. typhimurium* receptors, little structural variation has been noted within the subunits. This is consistent with the signalling model suggested by the structures of Milburn *et al* (1991) and Yeh *et al.* (1993) in which the monomer behaves as a relatively rigid unit during the conformational change. This model, however, contradicts the biochemical data available (Milligan & Koshland, 1991; Stoddard *et al.*, 1992) which indicate that signalling can occur *via* conformational changes within a single subunit and does not involve important conformational changes between subunits. We do note small alterations in the positioning of helices *A* and *E* within a single subunit, but the structural differences are subtle at best. It seems unlikely that crystal packing forces or crystallization conditions are forcing a particular conformation because the similar structures are obtained across a variety of space groups, crystallization conditions and species. It is possible, however, that the true unliganded conformation is stable only in the context of the intact structure and when the protein is truncated, only the liganded conformation is stable. In this case, the crystal structures would merely represent slight variations of the same conformation. So far, however, attempts to crystallize the intact protein have been unsuccessful.*

The authors would like to thank Duilio Cascio and Kam Zhang for technical help, Gil Prive for careful reading of the manuscript and Sung-Ho Kim for providing coordinates. JB is particularly indebted to Chris Hill for his patient teaching and expert advice throughout this project; and to David Eisenberg, in whose laboratory all the crystallographic work was performed. This work was supported by National Institutes of Health grant GM-31299. JB and AP were both supported by postdoctoral fellowships from the American Cancer Society.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 2ASR, R2ASRSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: AM0018). A list of deposited data is given at the end of this issue.

References

BORKAVICH, K., KAPLAN, N., HESS, J. & SIMON, M. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 1208–1212.

BRÜNGER, A. (1990). *X-PLOR. Version* 2.1. Yale Univ., New Haven, CT, USA.

CHOTHIA, C. & LESK, A. (1986). *EMBO J.* **5**, 823–826.

CHOTHIA, C., LEVITT, M. & RICHARDSON, D. (1981). *J. Mol. Biol.* **145**, 215–250.

COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4 (1994). *Acta Cryst.* D**50**, 760–763.

GARDINA, P., CONWAY, C., KOSSMAN, M. & MANSON, M. (1992). *J. Bacteriol.* **174**, 1528-1536.

HAMLIN, R. (1985). *Methods Enzymol.* **114**, 416–452.

HAZELBAUER, G. (1975). *J. Bacteriol.* **122**, 206–214.

KRAULIS, P. (1991). *J. Appl. Cryst.* **24**, 946–950.

KRIKOS, A., MUTOH, N., BOYD, A. & SIMON, M. (1983). *Cell*, **33**, 615–622.

LUTHY, R., BOWIE, J. & EISENBERG, D. (1992). *Nature (London)*, **356**, 83–85.

LYNCH, B. & KOSHLAND, D. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 10402–10406.

MACNAB, R. (1987). *Motility and Chemotaxis.* In *Escherichia Coli and Salmonella Typhimurium Cellular and Molecular Biology*, edited by F. NEIDHARDT, J. INGRAHAM, B. MAGASANIK, K. LOW, M. SCHAECHTER & H. UMBARGER, pp. 732–759. Washington, DC: American Society for Microbiology.

MANIATIS, T., FRITSCH, E. & SAMBROOK, J. (1982). *Molecular Cloning: a Laboratory Manual.* New York: Cold Spring Harbor Laboratory Press.

MILBURN, M., PRIVE, G., MILLIGAN, D., SCOTT, W., YEH, J., JANCARIK, J., KOSHLAND, D. & KIM, S.-H. (1991). *Science*, **254**, 1342–1347.

MILLIGAN, D. & KOSHLAND, D. (1991). *Science*, **254**, 1651–1654.

MORRIS, A., MACARTHUR, M., HUTCHINSON, E. & THORNTON, J. (1992). *Proteins Struct. Funct. Genet.* **12**, 345–364.

MOWBRAY, S., FOSTER, D. & KOSHLAND, D. (1985). *J. Biol. Chem.* **260**, 11711–11718.

PAKULA, A. & SAUER, R. (1989). *Proteins Struct. Funct. Genet.* **5**, 202–210.

PAKULA, A. & SIMON, M. (1992). *Nature (London)*, **355**, 496–497.

READER, R., TSU, W.-W., SPRINGER, M., GOY, M. & ADLER, J. (1979). *J. Gen. Microbiol.* **111**, 363–374.

SATOW, Y., COHEN, G., PADLAN, E. & DAVIES, D. (1986). *J. Mol. Biol.* **190**, 593–604.

SCOTT, W., MILLIGAN, D., MILBURN, M., PRIVE, G., YEH, J., KOSHLAND, D. & KIM, S.-H. (1993). *J. Mol. Biol.* **232**, 555–573.

STODDARD, B., BUI, J. & KOSHLAND, D. (1992). *Biochemistry*, **31**, 11978–11983.

TERWILLIGER, T. & EISENBERG, D. (1983). *Acta Cryst.* A**39**, 813–817.

WANG, E. & KOSHLAND, D. (1980). *Proc. Natl Acad. Sci. USA*, **71**, 7342–7346.

WEBER, P. & SALEMME, F. (1980). *Nature (London)*, **287**, 82–84.

YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). *Gene*, **33**, 103–109.

YEH, J., BIEMANN, H.-P., PANDIT, J., KOSHLAND, D. & KIM, S.-H. (1993). *J. Biol. Chem.* **268**, 9787–9792.

ZHANG, K. (1993). *Acta Cryst.* D**49**, 213–222.

ZHANG, K. & MAIN, P. (1990). *Acta Cryst.* A**46**, 377–381.