Protein Science (1992), *1*, 1363–1376. Cambridge University Press. Printed in the USA.

Solution structure of the phosphocarrier protein HPr from *Bacillus subtilis* by two-dimensional NMR spectroscopy

MICHAEL WITTEKIND,[1,4] PONNI RAJAGOPAL,[1] BRUCE R. BRANCHINI,[2] JONATHAN REIZER,[3] MILTON H. SAIER, JR.,[3] AND RACHEL E. KLEVIT[1]

[1] Department of Biochemistry, University of Washington, Seattle, Washington 98195
[2] Department of Chemistry, Connecticut College, New London, Connecticut 06320
[3] Department of Biology, University of California, La Jolla, California 92093

(RECEIVED July 23, 1992; REVISED MANUSCRIPT RECEIVED August 26, 1992)

Abstract

The solution structure of the phosphocarrier protein, HPr, from *Bacillus subtilis* has been determined by analysis of two-dimensional (2D) NMR spectra acquired for the unphosphorylated form of the protein. Inverse-detected 2D (^1H-^{15}N) heteronuclear multiple quantum correlation nuclear Overhauser effect (HMQC NOESY) and homonuclear Hartmann-Hahn (HOHAHA) spectra utilizing ^{15}N assignments (reported here) as well as previously published ^1H assignments were used to identify cross-peaks that are not resolved in 2D homonuclear ^1H spectra. Distance constraints derived from NOESY cross-peaks, hydrogen-bonding patterns derived from ^1H-^2H exchange experiments, and dihedral angle constraints derived from analysis of coupling constants were used for structure calculations using the variable target function algorithm, DIANA. The calculated models were refined by dynamical simulated annealing using the program X-PLOR. The resulting family of structures has a mean backbone rmsd of 0.63 Å (N, C^α, C′, O atoms), excluding the segments containing residues 45–59 and 84–88. The structure is comprised of a four-stranded antiparallel β-sheet with two antiparallel α-helices on one side of the sheet. The active-site His 15 residue serves as the N-cap of α-helix A, with its $N^\delta 1$ atom pointed toward the solvent to accept the phosphoryl group during the phosphotransfer reaction with enzyme I. The existence of a hydrogen bond between the side-chain oxygen atom of Tyr 37 and the amide proton of Ala 56 is suggested, which may account for the observed stabilization of the region that includes the β-turn comprised of residues 37–40. If the $\beta\alpha\beta\beta\alpha\beta(\alpha)$ folding topology of HPr is considered with the peptide chain polarity reversed, the protein fold is identical to that described for another group of $\beta\alpha\beta\beta\alpha\beta$ proteins that include acylphosphatase and the RNA-binding domains of the U1 snRNP A and hnRNP C proteins.

Keywords: histidine-containing protein; HPr; NMR; phosphoenolpyruvate:sugar transport system; phosphotransfer; protein structure

Histidine-containing protein (HPr) is an integral component of the PTS in bacteria. It serves as a phosphoryl group acceptor from enzyme I through an N^δ-phosphohistidine intermediate and as a phosphoryl group donor to the enzymes IIA (Saier & Reizer, 1992). HPrs from a variety of both gram-positive and gram-negative bacteria have been characterized. They are small (85–90 residues), soluble, acidic, monomeric proteins and thus are quite amenable to structural studies using 2D NMR spectroscopy.

Fairly complete ^1H resonance assignments have been reported for *Escherichia coli* HPr (Klevit et al., 1986;

Reprint requests to: Rachel E. Klevit, Department of Biochemistry, University of Washington, Seattle, Washington 98195.

[4] Present address: Macromolecular NMR Department, Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, New Jersey 08543.

Abbreviations: H-bond, hydrogen bond; HMQC, heteronuclear multiple quantum correlation spectroscopy; hnRNP, heteronuclear ribonucleoprotein; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; HPr, histidine-containing protein; INEPT, insensitive nucleus enhanced polarization transfer; NMR, nuclear magnetic resonance spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; P.E.COSY, primitive exclusive coherence spectroscopy; PTS, phosphoenolpyruvate:sugar transport system; rmsd, root mean squared deviation; snRNP, small nuclear ribonucleoprotein; TOCSY, total coherence spectroscopy; TPPI, time-proportional phase incrementation; 2D, two-dimensional.

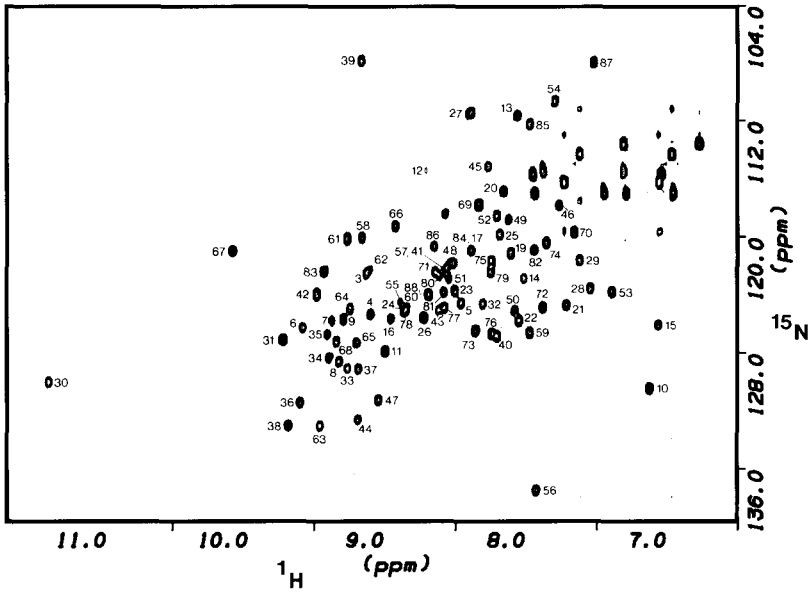

Fig. 1. Contour plot of the fingerprint region of an ^{15}N-^{1}H HMQC-HOHAHA spectrum of uniformly ^{15}N-labeled HPr from *Bacillus subtilis*. Experimental details are described in Materials and methods. Individual cross-peaks are labeled with their sequential assignments (listed in Table 1). The unlabeled peaks arise from side-chain ^{15}NHs of glutamine and asparagine residues.

Hammen et al., 1991), *Bacillus subtilis* HPr (Wittekind et al., 1990), and *Staphylococcus aureus* HPr (Kalbitzer et al., 1991), which share 33-63% sequence identity. A folding topology consisting of a four-stranded antiparallel β-sheet (topology A-D-B-C) and two or three antiparallel α-helices that cross one face of the sheet is indicated for the HPrs on the basis of qualitative analysis of observed NOE patterns (Klevit & Waygood, 1986; Wittekind et al., 1990; Hammen et al., 1991; Kalbitzer et al., 1991). However, the structure of *E. coli* HPr solved by X-ray diffraction has a significantly different topology (el Kabbani et al., 1987). Concurrent with the NMR structure determination reported here, the crystal structure of an active single-site mutant of *B. subtilis* HPr (S83C) was determined and has recently been reported (Herzberg et al., 1992). Unlike the earlier crystal structure, it has the folding topology described by the NMR studies.

As part of our program to gain insights into the effects of phosphorylation on HPr, we have determined the structure of *B. subtilis* HPr using all unambiguous NMR constraints from 2D spectra. The details of that structure are reported here.

Results

The ^{1}H assignments of *B. subtilis* HPr (88 residues), based on analysis of 2D ^{1}H NMR spectra, have been previously reported (Wittekind et al., 1990) and formed the basis for identification of cross-peaks in the homonuclear 2D ^{1}H NOESY spectra. However, spectral overlap resulted in ambiguities in the identification of the protons giving rise to some cross-peaks. In order to resolve ambiguities, HPr uniformly labeled with ^{15}N was prepared, and ^{15}N assignments were made from inverse-detected 2D ^{1}H-^{15}N-correlated spectra (HMQC, ^{1}H HMQC-NOESY, and ^{1}H HMQC-HOHAHA). The fingerprint region of an HMQC-HOHAHA spectrum labeled with the assignments is shown in Figure 1, and the ^{15}N assignments are listed in Table 1. These assignments allowed for the identification of additional NOESY cross-peaks in the ^{1}H, ^{15}N HMQC-NOESY.

Analysis of the heteronuclear spectra, plus additional ^{1}H homonuclear spectra such as TOCSY spectra collected in H_2O, resulted in additional ^{1}H assignments (for side-chain protons). During this analysis we also identified two incorrect sequential assignments in our original set: the spin system of Arg 17 and Pro 18 were switched, and the spin systems for Met 48 and Met 51 were switched.[5] A table containing all ^{1}H assignments for *B. subtilis* HPr is included as supplemental information on the Diskette Appendix.

Six 2D ^{1}H NOESY spectra with different mixing times were collected for a sample in H_2O. NOE intensity build-up rates were converted to target distances, and upper and lower bounds were assigned for cross-peaks that were well resolved (see Materials and methods). Analysis of the build-up rates made it possible to identify cross-peaks arising from spin-diffusion pathways, as characterized by a slow initial rate followed by a rapid increase at longer mixing times. Interactions giving rise to this class of peaks

[5] Following submission of this manuscript, we discovered that the overproduced version of *B. subtilis* HPr used in these studies as well as in the X-ray diffraction studies (Herzberg et al., 1992) actually contains a valine residue at position 51, rather than the methionine found in the natural sequence. This has been confirmed by both mass spectrometry and NMR (Reizer et al., unpubl.). Therefore, the chemical shifts are reported for the residue Val 51 in Table 1 and in the supplemental information on the Diskette Appendix. The structures described here were calculated with a methionine at position 51. However, because the calculations included no constraints involving the side chain of residue 51, the difference in its identity will have no effect on the structures reported here.

Table 1. *^{15}N chemical shifts and amide exchange rates for B. subtilis HPr at pH 6.9 and 30 °C*[a]

Residue	^{15}NH (ppm)	$J_{NH\alpha}$	$t_{1/2}$ (min)	Residue	^{15}NH (ppm)	$J_{NH\alpha}$	$t_{1/2}$ (min)
Q3	120.5			S46	115.8	9.0	
K4	123.6	7.6	36	I47	129.2	<3	
T5	122.5	9.1		M48	119.8	2.9	*
F6	124.2	9.7	55	G49	116.9		
K7	123.7	9.6	<30	V50	123.1	4.1	
V8	126.5	6.6	192	V51	120.8	4.8	
T9	123.6	10.4		S52	116.5	3.6	
A10	128.4	3.1	<20	L53	121.7	5.9	
D11	125.8	<3		G54	108.6		
S12	113.4			I55	122.5	5.7	
G13	109.6			A56	135.4	8.6	20
I14	120.8	9.6		K57	120.3	3.4	*
H15	124.0	9.2		G58	117.9		<20
A16	123.6	<3		A59	124.5	4.5	57
R17	119.0	3.2		E60	121.9	8.5*	<60*
P18				I61	118.0	9.8	888
A19	119.0	3.5	<20	T62	120.2	10.0	257
T20	114.8	4.3	<20	I63	131.0	10.3	365
V21	122.6	5.7	<20	S64	122.9	9.9	122
L22	123.7	4.0	88	A65	125.2	9.0	67
V23	121.6	5.1	433	S66	117.2	9.5	23
Q24	122.8	3.9	<195*	G67	118.9		
T25	117.7	5.6	29	A68	125.1		
A26	123.5	2.9	117	D69	115.7	9.6	
S27	109.4	3.6	39	E70	117.6	<3	41
K28	121.5	6.6		N71	120.4	5.9	*
Y29	119.5	9.6	99	D72	122.7	4.9	
D30	128.0	7.0		A73	124.4	3.3	533
A31	125.0	<3	19	L74	118.3	4.5	788
D32	122.5	7.9	25	N75	119.6	3.7	128
V33	127.0	9.6		A76	124.6	4.9	198
N34	126.2	9.8	210	L77	122.8	3.4	1,066
L35	124.6	9.6	578	E78	123.1	3.5	<1,500*
E36	129.4	10.0	578	E79	120.4	3.7	433
Y37	127.0	8.0	173	T80	120.7	3.9	270
N38	130.9	6.5		M81	121.8	2.8	248
G39	105.8			K82	118.8	5.6	169
K40	124.8	9.3	131	S83	120.3	4.3	<20
T41	120.0	10.7	*	E84	118.8	7.5	<20
V42	121.9	9.5	248	G85	110.2		
N43	123.0	5.0		L86	118.6	8.0	<40
L44	130.5	6.9	37	G87	105.9		<20
K45	113.1	8.0		E88	122.0	*	<60*

[a] Blank entries were not observed or not measured; * denotes exact values not obtained due to overlap.

were either not included in the structure calculation or given upper bounds of 7 Å. A total of 629 experimentally observed pairwise distance constraints were identified (see Materials and methods). A histogram showing the number, type, and distribution of these distance constraints is shown in Figure 2.

$^3J_{NH\alpha}$ coupling constants were determined for many residues of HPr from HMQC-J experiments and are listed in Table 1. These values, along with NOE information and $^3J_{\alpha\beta}$ (determined from P.E.COSY experiments), were used to determine allowed ranges for dihedral angles phi, psi, and chi I, as described in the Materials and methods. Amide protons protected from exchange with solvent were identified from a series of short 2D (^1H,^{15}N)-HMQC spectra obtained with a lyophilized sample of HPr freshly dissolved into D_2O. The fast HMQC experiments enabled the identification of additional exchange-protected amide protons from those originally observed in longer ^1H COSY experiments. Twenty-five slowly exchanging amides were identified from the earlier experiments at pH 6.9 and 30 °C (see Fig. 7; Wittekind et al., 1990), whereas 56 amide resonances were observed in spectra obtained during the 17 min immediately following dissolution of an HPr sample in D_2O (under the same conditions). Of these 56 amides, 26 have exchange half-times longer than 1 h. The measured exchange half-times are listed in Table 1. As illustrated in Figure 3, the protected amides are mostly in regions of regular secondary structure originally identified on the basis of NOE connectivity patterns (Wittekind et al., 1990). There are, however, exceptions: most notably, Ala 56, and residues at the extreme C-terminus. These will be discussed later in the context of the determined structure.

Eighteen hundred initial structures were calculated with DIANA, version 1.0 (Güntert et al., 1989), using both upper and lower bounds as distance constraints and dihedral angle constraints. Calculation of this relatively large number of structures was necessary in order to generate enough structures exhibiting low target function values. DIANA's inefficiency in arriving at the correct folding topology of HPr is most probably due to the high β-sheet content and the β-α-β topology of the protein. This problem is addressed in the most recent version of the program by employing the REDAC procedure (DIANA 2.0; Güntert & Wüthrich, 1991), which is reported to improve the rate of convergence in β-sheet proteins.

The DIANA structures were further refined by a dynamical simulated annealing procedure using the program X-PLOR (Brünger, 1990). Information about the 15 structures exhibiting the smallest angle and distance violations ({SA}) is summarized in Table 2. Deviations from idealized geometries are small. The structures have an average backbone (N, C^α, C′, O) rmsd of 0.81 Å over all residues and of 0.63 Å when only the backbone atoms of residues 2–44 and 60–83 are included. This value was obtained by calculating a pairwise backbone rmsd over these residues for each individual structure in {SA} versus the average structure $\overline{SA}$ and calculating the average of these values. An extensive minimization of $\overline{SA}$ that included the experimental constraints resulted in the minimized average structure SA_r. This structure has no distance bounds violations larger than 0.5 Å and has no angle constraint violations larger than 1°. The average magnitude of the distance violations for the minimized average structure is smaller than for any of the individ-

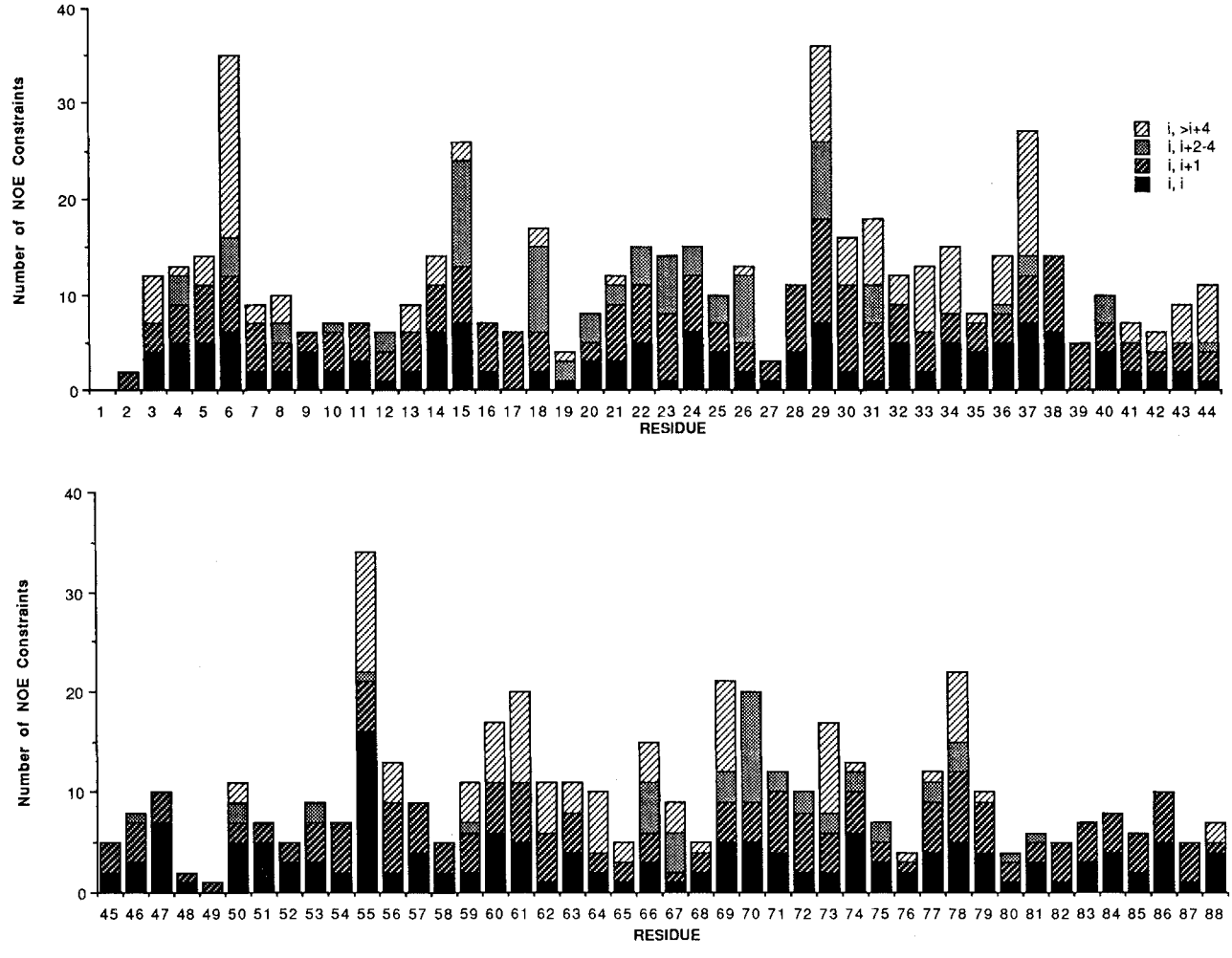

Fig. 2. Histograms showing the number and type of NOE constraints used for each residue in the structure determination. ■ Intraresidue *(i,i)* constraints; ▨ sequential $(i, i+1)$ constraints; ▩ medium-range $(i, i+2, 3, 4)$ constraints; ▨ long-range $(i, >i+4)$ constraints.

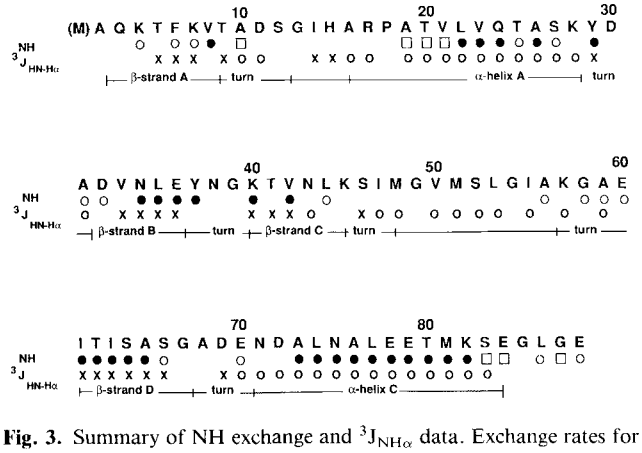

Fig. 3. Summary of NH exchange and $^3J_{NH\alpha}$ data. Exchange rates for individual amide protons are classified as: fast (unmarked), medium ($t_{1/2} < 20$ min, □); slow (20 min $< t_{1/2} <$ 60 min, ○); very slow ($t_{1/2} >$ 60 min, ●). $^3J_{NH\alpha}$ coupling constants are classified as: O, small ($^3J_{NH\alpha} < 6$ Hz); X, large ($^3J_{NH\alpha} > 9$ Hz). The elements of secondary structure indicated are those identified on the basis of a qualitative analysis of the NOEs, as described in Wittekind et al. (1990).

ual structures in {SA}, although the number of violations per structure is slightly higher.

Backbone atom representations of the 15 structures in {SA} are shown superimposed in Figure 4A and in Kinemage 1. Figure 4B shows a ribbon representation of the structure that highlights the elements of regular secondary structure. Comparison of this ribbon structure with that proposed on the basis of a qualitative analysis of the NOE data (Wittekind et al., 1990) shows that the models have identical topology. The structure is comprised of two well-defined α-helices (called helix A and helix C in Wittekind et al., [1990]) positioned antiparallel to one another on one face of a four-stranded antiparallel β-sheet, with topology A–D–B–C. The angle between the two helices is nearly 180° and the sheet exhibits a right-handed twist (see Kinemage 2). The structure of the β-sheet is well defined, as are the positions of the two α-helices on the face of the β-sheet. In the view shown, it is clear that the extreme C-terminus (upper left

Table 2. *Structural statistics*[a]

Structural statistics	{SA}	$\overline{SA}$	$\overline{SA}_r$
Root mean square deviations from experimental constraints (Å)			
All distance constraints[b] (705)	0.083 ± 0.003	0.129	0.064
Number of violations			
All distance constraints	105.3 ± 5.8	63	109
Distance constraints (above 0.5 Å)	1.2 ± 1.1	3	0
Angles above 1°	0 ± 0	0	0
F_{NOE} (kcal mol^{-1})[c]	238 ± 19	576	139
F_{TOR} (kcal mol^{-1})[c]	1.44 ± 0.46	1.00	1.18
F_{repel} (kcal mol^{-1})[d]	63.3 ± 3.5	13,477.5	63.3
E_{L-J} (kcal mol^{-1})[e]	−190.6 ± 13.7	3.4×10^7	−156.2
Mean rmsd from idealized geometry[f]			
Bonds (Å) (1,292)	0.005 ± 0.000	2.148	0.005
Angles (degrees) (2,343)	1.908 ± 0.008	25.219	1.896
Impropers (degrees) (458)	1.237 ± 0.015	7.642	1.234

Atomic rmsd (Å)[g]	Residues 2–44, 60–83	All residues
Backbone (N, Cα, C′)	0.599 ± 0.246	0.800 ± 0.457
Backbone (N, Cα, C′, O)	0.627 ± 0.266	0.814 ± 0.482
Side chains (heavy atoms)	1.118 ± 0.459	1.369 ± 0.813
All atoms (heavy atoms)	0.901 ± 0.349	1.121 ± 0.632
All atoms (including hydrogens)	1.081 ± 0.359	1.314 ± 0.677

[a] {SA} are the 15 final dynamical simulated annealing structures; $\overline{SA}$ is the mean structure derived by averaging the coordinates of the {SA} structures best fitted to one another (over backbone atoms of residues 2–44, 60–88); $\overline{SA}_r$ is the structure obtained by subjecting $\overline{SA}$ to 2,000 rounds of Powell minimization.

[b] All NOE constraints plus H-bond constraints. Each H-bond was enforced by two distance constraints as outlined in Materials and methods.

[c] Square well NOE (F_{NOE}) and torsion angle (F_{TOR}) potentials (see Equations 2 and 3 in Clore et al. [1986]) were calculated with force constants of 50 kcal mol^{-1} Å^2 and 200 kcal mol^{-1} rad^{-2}, respectively.

[d] The values of the van der Waals repulsion term F_{repel} are calculated with a force constant of 4 kcal mol^{-1} Å^{-4} with the hard sphere van der Waals radii set to 0.8 times the standard values used in the CHARMM empirical energy function (Brooks et al., 1983).

[e] E_{L-J} is the Lennard-Jones van der Waals energy term calculated with the CHARMM empirical energy function. It is not included in the target function for simulated annealing or restrained minimization.

[f] The force constant values for bond, angle, and improper torsions were set to 1,000 kcal mol^{-1} Å^{-2}, 500 kcal mol^{-1} rad^2, and 500 kcal mol^{-1} rad^{-2}, respectively. Improper torsion angles are used to maintain planarity and chirality.

[g] The atomic rms differences were obtained by calculating a pairwise backbone rmsd value over residues 2–44 and 60–83 for each individual structure in {SA} versus the average structure $\overline{SA}$ and calculating the average and SD values.

in this view) and the residues that link β-strands C and D (residues 45–59) are not as well defined relative to the rest of the protein.

Elements of regular secondary structure were identified according to their dihedral angles and H-bonding patterns, and the ends of helices were identified using the C^α criterion of Richardson and Richardson (1988). His 15 is the N-cap of α-helix A, which continues through Ser 27, and Asp 69 is the N-cap of α-helix C, which continues through Glu 84. Residues after Glu 84 lack medium size d_{NN} NOEs and have larger $^3J_{NH\alpha}$ coupling constants indicating a nonhelical or more dynamically averaged structure. Strand A of the antiparallel β-sheet begins near the N-terminus of the protein, at Gln 3, and extends to Val 8. Strand B, an inner strand of the β-sheet, begins at Asp 32 and leads into a reverse turn made up of residues Tyr 37, Asn 38, Gly 39, and Lys 40. The four residues immediately following the turn, Thr 41 through Asn 43, form strand C. Glu 60 through Ser 66 comprise strand D, an inner strand of the β-sheet.

The polypeptide backbone reverses direction six times (see Fig. 4): following β-strand A, following helix A, between strands B and C, following strand C, preceding strand D, and following strand D. Analysis of the structures in these regions revealed that only one of these regions corresponds to a described reverse turn geometry; residues 37–40 constitute a type I′ β-turn (Richardson, 1981). Residues 56–59 also assume a well-defined turn structure immediately preceding strand D, but its dihedral angles do not place it among common turn structures.

A superposition of the Ramachandran plots of the 15 structures in {SA} is shown in Figure 5A, and a plot for the minimized average structure, SA_r, is shown in Figure 5B. All non-glycine residues exhibit negative phi values except Asn 38 and Ser 12. Asn 38 is residue 2 of the turn linking strands B and C of the β-sheet. The measured $^3J_{HN\alpha}$ coupling constant for Asn 38 is 6.5 Hz and was interpreted as $\phi = 60 \pm 30°$. This decision was based on the relatively fast intraresidue HN-Hα NOE build-up rate (corresponding to a distance of ~2.4 Å) observed for this residue (Ludvigsen & Poulsen, 1992). In order to test this interpretation, 300 DIANA structures were calculated omitting the Asn 38 phi angle constraint. Of the 20 structures with target functions less than 100 Å^2, 17 had a positive phi value for Asn 38 between 0 and 100°. The other three structures in this set had local distortions in this region of the molecule, giving rise to multiple residual distance violations. These results indicate a positive ϕ value for Asn 38, consistent with its role as the second residue in a type I′ β-turn (Richardson, 1981). Ser 12 is in a turn connecting β-strand A with the active-site region at the N-terminus of the A helix. It is one of the least well constrained residues in the model (see Figs. 2 and 6, below).

Plots of rmsd versus residue are shown in Figure 6. The average rmsd in the backbone dihedral angle, ϕ, for the 15 structures is plotted in the upper panel. This plot illustrates that the backbone conformation is generally well defined by the NMR constraints throughout the structure (i.e., there is no long stretch of residues with significantly large deviations in ϕ). The isolated residues with large deviations in ϕ all correspond to residues for which $^3J_{NH\alpha}$

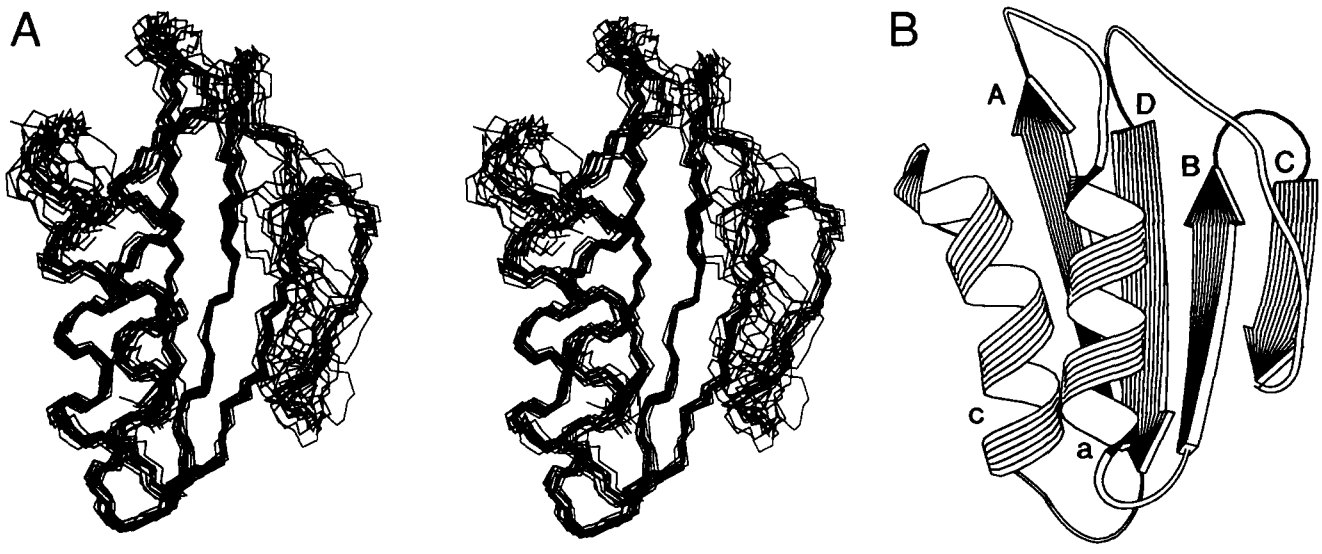

Fig. 4. A: Stereo view of the 15 structures that make up {SA} for *Bacillus subtilis* HPr. The entire polypeptide backbone (N, C^{α}, C′) for residues 2–88 is shown. Superposition was performed using residues 2–44 and 60–83. **B:** Ribbon representation of the minimized average structure (SA_r) of *B. subtilis* HPr. This representation highlights the elements of regular secondary structure, where the arrows represent β-strands and the curling ribbons represent α-helices. Letters A–D identify β-strands A–D, respectively, and letters a and c identify α-helices A and C. The two terminii of the protein are not clearly visible in this view; the N-terminus is at the beginning of β-strand A and the C-terminus is behind the end of α-helix C. The drawing was generated by the program RIBBON (J. Priestle).

could not be measured (see Table 1), and consequently no ϕ angle constraint was included for these residues. The average backbone position rmsd is plotted in the lower panel. This plot shows that the elements of regular secondary structure are relatively well ordered within the family of structures, whereas the segments comprised of residues 10–13, 44–54, and the C-terminus show larger deviations among the structures defined by the NMR data. The regions of large deviations correlate exactly to those residues for which no long-range NOEs were identified

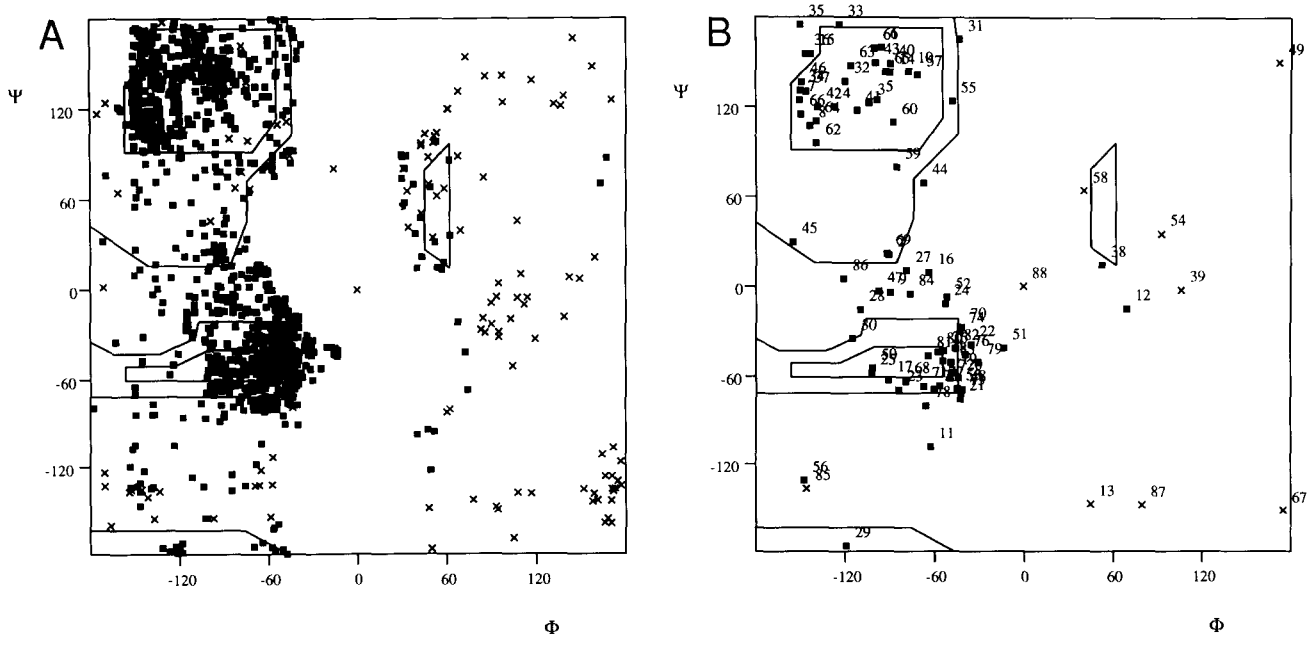

Fig. 5. A: Ramachandran plot for all residues in the 15 structures that make up {SA}. Glycine residues are indicated by ×. **B:** Ramachandran plot for all residues of the minimized average structure, SA_r.

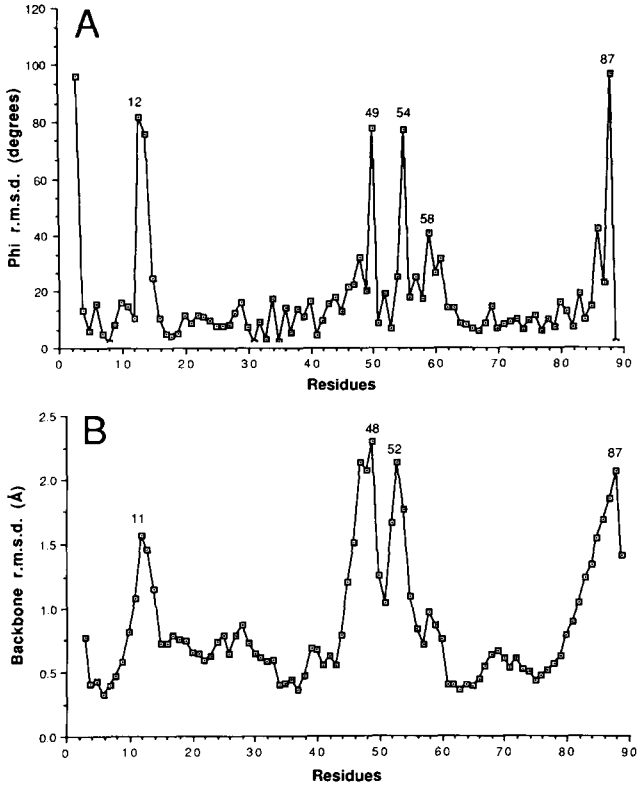

Fig. 6. Plots of local rmsd. **A:** Average pairwise rmsd for the dihedral angle, ϕ. **B:** Average pairwise rmsd from the average structure, $\overline{SA}$ for backbone heavy atoms (N, C^{α}, C', O).

(see Fig. 2). Thus, although the local conformations of these segments are fairly well defined (compare the upper panel), the *positions* of the segments relative to the rest of the protein structure are not defined by the NMR data.

The present structural models contain many well-defined side chains. There are 37 side chains with rmsd less than 1.0 Å. The majority of these well-defined side chains are in the interior of the protein, on the hydrophobic faces of the β-sheet and the α-helices. An example of the definition of the side chains of residues on the interior face of the β-sheet is illustrated in the superposition shown in Figure 7. Figure 8 and Kinemage 3 illustrate a close correspondence between the side-chain rmsd for a given residue and its surface accessibility. As shown in Figure 9, the imidazole ring of His 15 is well defined, with an average rmsd of 0.94 Å, even though it is on the surface of the protein. The definition of His 15 is due to NOE interactions between the $C^{\epsilon}H$ of His 15 and the $C^{\gamma}H$ and $C^{\delta}H$ resonances of Pro 18, which is constrained by virtue of its participation in α-helix A.

Discussion

We have previously described the folding topology of HPrs from *E. coli* (Klevit & Waygood, 1986) and from *B. subtilis* (Wittekind et al., 1990) based on qualitative analysis of NOE data. Recently, we reported a semiquantitative structure determination of *E. coli* HPr, based on a limited set of constraints (Hammen et al., 1991). Here we describe a more detailed and quantitative structure determination for *B. subtilis* HPr, based on a more complete set of experimental constraints.

The structural model that emerged has a global fold that has been described as an "open-faced sandwich" of two α-helices positioned against a four-stranded antiparallel β-sheet (Richardson, 1981). This folding topology is consistent with our earlier analyses of HPr. The $\beta\alpha\beta\beta\alpha\beta\alpha$ folding topology of HPr is interesting in that it is related to another group of proteins that also exhibit

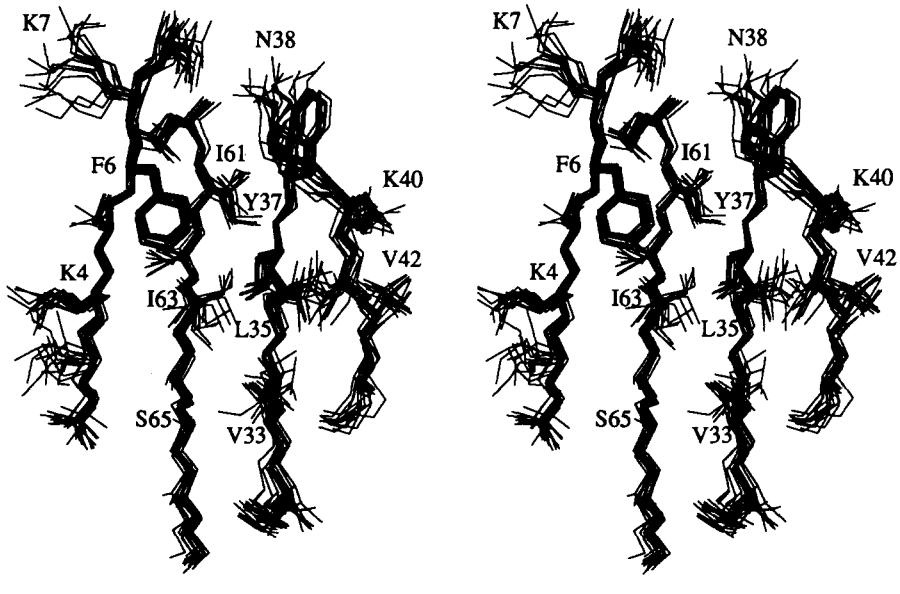

Fig. 7. Stereo view of the interior side of the β-sheet. All C and N atoms and the O^{γ} of Tyr 37 of all 15 structures in {SA} are shown.

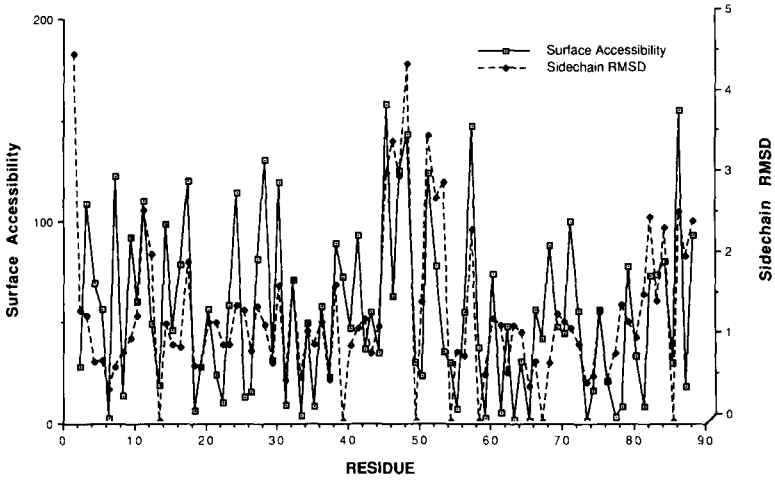

Fig. 8. Plot of side-chain rmsd and surface accessibility for residues of HPr. ---■---, mean pairwise rmsd from the average structure, SA, for all heavy atoms; —⊡— side-chain surface accessibility, calculated with a 1.4-Å probe for the minimized average structure, SA_r (see Materials and methods).

a $\beta\alpha\beta\beta\alpha\beta$ fold. This group includes acylphosphatase (Pastore et al., 1992), the RNA-binding domains of U1 snRNP A (Nagai et al., 1990), and hnRNP C proteins (Wittekind et al., 1992), the activation domain of procarboxypeptidase B (Coll et al., 1991; Vendrell et al., 1991), and a portion of ferredoxin structure (Adman et al., 1973). As recently pointed out, the HPr fold is not topologically equivalent to the one shared by these other proteins (Pastore et al., 1992), if one compares the two folds in the usual way. However, if the C α-helix of HPr is disregarded, we note that the $\beta\alpha\beta\beta\alpha\beta$ folding topologies of the two groups of proteins are actually identical, except that the chain polarity is in the opposite direction (see Fig. 10). Another way of relating the two folds is that they can be interconverted by changing the direction of the chain in the β-turn connecting β-strands B and C, resulting in a different interdigitation of the β-strands. It is clear that the $\beta\alpha\beta\beta\alpha\beta$ fold represents an efficient way of packing the polypeptide chain into a compact, stable unit that can be utilized as a framework for protein domains carrying out widely different functions.

Protected amide protons

Amide protons that are somewhat protected from exchange with solvent (20–60 min at pH 6.9, 30 °C) were detected in a series of fast (^{1}H, ^{15}N)-HMQC spectra. As indicated in Figure 3, most of these amide protons are at the ends of elements of regular secondary structure and probably indicate a "fraying" at the edges of the H-bonding network. Kinemage 4 shows the protected amide protons in the three-dimensional structure (SA_r) of *B. subtilis* HPr, color-coded according to their exchange rates. Most of the amides of residues at the ends of β-strands appear to participate in cross-strand main-chain H-bonds that would be predicted from the topology of the β-sheet. The exceptions are Ala 31, at the N-terminal end of strand B and Glu 60, at the N-terminal end of

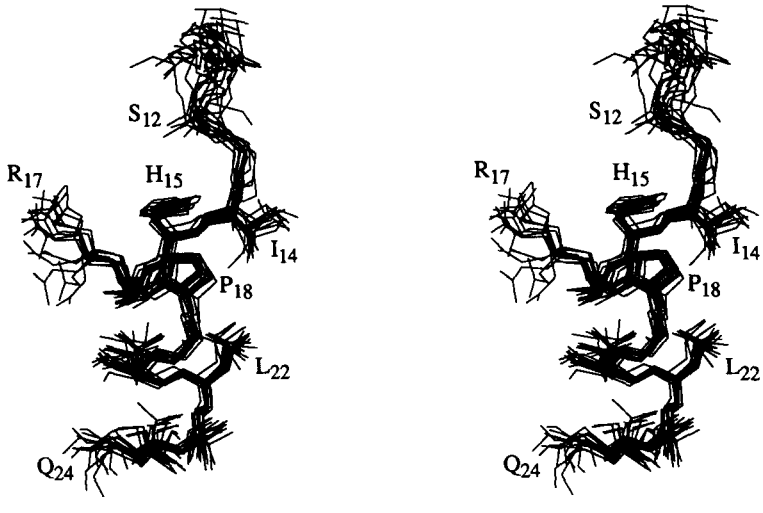

Fig. 9. Stereo view of the region near the site of phosphorylation, His 15. Superposition of all 15 structures in {SA} is shown; the backbone atoms (N, C^α, C′) and the heavy atoms of the side chains for residues 10–24 are illustrated.

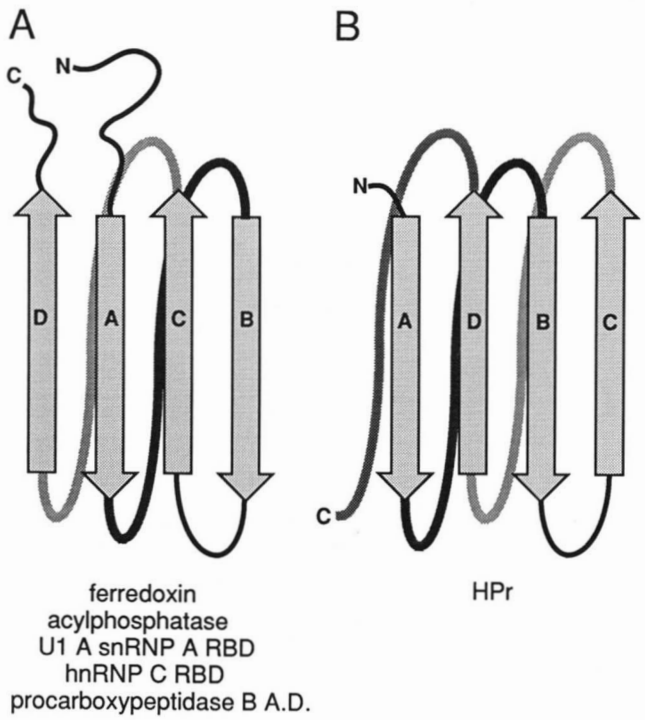

Fig. 10. Folding patterns of $\beta\alpha\beta\beta\alpha\beta$ proteins with either **(A)** D–A–C–B or **(B)** A–D–B–C type β-sheet topologies. The α-helices and connecting loops are represented by the gray lines, whereas the β-strands are depicted as arrows with the direction of the peptide chain indicated by the arrowheads. The amino- and carboxy-termini are denoted by N and C, respectively, and A, B, C, and D designate the order in which the β-strands appear in the primary sequence. The topologies are the same if the peptide chain polarity is disregarded.

strand D. Inspection of the structures in {SA} suggest a source for Ala 31's protection; in 6 of the 15 structures, as well as in SA_r, one of the carboxylate oxygens of Asp 69 is within H-bond distance and geometry to the NH of Ala 31. The amide of Glu 60 has a reasonable H-bonding distance and geometry vis-à-vis the carbonyl oxygen of Gly 58 in 11 of the 15 structures and in SA_r (see Kinemage 4). Some amide protons in turns (Ala 10, Gly 58, and Glu 70) are also transiently protected, but the source of protection is not clear from the structures.

Several transiently protected amide protons are not part of regular secondary structural elements, and their H-bonding partners were not included in the calculations. These include Lys 7, Ala 56, and residues 86–88. Lys 7 is on β-strand A, one of the external strands of the β-sheet, with its amide proton pointing toward solvent. The analogous amide proton in the *E. coli* protein exchanges too rapidly to be observed (Hammen et al., 1991). As shown in Kinemage 4, the NH of Lys 7 is within H-bonding distance and geometry to the carbonyl of Gly 87 in SA_r. This putative interaction does not appear often in {SA}, however, and its confirmation will require more constraints for the C-terminus of the protein. Similarly, the mode of protection of the three C-terminal amides cannot be inferred from the structure at its present level of definition.

Ala 56 is on the irregularly structured segment that connects β-strand C with β-strand D. This residue is constrained by 13 interproton distances (Fig. 2) and has a local backbone rmsd of 0.67 Å. In the family of structures determined, the H-bond acceptor that is found most consistently within H-bonding distance to the amide of Ala 56 is the phenolic oxygen of Tyr 37 (with an rmsd of 0.51 Å, see Fig. 7), in the reverse turn between β-strands B and C. This potential H-bond occurs in 12 of the 15 structures and is illustrated for SA_r in Kinemage 4. In addition to being transiently persistent, the chemical shift of the ^{15}N of Ala 56 is the most downfield in the protein, at 135.4 ppm. Although such a downfield shift could be due to H-bonding, it is unlikely that this H-bond is the strongest in the protein (the exchange half time for Ala 56 is only 20 min), indicating that there must be an additional contribution to the downfield shift of this nucleus. Deshielding due to the ring current associated with the tyrosine ring seems a likely candidate. We note that Tyr 37 is conserved in all gram-positive HPrs sequenced to date (Reizer et al., 1989b).

Active site

The active site His 15 residue is the N-cap (Richardson & Richardson, 1988) for α-helix A, as illustrated in Figure 9 and in Kinemage 5. According to NOEs, coupling constants, and H-bonding patterns, α-helix A begins at Ala 16 with Ile 14 and His 15 in extended conformations (see Fig. 5B). The backbone atoms of residues 10–13 are less well defined by the NMR data (Fig. 6), although an NOE involving the C^β protons of Ser 12 and the C^δ proton of the imidazole ring of His 15 is clearly observed. In all structures, the aromatic plane of the side chain of His 15 is positioned normal to the α-helix A axis, with its N^δ atom pointing away from the protein and toward solvent. The position of the N^δ atom is consistent with its role as the acceptor site for the phosphoryl group during the phosphotransfer reaction with enzyme I.

The solution structure suggests two elements that could be responsible for the low pK_a of 5.7 of His 15 in HPr (Kalbitzer et al., 1982). The proximity of His 15 to Arg 17 could lower its pK_a through destabilization of the protonated imidazole via the positively charged side chain. Second, an interaction between His 15 and the positive end of the helical dipole of α-helix A could contribute to a destabilization of the protonated form. There is evidence for the interaction of histidines with the N-terminal ends of helix dipoles in several proteins. When a histidine was placed at the N2 position of an α-helix in T4 lysozyme, its pK_a was 6.2, as opposed to 6.8 in unfolded lysozyme (Nicholson et al., 1991). Similarly, when histidine residues were placed at the N-termini of either of the

α-helices of barnase, the pK_a's were lowered by ~0.8 pH units compared to the unfolded protein (Sancho et al., 1992). The side chain of His 15 (which is in the deprotonated state under the NMR conditions) is oriented directly over the helical axis (see Fig. 9, Kinemage 5), suggesting that the helix dipole is responsible in some measure for the low pK_a of His 15. There is, however, also evidence for a contribution of the conserved Arg 17 to the pK_a of His 15 in *E. coli* HPr. When Arg 17 is replaced by a neutral side chain in that species, the pK_a* of His 15 increases to 6.1 (Lee, Pullen, Hammen, Waygood, & Klevit, unpubl.). Therefore, it is likely that both interactions are important in the HPrs.

Comparison with the crystal structure

Recently, the structure of an active single-site mutant of *B. subtilis* HPr (S83C) was solved by X-ray diffraction (Herzberg et al., 1992). The overall folding topologies of the X-ray and NMR structures are the same: both structures have the four-stranded β-sheet described here, comprised of identical residues, and the two long α-helices A and C. The crystal structure also contains a third short α-helix from Ile 47 to Met 51.[5] There is an indication of a helical turn in the NMR structure (although it is less well defined in this region), but between residues Val 50 and Leu 53. Residues 47, 48, and 50–53 all have small $^3J_{NH\alpha}$ coupling constants, but only a single *i*, *i* + 3 medium-range NOE between residues 50 and 53 could be unambiguously identified. In addition, no protected amide protons were observed for residues Lys 45 through Ile 55, indicating that the residues do not exist as a stable α-helix in solution. This region is involved in intermolecular contacts in the crystal, which may be responsible for this apparent difference between the solution and crystal structures.

The NMR and X-ray structures are similar at the active site. The His 15 side chain is oriented in the same way in both, with the imidazole ring covering the N-terminus of α-helix A. From the NMR data, the His 15 χ_2 angle, which determines whether the N^δ atom points toward or away from the solvent, is unambiguously established; this angle could not be determined from the X-ray data. The X-ray structure shows an interaction between Arg 17 and a sulfate ion (the protein was crystallized from 68 to 74% saturated ammonium sulfate). In the NMR structure, the Arg 17 side chain is highly solvent exposed and not well defined.

Comparison of B. subtilis and E. coli HPr

The HPr proteins from *B. subtilis* and *E. coli* share 34% sequence identity and perform the same function in their respective sugar transport systems. However, *B. subtilis* HPr can be phosphorylated by an ATP-dependent kinase at Ser 46, whereas *E. coli* HPr cannot be phosphorylated by this enzyme even though there is a serine at position 46 (Reizer et al., 1989a). We hope to be able to understand both the functional similarities and differences from the structures of these two related proteins. Because the structure determination for *E. coli* HPr is not yet as far advanced as the one reported here (Hammen et al., 1991) a detailed comparison of structures is not warranted at this time. However, comparison of the NMR observations themselves follows as they have been determined in analogous manners for the two proteins and contain structural information.

A residue-by-residue comparison of the $^3J_{NH\alpha}$ coupling constants for the two species of HPr reveal a remarkable similarity (see Table 2 in Hammen et al. [1991] and Table 1, this report). The 60 residues for which coupling constants could be measured in both proteins have an average residue-by-residue difference of 0.92 Hz (SD = 0.90 Hz). There are only four residues whose coupling constants are more than 2 SDs different from this average: residues 50, 51, 52, and 83 with differences of 4.0, 3.5, 3.7, and 4.7 Hz, respectively. The first three of these residues are in the irregularly structured stretch that connects β-strand C with β-strand D. In our recent studies of *E. coli* HPr, a transiently protected amide proton at residue 53 and a small $^3J_{NH\alpha}$ coupling constant for residue 50 were observed, consistent with a helical turn structure, but $^3J_{NH\alpha}$ coupling constants of 6.4, 7.3, and 7.1 Hz for residues 51, 52, and 53 indicate that the structure is dynamically averaged in this region (Hammen et al., 1991). In *B. subtilis* HPr, residues 50–53 have small $^3J_{NH\alpha}$ coupling constants, with the constants for residues 51 and 52 being significantly smaller than in the *E. coli* protein. In the refined *B. subtilis* HPr structures, there is a suggestion of a helical turn between residues 50 and 53, but the overall definition in this region is poor. However, the small coupling constants indicate that, in this species, the region is not significantly averaged. We note that this region lies directly between the two sites of phosphorylation, i.e., His 15 and Ser 46, and residues in this region experience some of the largest chemical shift perturbations upon phosphorylation of Ser 46 in *B. subtilis* HPr (Wittekind et al., 1989).

A residue-by-residue comparison of ^1H–^2H exchange rates for the two HPr species also reveals striking similarities. In both proteins, the most slowly exchanging amide protons are in the C helix. The most persistent amides are Val 74 and Val 78 (1,619 and 1,411 min) in *E. coli*, and Val 78 and Leu 77 (1,500 and 1,066 min) in *B. subtilis*. These residues are on the side of α-helix C facing the hydrophobic core, and their low solvent exchange rates are most likely due to reduced solvent accessibility (see Fig. 8). There are, however, a number of residues with quite different exchange behavior in the two proteins. The H-bonded amides in β-strand A are at positions 4, 6, and 8. Their exchange half-times are 188, 702, and 172 min, respectively, in *E. coli* and 36, 55, and 192 min

in *B. subtilis*. The differences in exchange rates are most likely due to the fact that the *B. subtilis* protein used in our NMR studies lacks residue 1, whereas this residue is involved in a strong $d_{\alpha\alpha}$ cross-β-strand NOE in *E. coli* HPr. Thus, β-strand A is held strongly in the sheet in the *E. coli* protein, but appears to experience more "end-fraying" in the *B. subtilis* structure. Similarly, the large difference in exchange half-times observed at the extreme C-terminus of the two proteins can probably be attributed to the different lengths of the two proteins, as the *B. subtilis* protein is 5 residues longer at its C-terminal end.

Of more interest is the large difference in exchange behavior observed for residues 37, 40, and 42 (10, <10, and 20 min for *E. coli* and 173, 131, and 248 min for *B. subtilis*). These residues make up the β-turn between β-strands B and C, which is a well-defined type I′ turn in the *B. subtilis* structure. Although not currently well defined in the *E. coli* structure, the $^3J_{NH\alpha}$ coupling constants for the turn residues are not significantly different in the two proteins. As suggested earlier, the phenolic oxygen of Tyr 37 in the turn may be involved in an H-bond with the amide proton of Ala 56 in *B. subtilis* HPr, whereas there is no indication of an analogous role for Ser 37 in *E. coli*. This side chain-to-backbone H-bond may be responsible for the increase in stability of this region as evidenced by the differential exchange behaviors in the two proteins.

Conclusions

The structure described here has been determined using a large number of constraints obtained from NMR measurements. The topology obtained is identical to that originally suggested by a qualitative analysis of the NOEs. The higher definition of the structures presented here can now serve as the foundation for further studies on HPr. In particular, we hope to gain insights into the structural and dynamical consequences of phosphorylation at each of the two sites, His 15 and Ser 46, in *B. subtilis* HPr.

Materials and methods

Sample preparation

Expression of *B. subtilis* HPr was carried out in *E. coli* as previously described (Reizer et al., 1989b) except that prototrophic *E. coli* host strain GM-1 was used (Coulandre & Miller, 1977). The HPr purified from this strain lacks the methionine residue at position 1 (Wittekind et al., 1990). For expression of uniformly labeled ^{15}N-enriched HPr protein, a minimal growth medium (McIntosh et al., 1987) was used with $^{15}NH_4Cl$ (99% isotopic enrichment, Cambridge Isotope Laboratories, Woburn, Massachusetts). Two samples of unlabeled HPr, one in D_2O and the other in 90%/10% H_2O/D_2O, were prepared for NMR analysis as previously described (Wittekind et al., 1990). A single sample of uniformly labeled ^{15}N HPr in 90%/10% H_2O/D_2O was used for all inverse-detected experiments. All protein samples were approximately 4 mM.

NMR spectroscopy

All spectra were obtained on a Bruker AM-500 spectrometer. Unless otherwise noted, the spectra were collected at 30 °C with spectral widths of 6,579 and 2,000 Hz for ^{1}H and ^{15}N, respectively. For all ^{1}H 2D experiments, TPPI (Marion & Wüthrich, 1983) was used for quadrature detection in the F1 dimension. The software packages FTNMR and FELIX (Hare Research, Bothell, Washington) were used to transform the spectra.

Two-dimensional ^{1}H NOESY spectra were collected as described (Wittekind et al., 1990); with τ_m = 5, 25, 50, 75, 100, and 150 ms (for sample in H_2O) and τ_m = 30, 60, and 120 ms (for sample in D_2O).

^{1}H,^{15}N-correlated HMQC, HMQC-NOESY, and HMQC-HOHAHA spectra were collected using published pulse sequences (Gronenborn et al., 1989) with an acquisition time in the ^{15}N dimension of 80 ms. Mixing times of 120 and 90 ms were used for the HMQC-NOESY and HMQC-HOHAHA spectra, respectively.

Amide protons resistant to exchange with solvent were detected by rapid measurement of ^{1}H,^{15}N-correlated HMQC spectra, after the ^{15}N-labeled protein was dissolved in D_2O. The Bruker AM500 spectrometer was modified with a timing device (Tschudin Associates, Kensington, Maryland) which allowed for buffered acquisition (Marion et al., 1989). Full 2D data sets (256 × 2K points) were collected in 20 min (four scans per serial file), using a refocused INEPT type of transfer between ^{1}H and ^{15}N (Clore et al., 1990). Ten minutes after D_2O was added to lyophilized protein, the first experiment was begun and subsequent experiments were performed over 24 h. Spectra were collected at 20, 40, 60, 82, 105, 130, 151, 195, 250, 370, 430, 490, 1,440, and 1,500 min after mixing. The data were processed using an unshifted sine-bell squared filter in both dimensions, with zero-filling to 1,024 points in t1, yielding a 1K × 1K data matrix. Volume integration was carried out on each cross-peak, and scaled information was plotted as a function of time. The resulting curves were fit to a single-exponential decay to calculate the apparent solvent exchange rates for each amide proton.

The $^3J_{HN\alpha}$ coupling constants were determined from HMQC-J spectra (method 3 in Kay & Bax [1990]) acquired on uniformly ^{15}N-labeled HPr in H_2O. Twelve hundred fifty TPPI increments were collected with a total acquisition time of 250 ms in the ^{15}N dimension, giving a spectral width in F1 of 2,500 Hz. The ^{15}N-^{1}H multiple-quantum linewidth for HPr was 7.0 ± 0.6 Hz. Data were processed and analyzed as previously described (Hammen et al., 1991).

$^3J_{\alpha\beta}$ coupling constants were obtained by analysis of P.E.COSY spectra (Mueller, 1987) obtained in D_2O. The spectra were apodized in both dimensions with a sine-bell squared filter shifted by 45°. Splittings for resolved cross-peaks were measured along F2 after zero filling to yield 1.6 Hz per point digital resolution.

Structure calculations

A NOESY cross-peak template file was created from the 150-ms NOESY spectra and was used to extract the integrated NOE intensity of each cross-peak from each of the six NOESY spectra collected in H_2O. An in-house software routine (written by Dr. Jon Herriott) was used to view the NOE build-up curves for each cross-peak and to calculate the build-up rates using a least-squares fit to the data.

Build-up rates were converted to target distances by normalizing to the average rate obtained for d_{NN} sequential NOE cross-peaks (set at 2.8 Å) from residues that correspond to the two well-defined α-helical regions of HPr (residues 21–27 and 73–84). Upper and lower bounds were then determined by calculating a range around each target distance as follows. For target distances ≤2.7 Å, the upper and lower bounds were calculated by adding and subtracting 0.5 Å from the target distance. For target distances >2.7 Å and ≤3.7 Å, bounds were calculated by adding 1.0 Å and subtracting 0.5 Å from the target distance. For target distances >3.7 Å, bounds were calculated by adding and subtracting 1.0 Å from the target distance. Cross-peaks that were resolved only in 2D ^1H NOESY recorded in D_2O or in (^1H–^{15}N)-HMQC-NOESY spectra were classified as being either strong, medium, weak, or very weak and were assigned bounds of 1.85–2.85 Å, 2.35–3.85 Å, 3.00–4.50 Å, and 3.00–5.50 Å, respectively. Whenever a methyl group was involved in an NOE interaction, an additional 0.5 Å was added to the upper bound (Wagner et al., 1987).

The program HABAS (Güntert et al., 1989) was used to determine allowed ranges for the backbone angles phi and psi. The input to this program included the distance constraints, the values of $^3J_{NH\alpha}$ obtained from analysis of HMQC-J data, and the values of $^3J_{\alpha\beta}$ determined by analysis of P.E.COSY spectra. Stereospecific assignments for prochiral β-methylene protons were made for 18 out of the 49 possible side chains. Chi I ranges were determined by HABAS or by manual inspection of the NOE and coupling data. Ranges for the angle constraints were often larger but never less than 60, 60, and 40° for phi, psi, and chi I, respectively.

Amide protons protected from exchange with solvent were identified from a series of fast (^1H,^{15}N)-HMQC spectra (described above). An H-bond interaction was included in the bounds file only when the identity of the H-bond acceptor for a given transiently protected amide proton was unambiguous. This decision was based on the observation of NOEs and coupling constants defining the residue as part of a segment of regular secondary structure. Appropriate cross-strand NOEs or $d_{\alpha N(i,i+3,4)}$ NOEs had to be present in order for H-bonds to be included for residues involved in antiparallel β-sheets or α-helices, respectively. Each H-bond was represented in the bounds file by two constraints: the ranges for the amide proton to carbonyl oxygen and amide nitrogen to carbonyl oxygen were 1.60–2.30 Å and 2.50–3.30 Å, respectively.

The final bounds file consisted of 286 i, i, 174 i, $i+1$, 64 i, $i+n$ ($n = 2, 3, 4$) and 106 i, $>i+4$ constraints, for a total of 629 measured constraints. Thirty-eight H-bonds were represented by an additional 76 constraints, making a total of 705 distance constraints, averaging 8.1 distance constraints per residue. Seventy-three phi, 57 psi, and 17 chi I constraints were included. The bounds files used for the structure determination are available as supplemental information on the Diskette Appendix.

Random starting structures were generated by the variable target-function minimization program DIANA (Güntert et al., 1991). Contributions to the target function from distance, dihedral angle, and van der Waals violations were weighted in a ratio of 1:10:0.2 for 4,150 iterations, whereas the residue window size was gradually increased until all residues were included. The van der Waals contribution was then increased to 0.6 (500 iterations) and again to 1.0 (500 iterations). Of the resulting structures, inspection of the 60 conformers exhibiting target functions below an arbitrary cutoff value of 75 $Å^2$ revealed that these structures shared a common global folding pattern. These structures were further minimized using DIANA for 250 iterations with the contributions to the target function from distance, dihedral angle, and van der Waals' violations weighted in a ratio of 1:5:0.2 followed by 1,000 iterations at 1:5:0.6, and finally 5,000 iterations at 1:5:2. The structures exhibiting the fewest violations had target functions in the range of 20–40 $Å^2$ when evaluated using both upper and lower bounds.

The 60 structures were subjected to a dynamical simulated annealing protocol (Nilges et al., 1988) using X-PLOR (Brünger, 1990). The refinement was carried out using only upper bounds as distance constraints while retaining all dihedral angle constraints. The annealing protocol was identical to the method described for refinement of BDS-I (Driscoll et al., 1989), except slow cooling was carried out in phase 3. Briefly, the protocol consists of 200 steps of Powell minimization during phase 1 followed by phase 2, which consists of 3.75 ps of molecular dynamics at 1,000 K. During this dynamics phase, the NOE and dihedral angle force constants were increased from 0.5 to 50 kcal mol^{-1} $Å^2$ and 2.5 to 200 kcal mol^{-1} rad^{-2}, respectively, by doubling the values every 75 ps. The van der Waals force constant was gradually increased from a starting value of 0.01 kcal mol^{-1} $Å^{-4}$ to a final value of 4.0 kcal mol^{-1} $Å^{-4}$ at the end of this stage. This was followed by phase 3 in which molecular

dynamics was performed while the system was cooled slowly to 300 K by lowering the temperature by 25 K every 50 fs. The protocol was completed in phase 4 by performing 200 cycles of Powell minimization. All other details are identical to the BDS-1 protocol (Driscoll et al., 1989) except that the force constant for maintaining bond lengths was 1,000 kcal mol^{-1} $Å^{-2}$ instead of 600 kcal mol^{-1} $Å^{-2}$.

The 15 best structures {SA}, as evaluated by residual distance and angle violations, were superimposed and the coordinates averaged to generate an average structure $\overline{SA}$. After 2,000 rounds of Powell minimization to restore proper bond lengths and to minimize van der Waals interactions, the resulting structure SA_r had no residual NOE violations over 0.5 Å and no residual angle violations over 1°. The coordinates for SA_r have been deposited at Brookhaven Protein Data Bank.

Surface accessibility (Lee & Richards, 1971) of sidechain atoms was calculated with X-PLOR using a 1.4-Å sphere. Hydrogen bonds in each structure were identified using the default parameters in QUANTA (version 2.1, Polygen Corp.): maximum distance between H and acceptor = 3.0 Å; minimal angles at the acceptor, at the H, and at the donor = 90°

Acknowledgments

This work was supported by NIH grant DK-35187 (R.E.K.). R.E.K. was also supported by an American Heart Association Established Investigatorship Award; M.G.W. was supported by a Damon Runyon-Walter Winchell Cancer Research Fund Fellowship (DRG-1007). We thank Bev Castner for technical assistance and Phil Hammen and Jon Herriott for helpful discussions and critical reading of the manuscript.

References

Adman, E.T., Sieker, L.E., & Jensen, L.H. (1973). Structure of a bacterial ferredoxin. *J. Biol. Chem. 248*, 3987–3996.

Brünger, A.T. (1990). *X-PLOR Version 2.1 User Manual*. Yale University, New Haven.

Clore, G.M., Driscoll, P.C., Wingfield, P.T., & Gronenborn, A.M. (1990). Analysis of the backbone dynamics of interleukin-1 beta using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy. *Biochemistry 29*, 7387–7401. [Published erratum: *Biochemistry 30*, 312 (1991).]

Coll, M., Guasch, A., Aviles, F.X., & Huber, R. (1991). Three-dimensional structure of porcine procarboxypeptidase B: A structural basis for its inactivity. *EMBO J. 10*, 1–9.

Coulandre, C. & Miller, J. H. (1977). Genetic studies of lac repressor. *J. Mol. Biol. 117*, 525–575.

Driscoll, P.C., Gronenborn, A.M., Beress, L., & Clore, G.M. (1989). Determination of the three-dimensional solution structure of the antihypertensive and antiviral protein BDS-I from the sea anemone *Anemonia sulcata*: A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry 28*, 2188–2198.

el Kabbani, O.A., Waygood, E.B., & Delbaere, L.T. (1987). Tertiary structure of histidine-containing protein of the phosphoenolpyruvate:sugar phosphotransferase system of *Escherichia coli*. *J. Biol. Chem. 262*, 12926–12929.

Gronenborn, A.M., Bax, A., Wingfield, P.T., & Clore, G.M. (1989). A powerful method of sequential proton resonance assignment in proteins using relayed ^{15}N-^{1}H multiple quantum coherence spectroscopy. *FEBS Lett 243*, 93–98.

Güntert, P., Braun, W., Billeter, M., & Wüthrich, K. (1989). Automated stereospecific ^{1}H NMR assignments and their impact on the precision of protein structure determinations in solution. *J. Am. Chem. Soc. 111*, 3997–4004.

Güntert, P., Braun, W., & Wüthrich, K. (1991). Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol. 217*, 517–530.

Güntert, P. & Wüthrich, K. (1991). Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J. Biomol. NMR 1*, 447–456.

Hammen, P.K., Waygood, E.B., & Klevit, R.E. (1991). Reexamination of the secondary and tertiary structure of histidine-containing protein from *Escherichia coli* by homonuclear and heteronuclear NMR spectroscopy. *Biochemistry 30*, 11842–11850.

Herzberg, O., Reddy, P., Sutrina, S., Saier, M., Reizer, J., & Kapadia, G. (1992). Structure of the histidine-containing phosphocarrier protein HPr from *Bacillus subtilis* at 2.0 A resolution. *Proc. Natl. Acad. Sci. USA 89*, 2499–2503.

Kalbitzer, H.R., Hengstenberg, W., Rosch, P., Muss, P., Bernsmann, P., Engelmann, R., Dorschug, M., & Deutscher, J. (1982). HPr proteins of different microorganisms studied by hydrogen-1 high-resolution nuclear magnetic resonance: Similarities of structure and mechanisms. *Biochemistry 21*, 2879–2925.

Kalbitzer, H.R., Neidig, K.P., & Hengstenberg, W. (1991). Two-dimensional ^{1}H NMR studies on HPr protein from *Staphylococcus aureus*: Complete sequential assignments and secondary structure. *Biochemistry 30*, 11186–11192.

Kay, L.E. & Bax, A. (1990). New methods for the measurement of NH-CαH coupling constants in ^{15}N-labeled proteins. *J. Magn. Reson. 86*, 110–126.

Klevit, R.E., Drobny, G.P., & Waygood, E.B. (1986). Two-dimensional ^{1}H NMR studies of histidine-containing protein from *Escherichia coli*. 1. Sequential resonance assignments. *Biochemistry 25*, 7760–7769.

Klevit, R.E. & Waygood, E.B. (1986). Two-dimensional ^{1}H NMR studies of histidine-containing protein from *Escherichia coli*. 3. Secondary and tertiary structure as determined by NMR. *Biochemistry 25*, 7774–7781.

Lee, B. & Richards, F.M. (1971). The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol. 55*, 379–400.

Ludvigsen, S. & Poulsen, F.M. (1992). Positive φ-angles in proteins by nuclear magnetic resonance spectroscopy. *J. Biomol. NMR 2*, 227–233.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989). Rapid recording of 2D NMR spectra of proteins without phase cycling. *J. Magn. Reson. 85*, 393–399.

Marion, D. & Wüthrich, K. (1983). Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun. 113*, 967–974.

McIntosh, L.P., Griffey, R.H., Muchmore, D.C., Nielson, C.P., Redfield, A.G., & Dahlquist, F.W. (1987). Proton NMR measurements of bacteriophage T4 lysozyme aided by ^{15}N isotopic labeling: Structural and dynamic studies of larger proteins. *Proc. Natl. Acad. Sci. USA 84*, 1244–1248.

Mueller, L. (1987). P.E.COSY, a simple alternative to E.COSY. *J. Magn. Reson. 72*, 191–197.

Nagai, K., Oubridge, C., Jessen, T.H., Li, J., & Evans, P.R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature 348*, 515–520.

Nicholson, H., Anderson, D.E., Dao, P.S., & Matthews, B.W. (1991). Analysis of the interaction between charged side chains and the alpha-helix dipole using designed thermostable mutants of phage T4 lysozyme. *Biochemistry 30*, 9816–9828.

Nilges, M., Gronenborn, A.M., Brünger, A.T., & Clore, G.M. (1988). Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Eng. 2*, 27–38.

Pastore, A., Saudek, V., Ramponni, G., & Williams, R.J.P. (1992).

Three-dimensional structure of acylphosphatase. *J. Mol. Biol. 224*, 427–440.

Reizer, J., Deutscher, J., & Saier, M.J. (1989). Metabolite-sensitive, ATP-dependent, protein kinase-catalyzed phosphorylation of HPr, a phosphocarrier protein of the phosphotransferase system in gram-positive bacteria. *Biochimie 71*, 989–996.

Reizer, J., Sutrina, S.L., Saier, M.H., Stewart, G.C., Peterkofsky, A., & Reddy, P. (1989). Mechanistic and physiological consequences of HPr(ser) phosphorylation on the activities of the phosphoenolpyruvate:sugar phosphotransferase system in gram-positive bacteria: Studies with site-specific mutants of HPr. *EMBO J. 8*, 2111–2120.

Richardson, J.S. (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem 34*, 167–339.

Richardson, J.S. & Richardson, D.C. (1988). Amino acid preference for specific locations at the ends of helices. *Science 240*, 1648–1652.

Saier, M.H., Jr. & Reizer, J. (1992). Proposed uniform nomenclature for the proteins and protein domains of the bacterial phosphoenolpyruvate:sugar phosphotransferase system. *J. Bacteriol. 174*, 1433–1438.

Sancho, J., Serrano, L., & Fersht, A.R. (1992). Histidine residues at the N- and C-termini of alpha-helices: Perturbed pKas and protein stability. *Biochemistry 31*, 2253–2258.

Vendrell, J., Billeter, M., Wider, G., Aviles, F.X., & Wüthrich, K. (1991). The NMR structure of the activation domain isolated from porcine procarboxypeptidase B. *EMBO J. 10*, 11–15.

Wagner, G., Braun, W., Havel, T.S., Schaumann, T., Go, N., & Wüthrich, K. (1987). Protein structure in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, disgeo and disman. *J. Mol. Biol. 196*, 611–639.

Wittekind, M., Görlach, M., Friedrichs, M.S., Dreyfuss, G., & Mueller, L. (1992). ^{1}H, ^{13}C, and ^{15}N NMR assignments and global folding pattern of the RNA binding domain of the human hnRNP C proteins. *Biochemistry 31*, 6254–6265.

Wittekind, M., Reizer, J., Deutscher, J., Saier, M.H., & Klevit, R.E. (1989). Common structural changes accompany the functional inactivation of HPr by seryl phosphorylation or by serine to aspartate substitution. *Biochemistry 28*, 9908–9912.

Wittekind, M., Reizer, J., & Klevit, R.E. (1990). Sequence-specific ^{1}H NMR resonance assignments of *Bacillus subtilis* HPr: Use of spectra obtained from mutants to resolve spectral overlap. *Biochemistry 29*, 7191–7200.