Sequence-Specific ^{1}H NMR Resonance Assignments of *Bacillus subtilis* HPr: Use of Spectra Obtained from Mutants To Resolve Spectral Overlap†

Michael Wittekind,‡ Jonathan Reizer,§ and Rachel E. Klevit*,‡

Department of Biochemistry, University of Washington, Seattle, Washington 98195, and Department of Biology, University of California—San Diego, La Jolla, California 92093

Received December 5, 1989; Revised Manuscript Received March 21, 1990

ABSTRACT: On the basis of an analysis of two-dimensional ^{1}H NMR spectra, the complete sequence-specific ^{1}H NMR assignments are presented for the phosphocarrier protein HPr from the Gram-positive bacterium *Bacillus subtilis*. During the assignment procedure, extensive use was made of spectra obtained from point mutants of HPr in order to resolve spectral overlap and to provide verification of assignments. Regions of regular secondary structure were identified by characteristic patterns of sequential backbone proton NOEs and slowly exchanging amide protons. *B. subtilis* HPr contains four β-strands that form a single antiparallel β-sheet and two well-defined α-helices. There are two stretches of extended backbone structure, one of which contains the active site His_{15}. The overall fold of the protein is very similar to that of *Escherichia coli* HPr determined by NMR studies [Klevit, R. E., & Waygood, E. B. (1986) *Biochemistry 25*, 7774–7781].

HPr[1] is a heat stable protein of molecular weight ~9000 that functions as one of the phosphocarrier proteins in the phosphoenolpyruvate:sugar phosphotransferase system (PTS) that is responsible for the simultaneous uptake and phosphorylation of numerous sugars in both Gram-negative and Gram-positive bacteria (Postma & Lengler, 1985; Reizer et al., 1988a). In the PTS, enzyme I transfers a phosphoryl group from phosphoenolpyruvate to the N^{δ}-position of the imidazole ring of His_{15} of HPr. HPr(His-P) transfers this phosphate to the N^{ϵ}-position of the active site histidine of factor III^{sugar} or enzyme II^{sugar}, depending on the specific sugar being transported [for further details, see Saier (1985)].

In HPr from Gram-positive bacteria, Ser_{46} can be phosphorylated by a HPr-specific ATP-dependent protein kinase to form HPr(Ser-P) (Reizer et al., 1984; Reizer & Peterkofsky, 1987). This regulatory phosphorylation has an inhibitory effect on HPr function and has been proposed to give rise to a hierarchy of preferences for the utilization of PTS sugars (Deutscher et al., 1984). In addition, HPr and HPr(Ser-P) kinase as well as the HPr-specific protein phosphatase have been found in heterofermentative lactobacilli that lack the other PTS components, suggesting that HPr(Ser-P) has a role in the regulation of non-PTS-related functions (Romano et al., 1987; Reizer et al., 1988b).

The sequence-specific ^{1}H NMR assignments and model for the solution structure of HPr for the Gram-negative bacterium *Escherichia coli* have been previously reported (Klevit et al., 1986; Klevit & Waygood, 1986). In this paper, we describe a two-dimensional ^{1}H NMR analysis of HPr from the Gram-positive bacterium *Bacillus subtilis*. Although HPrs from *E. coli* and *B. subtilis* perform the same function in their respective PTSs and share 34% sequence identity (Gonzy-Treboul et al., 1989; Reizer, 1989), they substitute poorly for each other in PTS assays in vitro (Reizer et al., unpublished results). In addition, HPrs from the Gram-negative bacteria *E. coli* and *Salmonella typhimurium* have a serine at position 46 and do not undergo the regulatory phosphorylation reaction at this residue as do the Gram-positive HPrs (Reizer et al., 1984; Reizer & Peterkofsky, 1987). A study that compares the solution structures of the two HPrs will help to elucidate common structural features that are relevant to the mechanism and regulation of phosphate transfer.

One-dimensional NMR studies have been reported that establish pK_a values for titratable residues, aromatic ring system resonance assignments, and the pH and temperature stabilities of Gram-positive HPrs (Maurer et al., 1977; Kalbitzer et al., 1982, 1985). However, two-dimensional studies are required to obtain sequence-specific assignments in order to derive structural information. The assignments presented here, which have been obtained by an analysis of wild-type and mutant HPr spectra, have recently been used to show that the serine-to-aspartate substitution mutation at position 46 causes the same type of perturbations of the structure of *B. subtilis* HPr as does phosphorylation at Ser_{46} (Wittekind et al., 1989). Further utilization of these assignments, in conjunction with analysis of NOESY data, will facilitate a detailed study of the solution structure of unmodified HPr as well as that of phosphorylated and mutant forms of *B. subtilis* HPr.

Materials and Methods

Protein Purification and Sample Preparation. Site-directed mutagenesis to produce the S46A and S46D mutants, overexpression of wild-type and mutant *B. subtilis* HPrs in *E. coli*, and their purification have been described (Reizer et al., 1989).

† This work was supported by NIH Grants DK35187 (R.E.K.) and AI21702 and AI14176 (J.R.). R.E.K. was supported by American Heart Association Established Investigatorship and M.W. was supported by Damon Runyon-Walter Winchell Cancer Research Fund Fellowship DRG-1007.

‡ University of Washington.

§ University of California—San Diego.

[1] Abbreviations: PTS, phosphoenolpyruvate-dependent sugar phosphotransferase system; HPr, histidine-containing phosphocarrier protein of the PTS; PEP, phosphoenolpyruvate; ATP, adenosine 5′-triphosphate; [^{2}H]TSP, sodium 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; 2D NMR, two-dimensional nuclear magnetic resonance spectroscopy; COSY, two-dimensional *J*-correlated spectroscopy; DQF, double quantum filter; 2Q, two-dimensional double quantum spectroscopy; RELAY, two-dimensional relayed coherence transfer spectroscopy; NOESY, two-dimensional NOE spectroscopy; TOCSY, two-dimensional total coherence spectroscopy.

0006-2960/90/0429-7191$02.50/0

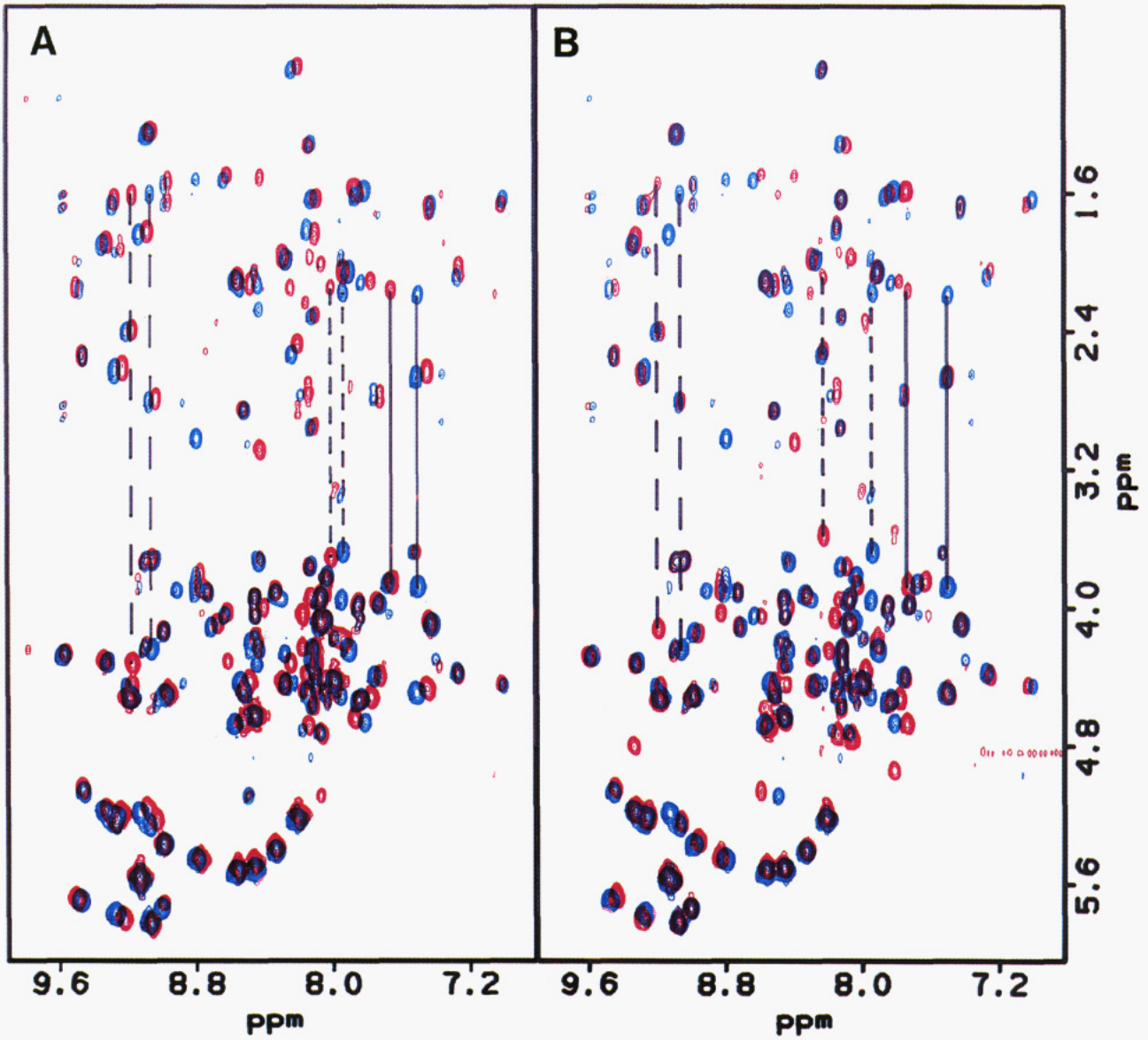

FIGURE 1: Spectra of wild-type and mutant forms of HPr. The regions of the ^1H_2O RELAY spectra containing the NH–C$^\alpha$H and NH–C$^\beta$H cross-peaks are overlaid. In both panels, the mutant spectrum is shown in red while the wild-type spectrum is shown in blue. (A) HPr mutant S46A versus wild-type HPr; (B) HPr mutant S46D versus wild-type HPr. The NH–C$^\alpha$H and NH–C$^\beta$H cross-peak pairs for several representative amino acid spin systems are connected with lines: Val_{21} (solid); Val_{50} (small dashes); Leu_{44} (large dashes).

The pooled fractions from the G-50 column were lyophilized, and NMR samples were prepared by dissolving 20 mg of HPr in a 4-mL volume of 5 mM potassium phosphate buffer, pH 6.9, followed by dialysis against the same buffer. The sample was lyophilized again and dissolved in 0.4 mL of either 99.96% ^2H_2O or 90% ^1H_2O/10% ^2H_2O, centrifuged to remove insoluble material, and transferred to a 5-mm NMR tube. The final concentration of HPr was ~5.6 mM, in 50 mM potassium phosphate buffer, pH 6.9.

NMR Spectroscopy. All ^1H NMR spectra were obtained on Bruker AM-500 and WM-500 spectrometers except the previously described home-built 500-MHz NMR spectrometer (Klevit & Drobny, 1986) was used to record the TOCSY spectra. The spectra were collected over 6579 Hz with quadrature detection into 1024 complex points, giving a digital resolution after transformation of 6.4 Hz/point in t_2. Typically, 64 scans were signal-averaged for each t_1 value with a recycle time of 1.5 s in ^2H_2O and 2.0 s in ^1H_2O. Solvent suppression was performed by presaturation during the recycle time or by utilization of the SCUBA sequence (Brown et al., 1988) with t_p = 60 ms. Unless otherwise noted, all spectra were obtained at 30 °C. Data processing was carried out with the software package FTNMR (Hare Research, Woodinville, WA). Chemical shift values are referenced to [^{2}H]TSP at 0.0 ppm.

COSY (Aue et al., 1976) and RELAY (Eich et al., 1985) spectra were collected and processed in the absolute value mode by use of a sine-bell window function in each dimension. Mixing times of 32 ms (^1H_2O) and 26 ms (^2H_2O) were used in the RELAY experiments (Bax & Drobny, 1985). Typically, 300–350 t_1 increments were recorded for each two-dimensional experiment.

NOESY (Bodenhausen et al., 1984), TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985), and DQF-COSY (Rance et al., 1983) experiments were collected in the phase-sensitive mode by use of time-proportional phase incrementation for quadrature detection in the ω_1 dimension (Marion & Wüthrich, 1983). With this implementation,

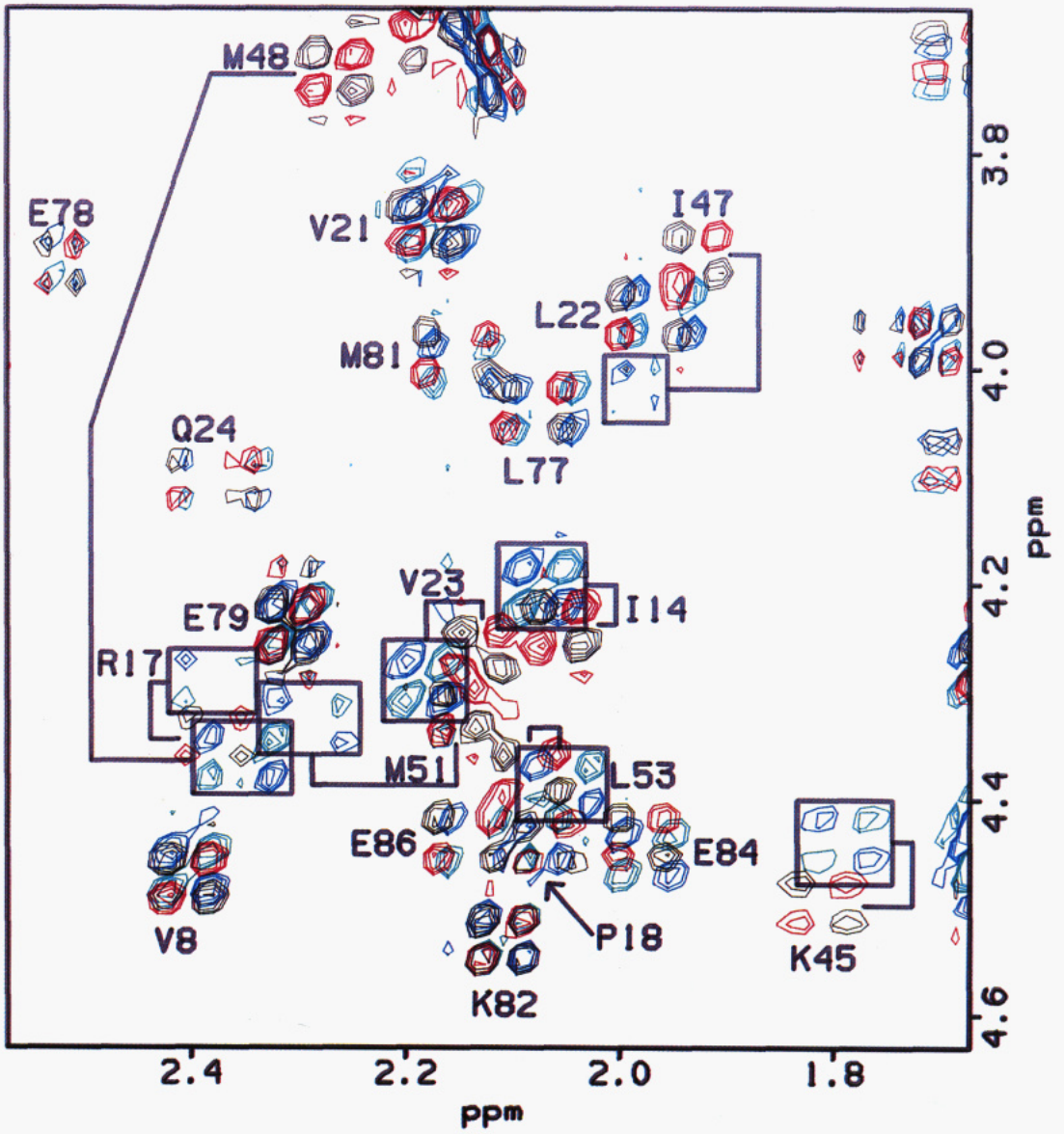

FIGURE 2: Comparison of the aliphatic regions of the spectra. The ^2H_2O DQF-COSY spectra of wild-type HPr (black and red) and the S46D mutant HPr (blue and green) are overlaid. The region depicted shows some of the $C^\alpha H$–$C^\beta H$ cross-peaks. Cross-peaks that show large shifts between the two spectra are indicated with lines parallel to the wild-type cross-peak along either ω_2 or ω_1 connected to boxes outlining the position of the same cross-peak in the S46D spectrum.

600–700 experiments were obtained for each two-dimensional experiment. NOESY data were collected with mixing times of 75 and 150 ms and apodized with a sine-bell window function, phase shifted by 45°, and skewed toward the beginning of the free induction decay in both t_2 and t_1. The TOCSY experiment was obtained by use of the MLEV17 mixing sequence (Bax & Davis, 1985) with a mixing time of 70 ms, and resolution enhancement during data processing was carried out as described for the NOESY experiment. Data from the DQF-COSY experiments were processed with a sine-bell window function in both dimensions.

Double quantum (2Q) spectra (Braunschweiler et al., 1983) were acquired by use of a 40-ms delay during the multiple quantum excitation period and processed in the absolute value mode.

For determination of persistent amide protons, a wild-type HPr sample was lyophilized from a buffered ^1H_2O solution and resuspended in ^2H_2O, and a RELAY experiment was begun within 2 h. The temperature was 30 °C and the duration of the experiment was approximately 14 h.

Results

Characterization of Spin Systems. The sequence-specific assignment procedure developed by Wüthrich and co-workers (Wüthrich et al., 1982; Wüthrich, 1986) was utilized in this study. At the initial stage of spin-system characterization, cross-peaks present within the NH–$C^\alpha H$ fingerprint region of the COSY spectra were correlated with the $C^\beta H$ spins by use of RELAY spectra recorded in ^1H_2O. To overcome spectral overlap problems, use was made of spectra acquired from mutant forms of HPr in which the serine at position 46 was replaced by various amino acid residues (Reizer et al., 1989; Wittekind et al., 1989). In Figure 1A, the amide region of the RELAY spectra of the wild-type and the serine-to-alanine substitution mutant (S46A) proteins are overlaid. Many of the cross-peaks are exactly coincidental in the two spectra while other cross-peaks show perturbations that range from less than a line width to 0.4 ppm. As illustrated for several examples, by following the positions of both NH–$C^\alpha H$ and NH–$C^\beta H$ cross-peaks, spin systems could be unequivocally identified

simultaneously in both wild-type and mutant spectra. Thus, in cases where several protons resonate at a given frequency in one of the spectra, it was often possible to identify the spin systems if one of the amides was even slightly shifted from its position in the mutant spectrum. A different set of perturbations is found when the wild-type spectrum is compared to the serine-to-aspartate substitution mutant (S46D) spectrum (Figure 1B).

The superposition of the aliphatic regions of the DQF-COSY spectra of the wild-type and S46D mutant proteins is shown in Figure 2. The phase-sensitive, four-color representation aids in the identification of the shifted $C^{\alpha}H-C^{\beta}H$ cross-peaks. Again, perfect superposition of cross-peaks for certain resonances (e.g., Lys_{82} and Glu_{79}) is observed while other cross-peaks undergo shifts in one or both dimensions (e.g., Ile_{14} and Val_{23}). Analysis of these spectra allowed for the identification of self-consistent sets of shifted $NH-C^{\alpha}H/NH-C^{\beta}H/C^{\alpha}H-C^{\beta}H$ cross-peaks and confirmed that the correct spin-subsystem identifications had been made.

Standard analysis of RELAY and TOCSY spectra obtained in ^2H_2O was utilized to assign an amino acid type or class designation to as many spin subsystems as possible. In some cases, complete assignments were made for all carbon-bound protons within a spin system. This was possible for the unique arginine (Arg_{17}) and proline (Pro_{18}) spin systems (Figure 3). The identification of these systems was useful because they are in a region of HPr devoid of regular secondary structure and provided a starting point during the subsequent sequence-specific assignment stage. Further assignments were made possible by analysis of the double quantum (2Q) spectra. These experiments were useful for both identifying $C^{\beta}H$ resonances of many spin systems and providing verification of the glycine spin-system assignments (Rance et al., 1989).

The relative spectral simplicity of the aromatic region (two tyrosines, one phenylalanine, and one histidine) allowed easy assignment of the ring system resonances of these amino acid spin systems by an analysis of the cross-peaks found in the DQF-COSY and RELAY spectra. These assignments are in agreement with those previously published on the basis of a one-dimensional analysis of this region of the spectrum (Kalbitzer et al., 1985). The identities of the four aromatic spin systems were established by the presence of clearly resolved cross-peaks in the ^2H_2O NOESY spectrum corresponding to NOEs between the $C^{\beta}Hs$ of previously delineated $NH-C^{\alpha}H-C^{\beta}H$ subsystems and ring proton resonances.

^1H Sequential Connectivities in the β-Sheet. Because downfield-shifted $C^{\alpha}H$ resonances have been shown to be derived from residues involved in the β-sheet structure of proteins (Wagner et al., 1983), including *E. coli* HPr (Klevit et al., 1986), the $C^{\alpha}H$ resonances exhibiting downfield shifts were chosen as a starting point for the sequential assignment procedure. However, assignment of these cross-peaks was hampered by two difficulties. First, only a few of the downfield-shifted $NH-C^{\alpha}H$ cross-peaks could be unambiguously associated with an amino acid type designation on the basis of analysis of the RELAY and TOCSY spectra. Second, overlap of the $NH-C^{\alpha}H$ cross-peaks and redundancy in the NH dimension made tracing the NOE connectivities impossible in some instances. A data set collected at a higher temperature (42 °C) was not helpful in resolving much of the chemical shift redundancy (data not shown).

Figure 4 illustrates how the NOESY and RELAY spectra obtained for the S46D mutant were used to resolve the chemical shift overlap within this region. The downfield portion of the fingerprint region is shown with the ^1H_2O NOESY spectrum plotted over the ^1H_2O RELAY spectrum. As an example, the spectrum of wild-type HPr (Figure 4A) shows severe overlap for the $NH-C^{\alpha}H$ cross-peaks of Ile_{61} and Ser_{64} in both ω_1 and ω_2. Two strong $d_{\alpha N}$ [$C^{\alpha}H(i)-NH(i+1)$] NOE cross-peaks occur at this same NH resonance position, and the backbone connectivity cannot be unambiguously traced. Figure 4B shows the same representation of the spectra obtained for the S46D mutant protein. The Ile_{61} and Ser_{64} cross-peaks are now resolved, and the $d_{\alpha N}$ connectivities can be followed easily. Other examples of the resolving power of these mutant spectra include resolution of the overlap of the NH resonance positions of Lys_4 and Thr_{62} and those of Tyr_{37} and Ala_{65}, allowing assignment of the $d_{\alpha N}$ cross-peaks at these positions. Of course, the spectra of the mutant HPr contain some chemical shift overlap that is not present in the spectra of the wild-type protein (e.g., the Glu_{36} and Phe_6 NH resonance positions). Therefore, it is through the combined use of the two sets of spectra that much of the overlap found in the wild-type and mutant protein spectra can be resolved.

Once a set of $NH-C^{\alpha}H$ coherence transfer cross-peaks connected by $d_{\alpha N}$ NOE connectivities was identified, the sequence-specific assignments for these residues were made on the basis of an amino acid type assignment involving one or more cross-peaks. Assignments were made for the four strands of the β-sheet structure in this manner by use of the italicized amino acid residues as starting points: A strand, Ala_2-Gln_3-Lys_4-*Thr_5*-*Phe_6*-Lys_7-Val_8; B strand, Ala_{31}-Asp_{32}-Val_{33}-Asn_{34}-Leu_{35}-Glu_{36}-*Tyr_{37}*-Asn_{38}; C strand, *Thr_{41}*-*Val_{42}*-Asn_{43}-Leu_{44}*; D strand, Glu_{60}-Ile_{61}-*Thr_{62}*-Ile_{63}-Ser_{64}-*Ala_{65}*-Ser_{66}-*Gly_{67}*.

Interstrand NOEs provided further verification of the β-sheet structure. $d_{\alpha\alpha}$ NOE connectivities characteristic of an antiparallel arrangement of strands were observed and are depicted in Figure 5 along with the other observed interstrand NOEs. The pattern of persistent amides is consistent with the placement of the A and C β-strands at the edges of the four-stranded β-sheet. The overall arrangement of the β-sheet is the same as that found in the NMR solution structure of *E. coli* HPr (Klevit & Waygood, 1986), although some minor differences are observed (see Discussion).

^1H Sequential Connectivities in the α-Helices. A similar approach to that just described for obtaining the β-sheet assignments was used to identify the amino acid residues in α-helical conformations. A set of d_{NN} [$NH(i)-NH(i+1)$] NOEs, indicative of residues involved in α-helices (Wüthrich, 1986), were present in the downfield portion of the ^1H_2O NOESY spectrum of wild-type HPr. However, ambiguities arose at several places where two or more spin systems had the same NH resonance frequency. Resolution of some of this overlap was again made possible by superposition of the wild-type and mutant spectra. Identical displacement in the NH chemical shift dimension for $NH-C^{\alpha}H$ and $NH-C^{\beta}H$ coherence transfer (RELAY) cross-peaks (giving connectivities within a particular amino acid spin system) and an $NH(i)-NH(i+1)$ NOESY cross-peak (giving sequential connectivity to other spin systems) indicated that the cross-peaks shared a common NH resonance. In this way, many ambiguities could be resolved.

To illustrate this point, Figure 6 shows the superposition of the wild-type and S46D spectra. There are a number of cross-peaks that are coincidental in the two spectra, while others are shifted. In this case, it can be seen that large displacements within the NH dimension allow an unambiguous assignment for some of the $NH-C^{\beta}H/NH-NH$ cross-peak pairs. Further verification of the identity of the shifted peaks was made possible by analysis of the extended spin systems in other spectra of the two proteins (data not shown). With

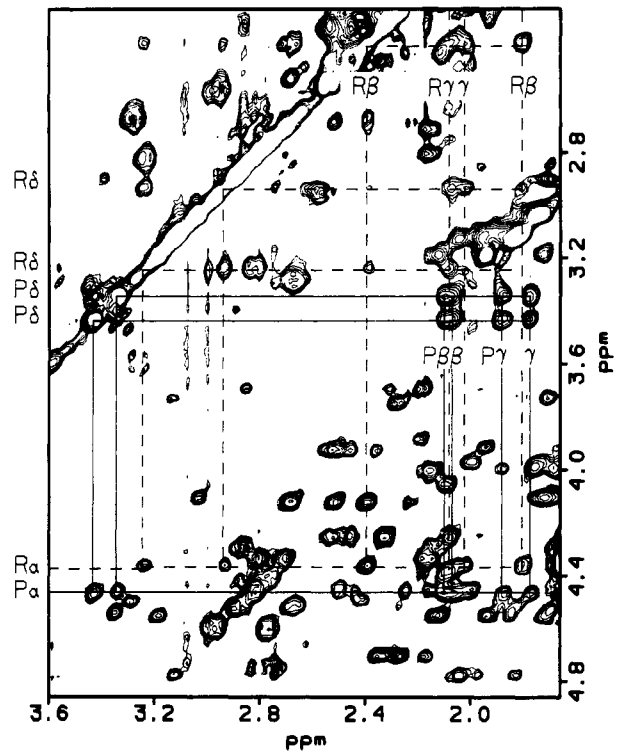

FIGURE 3: Aliphatic region of the TOCSY spectra recorded in ^2H_2O. The Arg_{17} (dashed lines) and Pro_{18} (solid lines) spin systems are traced. Resonance positions of the $C^\alpha H$, $C^\beta H$, $C^\gamma H$, and $C^\delta H$ protons are indicated by α, β, γ, and δ, respectively, preceded by either R (Arg_{17}) or P (Pro_{18}).

this type of analysis, the spin systems corresponding to residues Ala_{20} through Lys_{28}, making up the A α-helix, were assigned.

Another stretch of d_{NN} connectivities was identified and found to correspond to a long α-helix analogous to the C helix in the *E. coli* HPr structure. Assignment of the *B. subtilis* C helix was completed without relying on the mutant spectra because the S46D and S46A mutations have no effect on the chemical shift positions of the spin systems of the residues in the C helix. Sequential assignments for Asn_{75}/Ala_{76} and Thr_{80}/Met_{81} in the C helix were made by identifying $d_{\alpha_i N_{i+3}}$, $d_{\alpha_i \beta_{i+3}}$, and $d_{\beta_i N_{i+3}}$ connectivities because the sequential d_{NN} cross-peaks for these residue pairs were on the diagonal.

^1H Sequential Connectivities in Regions of Nonregular Secondary Structure and Turns. Figure 7 summarizes the sequential connectivities for *B. subtilis* HPr and outlines the locations of the A and C α-helices and the four strands of the β-sheet. It can be seen that there are two regions of the protein connected by sequential NOE connectivities that do not appear to adopt a regular secondary structure. The segment Gly_{13}-Ala_{19}, containing the active site His_{15}, appears to be in a predominantly extended conformation because all of the observed backbone NOEs are of the $d_{\alpha N}$ type. The sequential NOEs that connect the segment starting with Lys_{45} and extending through Ala_{56} also do not represent a region of well-defined secondary structure. However, the presence of a single turn of an α-helix in this segment is suggested by the d_{NN} connectivities connecting residues Val_{50} through Leu_{53} and the $d_{\alpha N_{i+3}}$ and $d_{\beta N_{i+3}}$ connectivities between Val_{50} and Leu_{53} (see Discussion).

The remaining stretches of NOE connectivities can be classified as turns that connect the regions described above. The sequence Thr_9–Ser_{12} makes up turn 1 that connects β-strand A with the extended region containing the active site His_{15}. Turn 2 (Tyr_{29}–Ala_{31}) is a very abrupt reversal of the backbone that connects the end of α-helix A with β-strand B. Turn 3 (Tyr_{37}–Lys_{40}) is a β-turn that reverses the direction of the neighboring β-strands B and C. Turn 4 (Leu_{44}–Lys_{47}) joins β-strand C with the extended segment Lys_{47}–Ala_{56}, and turn 5 (Lys_{57}–Glu_{60}) connects this segment with β-strand D. Turn 6 (Ala_{68}–Asn_{71}) reverses the direction of β-strand D, joining it to the α-helix C. The sequence-specific assignments are listed in Table I, and Figure 8 shows the fingerprint-region NH–$C^\alpha H$ cross-peaks labeled with the assignments.

Long-Range Connectivities and Global Fold. Table II contains a list of NOEs observed between spin systems corresponding to amino acid residues that are nonadjacent in the primary sequence (NOEs connecting the strands of the β-sheet, summarized in Figure 5, are omitted from this table). Some of these help to define the positions of the α-helices relative to the β-sheet. There is only one global fold, shown in Figure 9, consistent with these observations, and it corresponds to a right-handed structure in which the α-helices lie on the same side of a four-stranded antiparallel β-sheet.

Discussion

On the basis of the functional and primary sequence similarities, Gram-positive HPr was expected to have secondary and tertiary structures (and hence NOE patterns) similar to those determined for Gram-negative HPr. However, the assignments described here were made de novo, and some assignments were impossible due to spectral overlap when the wild-type spectrum alone is used. To overcome this problem, extensive use was made of 2D NMR spectra of point mutants of HPr. Other groups have successfully used analysis of naturally occurring variants and site-directed mutant proteins predominantly for verification of assignments of the mutated residues [for examples, see Markley et al. (1984) and Gronenborn et al. (1986)]. In the present study, our emphasis has been to exploit the chemical shift perturbations that many of the unmutated residues undergo in the spectra of the mutant proteins. This approach has also been utilized to some extent by others to help assign the proton resonances of hirudin (Folkers et al., 1989). These results suggest that concurrent assignment of the mutant protein spectra along with that of the wild-type protein can be extremely helpful. The extra time spent acquiring and assigning the additional spectra is more than offset by the gain in resolution and the increase in reliability of the resulting assignments.

The overall fold of the protein has α-helices on one side and a single antiparallel β-sheet on the other with a well-defined hydrophobic core at the interface of these structural elements. This type of fold has been termed an "open-face sandwich" structure (Richardson, 1981). The global fold of *B. subtilis* HPr is the same as that determined for *E. coli* HPr on the basis of NMR analysis (Klevit & Waygood, 1986). However, several differences are found when details of the two structures are compared.

In the β-sheet, there is no evidence in the *B. subtilis* HPr spectrum for a β-bulge, between residues 35 and 63, as suggested by an NOE analysis of the *E. coli* β-sheet structure (Klevit & Waygood, 1986). In the *B. subtilis* spectrum, all of the amide protons forming hydrogen bonds between the B and D strands of the β-sheet are persistent to exchange and all $d_{\alpha N}$ connectivities are present along these strands. This conclusion is further supported by the presence of many cross-strand NOEs between the two β-strands, although the diagnostic $d_{\alpha\alpha}$ connectivity between residues 35 and 63, which is absent in the *E. coli* spectrum, is on the diagonal and is unobservable due to the redundancy of the $C^\alpha H$ resonances of Leu_{35} and Ile_{63} in the *B. subtilis* spectrum. It is likely that the unfavorable packing of the C^β-branched side chains of Val_{35} and Ile_{63} in the *E. coli* protein is relieved by the sub-

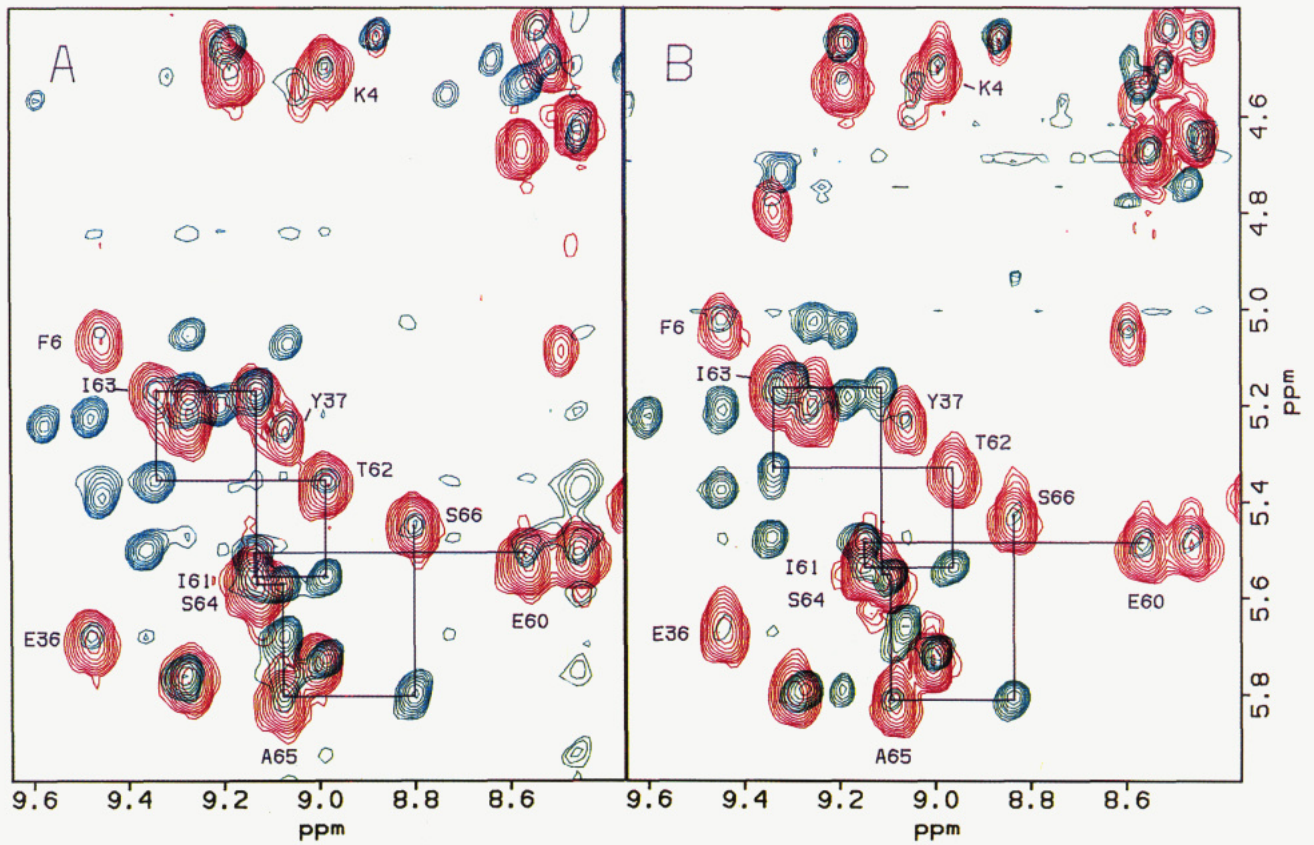

FIGURE 4: $d_{\alpha N}$ connectivities among the residues forming the β-sheet of wild-type and mutant HPrs. The downfield portions of the fingerprint regions of the RELAY (red) and NOESY (green) spectra are overlaid, and sequential $d_{\alpha N}$ connectivities from β-strand D (Glu_{60} through Ser_{66}) are traced for wild-type HPr (panel A) and the S46D mutant (panel B). Relevant sequence-specific assignments are indicated with the one-letter amino acid code followed by the residue number.

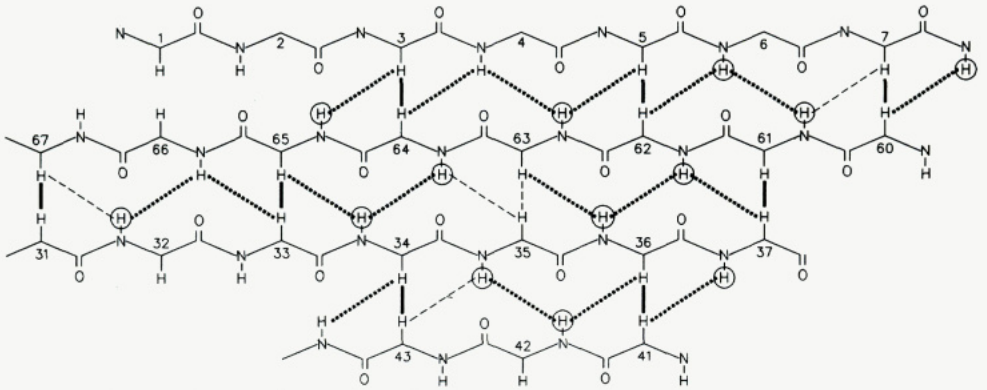

FIGURE 5: β-Sheet structure of HPr. Diagram of the antiparallel β-sheet determined by long-range NOEs. Short, solid lines represent strong $d_{\alpha\alpha}$ connectivities while the dark, dotted lines depict weak $d_{\alpha N}$ and d_{NN} NOEs. Potential cross-strand NOEs that were not detectable due to spectral overlap are denoted with thin, dashed lines. The amide protons that are circled represent persistent NHs.

stitution of leucine at position 35 in *B. subtilis* HPr (Richardson, 1981).

Another difference between the two β-sheet structures is the nature of the secondary structure of the N-terminus. Met_1 and Phe_2 are part of the β-sheet structure of *E. coli* HPr as evidenced by clearly resolved cross-strand NOEs. However, in the *B. subtilis* HPr spectra, the amide resonances of Met_1 and Ala_2 were not observed (see Table I). This may be due to dynamic properties of the amide protons or to spectral overlap. It is therefore unclear whether these two residues form part of the β-sheet structure of *B. subtilis* HPr.

The region of *B. subtilis* HPr that contains the active site His_{15} differs from the corresponding segment from *E. coli* HPr in the nature of the sequential connectivities connecting some of the residues. For example, *B. subtilis* HPr has $d_{\alpha N}$ connectivities connecting Gly_{13}-Ile_{14}-His_{15} while *E. coli* HPr has d_{NN} connectivities joining Gly_{13}-Leu_{14}-His_{15}. However, NOEs between the side-chain resonances of His_{15} and Arg_{17} indicate that this region may retain some of the kinked structure that has been suggested for the active site of *E. coli* HPr (Klevit & Waygood, 1986). In the other region of HPr made up of nonregular secondary structure, a single turn of an α-helix is suggested for *B. subtilis* HPr by d_{NN} connectivities and by one set of $d_{\beta N_{i+3}}$ and $d_{\alpha_i N_{i+3}}$ NOEs. However, the exact location of this helix is not conserved between the two HPrs, with the *E. coli* B α-helix being made up of residues 46–52 and the putative one-turn α-helix occurring at residues 50–53 in *B. subtilis* HPr. Neither of these helices is well-defined by the

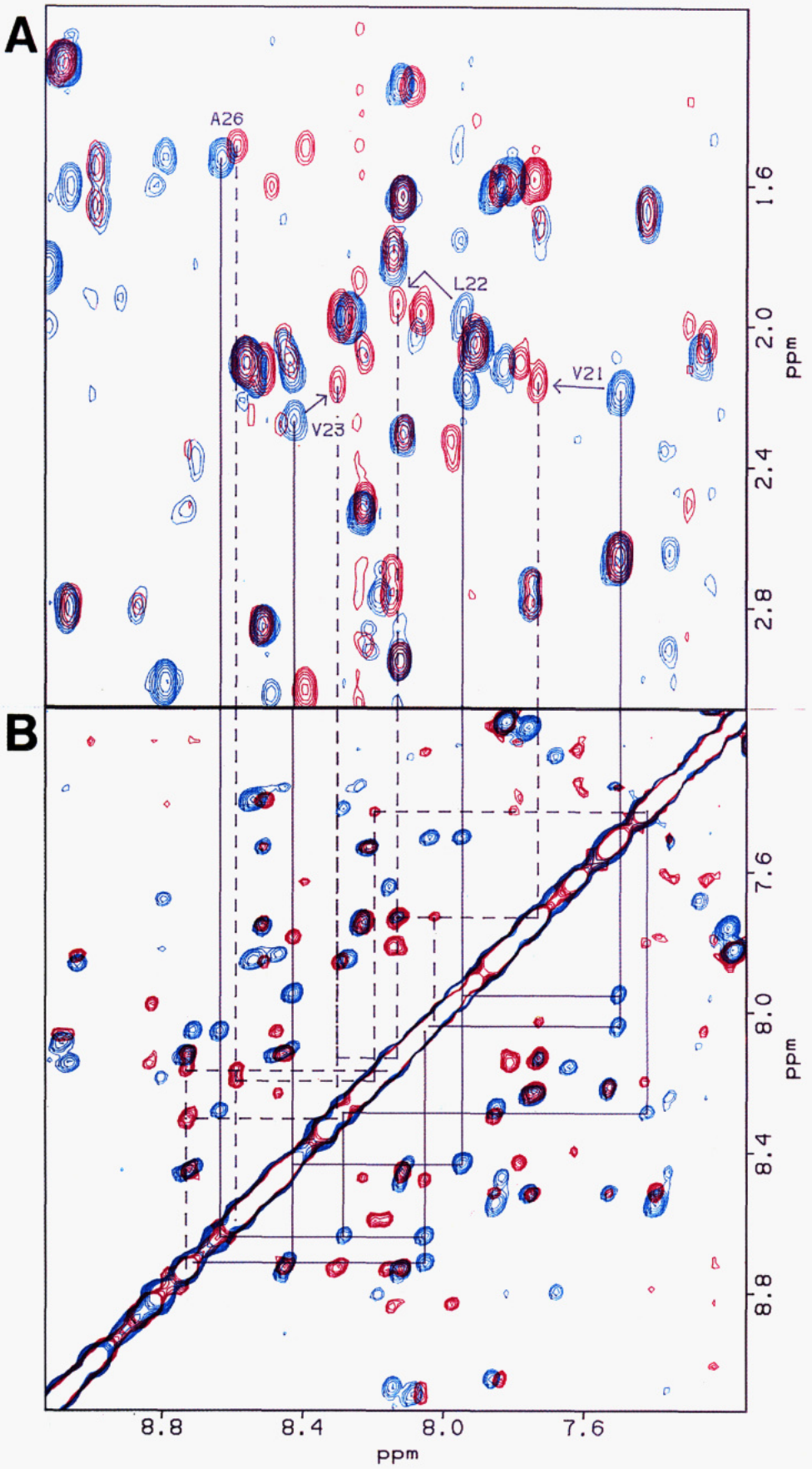

FIGURE 6: Spectral overlap resolution in amide region for spin systems of amino acids involved in α-helix A. (A) ^1H_2O RELAY spectrum showing NH–C$^\beta$H cross-peaks. (B) ^1H_2O NOESY spectrum featuring NH–NH cross-peaks. In both panels wild-type HPr spectrum is blue and S46D mutant spectrum is red. NH–C$^\beta$H to NH–NH cross-peaks are connected, and NH–NH connectivities are traced for both wild-type (—) and mutant (- - -) spectra for residues comprising helix A (residues 20–28). Arrows in the upper portion indicate NH–C$^\beta$H cross-peaks undergoing large position shifts. NOESY spectrum has been symmetrized for clarity. Weak NH–NH NOESY cross-peak corresponding to connectivity between Leu_{22} and Val_{23} in the S46D spectrum is not observable due to symmetrization treatment.

Table I: ^{1}H NMR Chemical Shifts (ppm) of Wild-Type *B. subtilis* HPr[a]

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	other
M1						
A2		4.35	1.63			
Q3	8.99	5.73	2.17, 1.90	2.29, 2.24		
K4	8.97	4.50	1.69, 1.55	1.29, 1.27	1.06, 1.04	($C^{\epsilon}H$) 2.89
T5	8.32	5.39	3.90	1.30		
F6	9.44	5.03	2.94, 2.55		7.13	($C^{\epsilon}H$) 7.22, ($C^{\zeta}H$) 6.99
K7	9.25	5.19	1.96, 1.93	1.79	1.46	($C^{\epsilon}H$) 3.68, 3.11
V8	9.19	4.47	2.41	1.06, 0.93		
T9	9.17	4.52	4.45	1.24		
A10	6.99	4.44	1.63			
D11	8.86	4.43	2.80			
S12	8.59	4.61	4.04, 3.98			
G13	7.94	4.51, 3.33				
I14	7.89	4.23	2.07	1.31, 1.05	($C^{\delta}H_3$) 0.66	($C^{\gamma}H_3$) 0.67
H15	6.72	4.98	3.56, 3.25		7.25	($C^{\epsilon}H$) 7.94
A16	8.79	3.92	1.60			
R17	8.24	4.34	2.37, 1.79	2.07, 2.01	3.23, 2.93	
P18		4.44	2.10, 2.07	1.87, 1.77	3.42, 3.33	
A19	8.11	3.75	1.31			
T20	8.02	3.82	4.57	1.30		
V21	7.49	3.86	2.17	1.05, 1.00		
L22	7.93	3.95	1.97, 1.40	1.72	0.73	
V23	8.42	4.26	2.12	1.02, 0.93		
Q24	8.70	4.10	2.38, 2.21	2.67, 2.50		
T25	8.03	4.07	4.50	1.33		
A26	8.62	4.03	1.52			
S27	8.26	4.25				
K28	7.41	4.08	1.73, 1.67	1.37		($C^{\epsilon}H$) 3.03
Y29	7.50	4.49	3.24, 2.66		7.35	($C^{\epsilon}H$) 6.80
D30	11.24	4.57	2.76			
A31	9.56	4.26	1.06			
D32	8.17	4.72	2.77, 2.72			
V33	9.11	5.16	1.83	0.95, 0.90		
N34	9.26	5.76	2.62, 2.53			
L35	9.27	5.21	1.67, 1.53	1.60	0.85	
E36	9.46	5.67	2.15, 1.99	2.22, 2.09		
Y37	9.06	5.23	3.21, 2.81		7.35	($C^{\epsilon}H$) 7.07
N38	9.56	4.29	2.87, 1.65			
G39	9.08	4.22, 3.72				
K40	8.07	4.75	2.01, 1.95	1.81, 1.55	1.49	($C^{\epsilon}H$) 3.10
T41	8.45	5.50	3.93	1.15		
V42	9.35	4.83	2.26	0.99, 0.90		
N43	8.48	5.04	3.06, 3.03			(δNH_2) 7.73, 7.18
L44	9.05	4.22	1.59, 1.18	1.94	0.98, 0.95	
K45	8.13	4.49	1.81			
S46	7.63	4.88	4.24, 3.86			
I47	8.90	3.89	1.93	1.64, 1.24	($C^{\delta}H_3$) 1.02	($C^{\gamma}H_3$) 1.04
M48	8.41	3.72	2.26, 1.09			
G49	8.07	4.09, 3.96				
V50	7.93	3.69	2.17	1.12, 0.84		
M51	8.43	4.31	2.15			
S52	8.11	4.56				
L53	7.26	4.37	2.08, 1.65		1.12	
G54	7.67	4.06, 3.84				
I55	8.78	3.01	1.53	1.37, 0.77	($C^{\delta}H_3$) −0.08	($C^{\gamma}H_3$) 0.58
A56	7.80	4.67	1.58			
K57	8.44	3.69	2.20, 1.68	1.38, 1.29		($C^{\epsilon}H$) 3.11
G58	9.02	4.56, 3.72				
A59	7.85	4.54	1.61			
E60	8.55	5.51	2.24, 2.10	2.41, 2.22		
I61	9.12	5.54	2.00	1.79, 1.15	($C^{\delta}H_3$) 0.98	($C^{\gamma}H_3$) 0.74
T62	8.97	5.34	4.13	1.13		
I63	9.33	5.15	1.94	1.83, 1.08	($C^{\delta}H_3$) 0.92	($C^{\gamma}H_3$) 0.75
S64	9.11	5.56	3.91, 3.81			
A65	9.06	5.80	1.25			
S66	8.78	5.43	3.88, 3.78			
G67	9.96	4.76, 4.11				
A68	9.20	4.24	1.60			
D69	8.20	5.19	3.38, 2.87			
E70	7.52	3.67	2.28, 1.95, 2.84			
N71	8.50	4.49	2.84			(δNH_2) 7.82, 7.17
D72	7.75	4.39	2.78, 2.69			
A73	8.22	2.52	0.89			
L74	7.72	3.97	1.70	1.86	0.95, 0.85	
N75	8.12	4.57	2.94, 2.94			(δNH_2) 7.83, 6.91
A76	8.10	4.32	1.64			
L77	8.44	4.03	2.08, 1.18	1.78	1.01, 0.89	

Table I (Continued)

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	other
E78	8.72	3.90	2.52, 2.33	2.48, 2.43		
E79	8.10	4.23	2.32, 2.31	2.52, 2.44		
T80	8.47	4.18	4.24	1.17		
M81	8.44	3.98	2.14, 1.59			
K82	7.82	4.52	2.11	1.83, 1.70	1.91	($C^{\epsilon}H$) 3.16
S83	9.30	4.30	4.16, 4.10			
E84	8.26	4.43	2.09, 1.98	2.50, 2.22		
G85	7.84	4.06, 3.99				
L86	8.52	4.43	2.14, 1.27	1.68	0.86, 0.81	
G87	7.38	4.82, 4.30				
Q88	8.56	4.68	2.24, 2.16	2.34, 2.31		

[a] At 30 °C in 50 mM phosphate buffer, pH 6.9, with ppm relative to [^2H]TSP at 0.00 ppm.

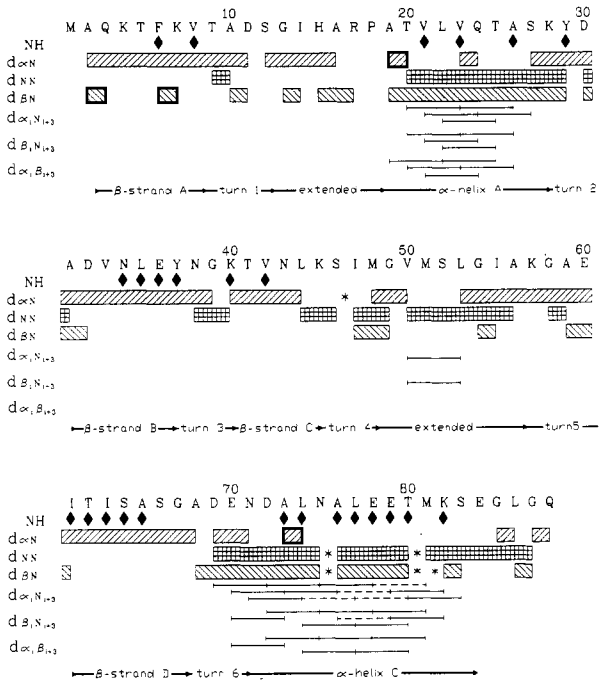

FIGURE 7: Summary of the sequential NOE connectivities observed for *B. subtilis* HPr. The three types of nearest-neighbor connectivities used for the sequential assignments ($d_{\alpha N}$, d_{NN}, and $d_{\beta N}$) are summarized along with the $d_{\alpha N}$, $d_{\beta N}$, and $d_{\alpha\beta}$ connectivities observed between residues i and $i+3$. The presence of persistent backbone amide protons is indicated by the solid diamonds. NOE connectivities that may exist but were unobservable due to spectral overlap are denoted by the asterisks. Secondary structural elements of *B. subtilis* HPr as deduced from the short-range and medium-range connectivities are indicated.

Table II: Medium- and Long-Range NOEs between Residues of Wild-Type *B. subtilis* HPr

residues belonging to	NOEs
active site–β-sheet	Gly_{13}–Phe_6
active site	His_{15}–Arg_{17}
extended–β-sheet	Ile_{55}–Tyr_{37}
	Leu_{53}–Tyr_{37}
	Ala_{59}–Tyr_{37}
C helix–β-sheet	Glu_{78}–Phe_6
turn 1–turn 5	Val_8–Gly_{58}
	Val_8–Ala_{59}
active site–extended	Gly_{13}–Ala_{56}
	Arg_{17}–Ile_{55}
	Pro_{18}–Ile_{55}
A helix–extended	Ala_{19}–Val_{50}
A helix–C helix	Tyr_{29}–Asp_{69}
	Tyr_{29}–Asp_{72}
	Tyr_{29}–Ala_{73}
turn 2–turn 6	Ala_{31}–Asp_{69}
	Ala_{31}–Glu_{70}
turn 6	Ser_{66}–Glu_{70}
	Gly_{67}–Glu_{70}

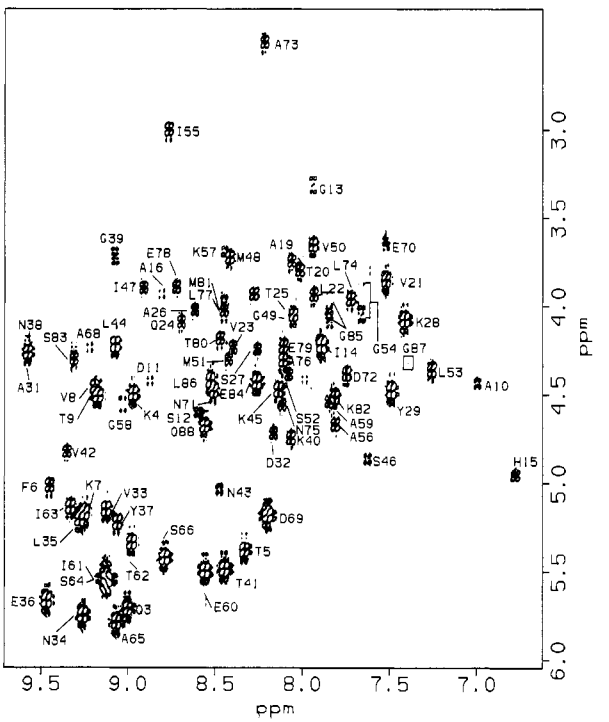

FIGURE 8: Sequence-specific assignments for the NH–$C^{\alpha}H$ cross-peaks of HPr. The fingerprint region of the phase-sensitive DQF-COSY spectrum in ^1H_2O is shown with assignments derived from the sequence-specific assignment procedure. This spectrum was obtained with the SCUBA pulse sequence (Brown et al., 1988) for water suppression.

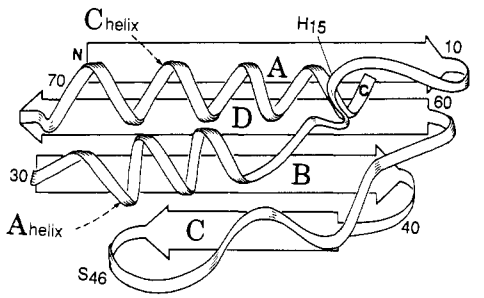

FIGURE 9: Tertiary folding of *B. subtilis* HPr. A model of the solution structure of HPr based on the NOE data is shown. The labels assigned to the β-strands and α-helices (large letters) are consistent with those used in the *E. coli* HPr model (Klevit & Waygood, 1986). The small letters N and C denote the amino- and carboxy-termini, respectively. The numbers represent the residue number of the amino acid at that position in the structure. The region between Ser_{46} and Glu_{60} may include a short α-helix between residues 50 and 53 (see text).

presence of persistent amides or an extensive network of i to $i+3$ connectivities as seen in the A and C α-helices. Overall, the segments of nonregular secondary structure (residues 12–19 and 46–57) appear to be in somewhat different con-

formations in the two HPr structures.

The availability of the complete sequence-specific ^1H NMR assignments for *B. subtilis* HPr forms the basis for the ongoing quantitative analysis of the NOESY data with the goal of refining the solution structure. The refined solution structure of *B. subtilis* HPr will be useful for defining the environment around the active site and interpreting the results of mutational studies (Eisermann et al., 1988; Reizer et al., 1989) and for comparison studies with the *E. coli* HPr structure. These results will form the basis for studies concerning the effects of phosphorylation at His_{15} and at Ser_{46} on the structure of HPr.

Added in Proof

A sample of the *B. subtilis* HPr used in the NMR studies was subjected to N-terminal sequence analysis. This protein, expressed in *E. coli*, lacks Met_1 and begins with Ala_2 (Hai Le Trong and Ken Walsh, personal communication). This result is consistent with our inability to detect the amide proton resonances from Met_1 and Ala_2 in the 2D NMR spectra.

Acknowledgments

We thank Gary Drobny for the use of the AM-500 and home-built 500-MHz spectrometers, Jon Herriott, Melissa Starovasnik, and Bruce Waygood for helpful discussions, and Bev Castner for technical assistance. J.R. thanks Milton Saier for encouragement and support.

References

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Bax, A., & Drobny, G. (1985) *J. Magn. Reson. 61*, 306–320.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) *Mol. Phys. 48*, 535–560.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.

Deutscher, J., Kessler, U., Alpert, C. A., & Hengstenberg, W. (1984) *Biochemistry 23*, 4455–4460.

Eich, G., Bodenhausen, G., & Ernst, R. R. (1985) *J. Am. Chem. Soc. 104*, 3732–3733.

Eisermann, R., Deutscher, J., Gonzy-Treboul, G., & Hengstenberg, W. (1988) *J. Biol. Chem. 263*, 17050–17054.

Folkers, P. J. M., Clore, G. M., Driscoll, P. C., Dodt, J., Kohler, S., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2601–2617.

Gonzy-Treboul, G., Zagoree, M., Rain-Guion, M. C., & Steinmetz, M. (1989) *Mol. Microbiol. 3*, 103–112.

Gronenborn, A. M., Clore, G. M., Schmeissner, U., & Wingfield, P. (1986) *Eur. J. Biochem. 161*, 37–43.

Kalbitzer, H. R., Hengstenberg, W., Rosch, P., Muss, P., Bernsmann, P., Engelmann, R., Dorsschug, M., & Deutscher, J. (1982) *Biochemistry 21*, 2879–2885.

Kalbitzer, H. R., Muss, H. P., Engelmann, R., Kiltz, H. H., Stuber, K., & Hengstenberg, W. (1985) *Biochemistry 24*, 4562–4569.

Klevit, R. E., & Drobny, G. (1986) *Biochemistry 25*, 7770–7773.

Klevit, R. E., & Waygood, E. B. (1986) *Biochemistry 25*, 7774–7781.

Klevit, R. E., Drobny, G., & Waygood, E. B. (1986) *Biochemistry 25*, 7760–7769.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Markley, J. L., Westler, W. M., Chan, T.-M., Kojiro, C. L., & Ulrich, E. L. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 2648–2656.

Maurer, W., Ruterjans, H., Schrecker, O., Hengstenberg, W., Gassner, M., & Stehlik, D. (1977) *Eur. J. Biochem. 75*, 297–301.

Postma, P. W., & Lengeler, J. W. (1985) *Microbiol. Rev. 49*, 232–269.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Rance, M., Chazin, W. J., Dalvit, C., & Wright, P. E. (1989) *Methods Enzymol. 176*, 114–134.

Reizer, J. (1989) *FEMS Microbiol. Rev. 63*, 149–156.

Reizer, J., & Peterkofsky, A. (1987) in *Sugar Transport and Metabolism in Gram-Positive Bacteria* (Reizer, J., & Peterkofsky, A., Eds.) pp 333–364, Ellis Horwood, Chichester, England.

Reizer, J., Novotny, M. J., Hengstenberg, W., & Saier, M. H. (1984) *J. Bacteriol. 160*, 333–340.

Reizer, J., Saier, M. H., Deutscher, J., Grenier, F., Thompson, J., & Hengstenberg, W. (1988a) *CRC Crit. Rev. Microbiol. 15*, 297–338.

Reizer, J., Peterkofsky, A., & Romano, A. H. (1988b) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2041–2045.

Reizer, J., Sutrina, S. L., Saier, M. H., Stewart, G. C., Peterkofsky, A., & Reddy, P. (1989) *EMBO J. 8*, 2111–2120.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 168–363.

Romano, A. H., Brino, G., Peterkofsky, A., & Reizer, J. (1987) *J. Bacteriol. 169*, 5589.

Saier, M. H., Jr. (1985) *Mechanisms and Regulation of Carbohydrate Transport in Bacteria*, Academic Press, Orlando, FL.

Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948–5949.

Wittekind, M., Reizer, J., Deutscher, J., Saier, M. H., & Klevit, R. E. (1989) *Biochemistry 28*, 9908–9912.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311–319.