Polypeptide Backbone Resonance Assignments and Secondary Structure of *Bacillus subtilis* Enzyme III^{glc} Determined by Two-Dimensional and Three-Dimensional Heteronuclear NMR Spectroscopy†

Wayne J. Fairbrother,‡ John Cavanagh,‡,§ H. Jane Dyson,‡ Arthur G. Palmer, III,‡ Sarah L. Sutrina,∥ Jonathan Reizer,∥ Milton H. Saier, Jr.,∥ and Peter E. Wright*,‡

Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037, and Department of Biology, University of California at San Diego, La Jolla, California 92093-0116

Received February 8, 1991

ABSTRACT: The enzyme III^{glc}-like domain of *Bacillus subtilis* II^{glc} (III^{glc}; 162 residues, 17.4 kDa) has been cloned and overexpressed in *Escherichia coli*. Sequence-specific assignment of the backbone ^1H and ^{15}N resonances has been carried out with a combination of homonuclear and heteronuclear two-dimensional and heteronuclear three-dimensional (3D) NMR spectroscopy. Amide proton solvent exchange rate constants have been determined from a series of ^1H–^{15}N heteronuclear single-quantum coherence (HSQC) spectra acquired following dissolution of the protein in D_2O. Major structural features of III^{glc} have been inferred from the pattern of short-, medium- and long-range NOEs in 3D heteronuclear ^1H nuclear Overhauser effect ^1H–^{15}N multiple-quantum coherence (3D NOESY-HMQC) spectra, together with the exchange rate constants. III^{glc} contains three antiparallel β-sheets comprised of eight, three, and two β-strands. In addition, five β-bulges were identified. No evidence of regular helical structure was found. The N-terminal 15 residues of the protein appear disordered, which is consistent with their being part of the Q-linker that connects the C-terminal enzyme III^{glc}-like domain to the membrane-bound II^{glc} domain. Significantly, two histidine residues, His 68 and His 83, which are important for phosphotransferase function, are found from NOE measurements to be in close proximity at the ends of adjacent strands in the major β-sheet.

The phosphoenolpyruvate/sugar phosphotransferase system (PTS)[1] of a number of bacteria is responsible for the concomitant uptake and phosphorylation of PTS sugar substrates. In addition, the PTS regulates the uptake of several non-PTS sugars, by both transcriptional and nontranscriptional mechanisms [for reviews, see Reizer et al. (1988), Saier (1989), and Meadow et al. (1990)]. Phosphorylation of the glucose-specific enzyme III (III^{glc}) of the PTS is thought to play a central role in these regulatory processes. The phosphorylated form of III^{glc} acts as an allosteric activator of adenylate cyclase, while the free (unphosphorylated) form functions as an allosteric inhibitor of the permeases and catabolic enzymes required to produce endogenous inducers of non-PTS operons. For example, III^{glc} has been shown to interact with lactose permease and glycerol kinase to inhibit their activities (Osumi & Saier, 1982; Nelson et al., 1983; Saier et al., 1983; Novotny et al., 1985; de Boer et al., 1986).

The soluble III^{glc} of *Escherichia coli* is phosphorylated at N3 of the imidazole ring of His 90 by phospho-HPr, the general phosphoryl carrier protein of the PTS (Dörschug et al., 1984; Meadow & Roseman, 1982; Kalbitzer et al., 1981). During the uptake of PTS sugars a phosphoryl group is transferred from phospho-III^{glc} to the sugar via the membrane-bound glucose permease, enzyme II^{glc}. The site of III^{glc} phosphorylation has recently been confirmed as His 90 by site-directed mutagenesis (Presper et al., 1989). Interestingly, mutation of the only other histidine (His 75) in *E. coli* III^{glc} to glutamine results in a protein that can accept but cannot transfer phosphoryl groups, suggesting that His 75 may also be a critical amino acid for III^{glc}-mediated signaling mechanisms (Presper et al., 1989).

Bacillus subtilis lacks a distinct III^{glc} and instead possesses a membrane-bound glucose permease with a C-terminal III^{glc}-like domain covalently linked to the integral membrane II^{glc} domain via a Q-linker (Sutrina et al., 1990). The 162-residue III^{glc}-like domain of the glucose permease of *B. subtilis* has recently been overexpressed in *E. coli* and shown to fold independently of the N-terminal domain (Sutrina et al., 1990). This III^{glc}-like C-terminal domain assumes both the functional and regulatory roles of the soluble III^{glc} when expressed in *crr* (III^{glc} deficient) mutants of *E. coli* [Sutrina et al., 1990; Dean et al., 1990; Reizer and Saier, (in preparation)]. Site-directed mutagenesis establishes that phosphorylation by phospho-HPr occurs at His 83. In addition, replacement of His 68 by alanine results in a protein that can accept but not transfer phosphoryl groups (Reizer and Saier, unpublished results), in exact analogy with the results from *E. coli*.

† This work was supported by Grant GM-36643 from the National Institutes of Health (P.E.W.). W.J.F. was supported by a Damon Runyon–Walter Winchell Cancer Research Fund Fellowship, DRG-1059. A.G.P. was supported by a National Science Foundation postdoctoral fellowship in chemistry, under Grant CHE-8907510 awarded in 1989.

* To whom correspondence should be addressed.

‡ Research Institute of Scripps Clinic.

§ Present address: Department of Chemistry, University of Cambridge, Cambridge CB2 1EW.

∥ University of California, San Diego.

[1] Abbreviations: PTS, phosphoenolpyruvate/sugar phosphotransferase system; III^{glc}, glucose specific enzyme III; NMR, nuclear magnetic resonance; 2D, two dimensional; 3D, three dimensional; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single-quantum coherence spectroscopy; HMQC, heteronuclear multiple-quantum coherence spectroscopy; NOE, nuclear Overhauser effect; 2QF-COSY, double-quantum filtered correlation spectroscopy; 2Q, double-quantum spectroscopy; R-COSY, relayed correlation spectroscopy; TPPI, time-proportional phase incrementation; pH*, pH meter reading without correction for the deuterium isotope effect.

0006-2960/91/0430-6896$02.50/0

Little is currently known about the structure of III^{glc}, the mechanism of phosphoryl transfer, or the conformational changes that may occur upon phosphorylation. The overexpressed III^{glc}-like domain of *B. subtilis* is stable, soluble, and of an appropriate size for solution structure determination using modern nuclear magnetic resonance (NMR) spectroscopic techniques.

Advances in NMR spectroscopy during the last decade (Clore & Gronenborn, 1989; Bax, 1989; Wüthrich, 1986, 1989; Ernst et al., 1987) have made possible the assignment of proton resonances of small proteins (M_r <10000) and subsequent determination of their structures in solution. Homonuclear ^{1}H 2D NMR methods are limited, however, in the case of larger proteins, where the increased number of protons leads to severe spectral overlap and the larger molecular mass results in increased rotational correlation times and consequently broader resonances. Several approaches have been taken to improve the resolution and sensitivity of the NMR method and thus extend its usefulness to larger systems. One approach that has proven successful is ^{1}H–^{15}N (or ^{1}H–^{13}C) heteronuclear correlated spectroscopy. Proteins that have been overexpressed, such as III^{glc}, can be labeled uniformly with ^{15}N (and/or ^{13}C), which allows the implementation of a number of heteronuclear 2D and 3D NMR techniques (McIntosh & Dahlquist, 1990; Fesik & Zuiderweg, 1990). The larger chemical shift dispersion of the ^{15}N nucleus in heteronuclear 2D spectra aids in the resolution of ambiguities involving amide protons, which are often present in 2D ^{1}H NMR spectra. Experiments have been based on one-bond correlations and on correlations relayed by using NOESY and TOCSY mixing schemes. The heteronuclear single-quantum coherence (HSQC) schemes have been shown to provide improved resolution in the ^{15}N dimension relative to the heteronuclear multiple-quantum coherence (HMQC) experiments (Bax et al., 1990a; Norwood et al., 1990). Heteronuclear 3D NMR experiments give even greater resolution by using the heteronuclear chemical shift as the third dimension (Fesik & Zuiderweg, 1990; Marion et al., 1989ab). The usefulness of 3D TOCSY-HMQC and NOESY-HMQC experiments for the assignment and secondary structure determination of a 153-residue protein, interleukin 1β, has recently been demonstrated (Driscoll et al., 1990a,b). This paper reports the sequence specific assignments of the backbone resonances of the III^{glc}-like domain of *B. subtilis* II^{glc} (162 residues, 17.4 kDa), obtained from 2D homonuclear and heteronuclear and 3D heteronuclear NMR experiments, and the elucidation of the secondary structure based on short-, medium-, and long-range NOEs.

Experimental Procedures

Preparation of III^{glc}. The overproduction and purification of the soluble III^{glc}-like domain of *B. subtilis* II^{glc} (hereafter referred to as III^{glc}) was as described previously [Sutrina et al., 1990; Reizer et al. (in preparation)]. Uniform ^{15}N labeling (>95%) of the protein was performed by growing the cells in a minimal medium with $(^{15}NH_4)_2SO_4$ as the sole nitrogen source. NMR samples contained 10 mM potassium phosphate, pH 6.6 (92% H_2O/8% D_2O or 99.9% D_2O) with protein concentrations of ~0.8 mM.

NMR Measurements. All NMR spectra were recorded on a Bruker AM600 spectrometer, equipped with an Aspect 3000 computer, digital phase shifting hardware and fast switching of the ^{1}H decoupler and ^{15}N transmitter power levels. Spectra were recorded at 298 and 308 K and were referenced to internal dioxane (3.75 ppm downfield of TSS).

Standard pulse sequences and phase cycling were employed for the following 2D homonuclear experiments: 2QF-COSY (Rance et al., 1983), R-COSY (Wagner, 1983), and double-quantum (2Q) (Braunschweiler et al., 1984; Rance & Wright, 1986). TOCSY spectra were recorded by the improved high-sensitivity method of Cavanagh and Rance (1990) with the DIPSI-2 sequence (Shaka et al., 1988) for isotropic mixing and spin-lock periods of 60 and 80 ms. NOESY spectra (Bodenhausen et al., 1984) with mixing times of 100 ms and a TOCSY relayed-NOESY spectrum with a DIPSI-2 spin-lock period of 60 ms and a NOESY mixing time of 100 ms (Rance, 1987; Kessler et al., 1988) were acquired with the normal pulse sequences followed by a short Hahn-echo period to improve the quality of the baseline (Rance & Byrd, 1983; Davis, 1989). Low-power phase-locked presaturation of the H_2O signal was used for all of the above experiments. A 150 ms NOESY spectrum in H_2O solution was also recorded at 298 K using a jump-and-return read pulse (Bax et al., 1987). Spectra were usually acquired with 4K complex data points with an F_2 spectral width of 12.5 kHz, an F_1 spectral width of ~7.0 kHz, and 400–900 increments in F_1. Typically, 64–96 scans were collected per t_1 increment. The carrier was set on the solvent resonance for all experiments, and a recycle delay of 1.2–1.5 s was used. All 2D spectra were recorded in the phase-sensitive mode using time-proportional phase incrementation (TPPI) for quadrature detection in F_1 (Redfield & Kunz, 1975; Bodenhausen et al., 1980; Marion & Wüthrich, 1983). An unlabeled sample of III^{glc} was used for all homonuclear 2D spectra recorded in H_2O solution. Homonuclear 2QF-COSY, 2Q, TOCSY, and NOESY spectra were also acquired with a uniformly ^{15}N-labeled sample in D_2O solution. In these cases ^{15}N decoupling was achieved by applying a composite 180° ^{15}N pulse in the middle of the t_1 period and GARP-1 phase modulation (Shaka et al., 1985) during acquisition.

Two-dimensional ^{1}H–^{15}N correlated spectra (HSQC, HSQC-NOESY, and HSQC-TOCSY) were acquired by using published pulse sequences (Bax et al., 1990a; Norwood et al., 1990). ^{15}N decoupling was carried out with GARP-1 phase modulation, and DIPSI-2 was used for isotropic mixing in HSQC-TOCSY experiments. An additional HSQC-TOCSY experiment was acquired by a new sensitivity-enhanced method, which gives up to $\sqrt{2}$ signal-to-noise improvement (Cavanagh et al., 1991). Solvent suppression was achieved either by low-power presaturation or by the use of a spin-lock purge pulse (Messerle et al., 1989) and a 3-ms homospoil pulse. The ^{15}N–^{1}H J coupling evolution and refocusing delays of the INEPT (Morris & Freeman, 1979) portions of the pulse sequences were set to 2.3 ms [slightly shorter than $1/(4J_{NH})$ in order to minimize relaxation effects]. A total of 4K complex points were acquired in F_2, with a spectral width of 12.5 kHz and the ^{1}H carrier placed on the H_2O signal; 450 increments were collected in F_1 with a spectral width of 3.3 kHz and the ^{15}N carrier at 107.3 ppm [the ^{15}N frequency scale was indirectly referenced to liquid NH_3 by using the ^{1}H frequency of the H_2O resonance (Live et al., 1984; Bax & Subramanian, 1986)].

The heteronuclear 3D TOCSY-HMQC and NOESY-HMQC pulse schemes used were similar to those reported previously (Fesik & Zuiderweg, 1990; Driscoll et al., 1990a; Marion et al., 1989a,b) except that spin-lock purge pulses were used for suppression of the solvent signal (Messerle et al., 1989). The actual pulse sequences used are illustrated in Figure 1. The two experiments were run consecutively in order to minimize differences in sample conditions.

Data Processing. All 2D spectra were processed on a Convex C-240 using FTNMR (Hare Research). Where spin-

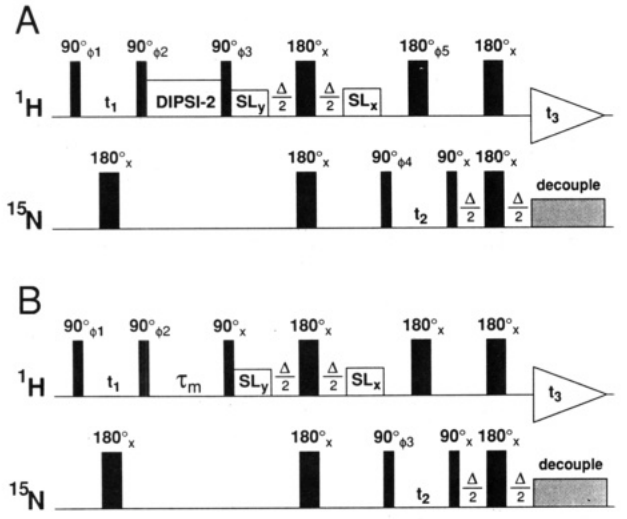

FIGURE 1: (A) Pulse scheme used for the 3D TOCSY-HMQC experiment. The phase cycling used was $\phi 1 = \phi 2 = 2(x,x,-x,-x)$, $\phi 3 = \phi 5 = 4(x),4(-x)$, $\phi 4 = x,-x$, and receiver = $2(x,-x),2(-x,x)$. Quadrature detection was achieved in F_1 by applying TPPI to $\phi 1$ and in F_2 by applying TPPI to $\phi 4$. The duration of the DIPSI-2 isotropic mixing sequence was 70 ms. (B) Pulse scheme used for the 3D NOESY-HMQC experiment. The phase cycling used was $\phi 1 = 2(x,x,-x,-x)$, $\phi 2 = 4(x),4(-x)$, $\phi 3 = x,-x$ and receiver = $x,-x,2(-x,x),x,-x$. Quadrature detection in F_1 was achieved by applying TPPI to $\phi 1$ and in F_2 by applying TPPI to $\phi 3$. The NOESY mixing time, τ_m, used was 100 ms. In both 3D experiments, water suppression was obtained by the use of spin-lock pulses, denoted by SL (Messerle et al., 1989). Spin-lock pulse lengths of SL_y = 2 ms and SL_x = 6 ms were used with a rf field strength of 8.6 kHz. The ^{15}N–^1H J^1 coupling evolution and refocusing time, $\Delta/2$, was set to 2.3 ms in both experiments. ^{15}N decoupling was carried out by using GARP-1 (Shaka et al., 1985) phase modulation. The ^{15}N carrier frequency was set at 116.6 ppm, approximately in the middle of the backbone–amide ^{15}N spectrum, and the ^1H carrier was set on the water resonance. The total acquisition times were 18.3, 19.0, and 114.7 ms in t_1, t_2, and t_3, respectively, corresponding to spectral widths of 7.0, 2.1, and 8.9 kHz in F_1, F_2, and F_3, respectively.

lock purge pulses or jump-and-return sequences were used, a low-pass filter was applied to the time domain data prior to Fourier transformation in order to improve the solvent suppression (Marion et al., 1989c). The 3D experiments were also processed on a Convex C-240 using FTNMR for the F_1–F_3 planes, and a separate routine written by Dr. M. Rance (Scripps Clinic) for the F_2 Fourier transform, to yield spectra of 256 × 64 × 1024 data points. The window function applied in the F_1 dimension was a 60°-shifted sine bell, while the F_3 dimension was multiplied by a cosine window function followed by a Lorentzian-to-Gaussian transformation. A doubly shifted sine bell window function, shifted by 60° at the beginning and by 10° at the end of the time domain interferogram, was used in the F_2 dimension (Kay et al., 1989). A simple spline baseline correction was applied to F_3 after Fourier transformation in order to remove DC baseline offsets.

NH Exchange Rates. Exchange of amide protons with solvent was monitored with a series of ^1H–^{15}N HSQC spectra. The NH solvent exchange experiment was started by addition of D_2O, at 273 K, to a sample of ^{15}N-labeled III^{glc} lyophilized from H_2O. The sample was inserted into the magnet at 308 K and thermally equilibrated for 6 min. An additional 2 min were spent shimming the magnet. The first ^1H–^{15}N HSQC experiment was therefore started 8 min after dissolution in D_2O. A series of 28 2D ^1H–^{15}N HSQC experiments were recorded over a period of ~6 days. The spectral widths used were 2.0 and 5.3 kHz in F_1 and F_2, respectively. The ^1H carrier was positioned at ~6.4 ppm such that the most upfield aliphatic peak was folded to be just upfield of the most upfield NH peak. A total of 100 t_1 increments of 2K data points were recorded. To minimize the acquisition time, dummy scans were only used prior to the first t_1 increment in each 2D experiment, and the data were stored directly in the memory buffer and not written to disk until the total acquisition was completed, i.e., all 100 t_1 increments were recorded consecutively without the need to write data to disk. A spectrum of a 0.8 mM ^{15}N labeled III^{glc} sample could be acquired in ~10 min, by using two scans per t_1 increment. A total of eight spectra were recorded within the first 2 h following dissolution in D_2O (four with two scans/t_1 increment and four with four scans/t_1 increment). Later experiments used 8 or 16 scans/t_1 increment. The resulting cross-peak volumes were measured with the FTNMR software using elliptical footprints. After scaling for the different number of scans per t_1 increment, pseudo-first-order rate constants were obtained by fitting a single-exponential decay to data for each cross peak with a Levenburg–Marquardt nonlinear least-squares algorithm (Press et al., 1986). Volumes measured in "empty" regions of the spectra were used to estimate the uncertainties in the integrated peak volumes.

RESULTS

Preliminary Homonuclear 2D Analysis. Initial analysis of the scalar connectivities in homonuclear 2D 2QF-COSY, R-COSY, 2Q and TOCSY spectra of unlabeled III^{glc} in H_2O solution enabled the identification of a number of spin systems. For instance, 13 of the 17 glycine spin systems were identified by the presence of a unique ($\omega_\alpha + \omega_{\alpha'}$ at ω_{NH}) remote peak in a 2Q spectrum and by distinctive NH/C^αH cross peaks in 2QF-COSY spectra. Spin systems were also identified for 5 of 6 alanines, 3 of 8 threonines, and 11 of 18 valines. In addition, 13 of the 37 AMX spin systems were identified, and observation of indirect couplings to C^γH protons defined 10 spin systems as being either Glu, Gln, or Met. Starting with this limited number of spin-system assignments, sequential assignments of 65% of the backbone resonances were possible by conventional NOE-based sequential assignment procedures (Billeter et al., 1982; Wüthrich, 1986). The TOCSY relayed-NOESY spectrum was found to be particularly useful in resolving ambiguities resulting from degenerate C^αH resonances in the C^αH–NH regions of the conventional NOESY spectra.

Th main problems preventing further assignment by homonuclear 2D methods were chemical shift degeneracy and resonance overlap. This is seen in Figure 2, which shows the aromatic/amide–aliphatic regions of the homonuclear 2D TOCSY (mixing time 80 ms) and 2D NOESY (mixing time 100 ms) spectra of III^{glc} obtained at pH 6.6 and 308 K. Although some of the many cross peaks observed in these spectra are well resolved, cross-peak overlap is severe in several regions. This is particularly so in the NOESY spectrum where a significant number of sequential NOE connectivities were found to be overlapped and therefore ambiguous (Figure 2B). In contrast to results reported for proteins of comparable size (Driscoll et al., 1990a), the degree of coherence transfer in the TOCSY experiment does not appear to be a limiting factor in the assignment process for III^{glc}, because a large number of relayed connectivities are observed from amide protons to C^βHs and beyond (Figure 2A).

Heteronuclear 2D Spectra. ^1H–^{15}N correlation spectra of uniformly ^{15}N-labeled III^{glc} were used to resolve ambiguities present in the ^1H 2D spectra and extend the assignments. Figure 3 shows the ^{15}N–^1H HSQC spectrum of a uniformly ^{15}N-labeled sample of III^{glc} obtained under the same conditions as the homonuclear 2D spectra in Figure 2. The assignments

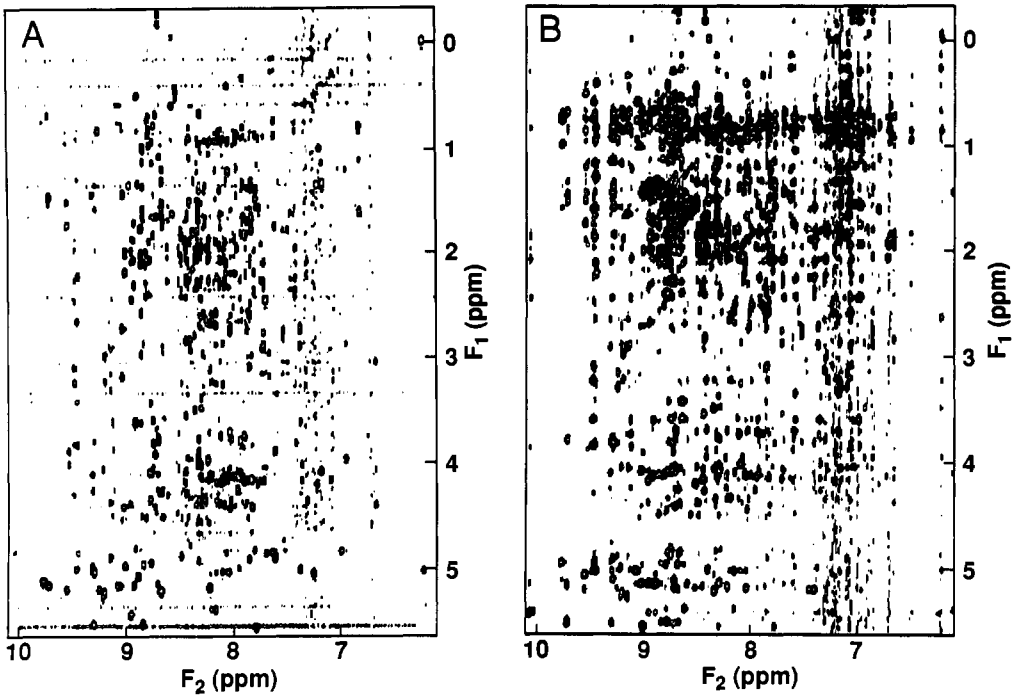

FIGURE 2: Amide–aliphatic regions of (A) ^{1}H 2D TOCSY and (B) ^{1}H 2D NOESY spectra of III^{glc} recorded at 600 MHz, pH 6.6, 308 K, and 10 mM potassium phosphate (92% H_2O/8% D_2O). The TOCSY spectrum was recorded with a DIPSI-2 mixing sequence and the sensitivity-enhanced method (Cavanagh & Rance, 1990). Both spectra were recorded with solvent presaturation and mixing times of 80 and 100 ms, respectively.

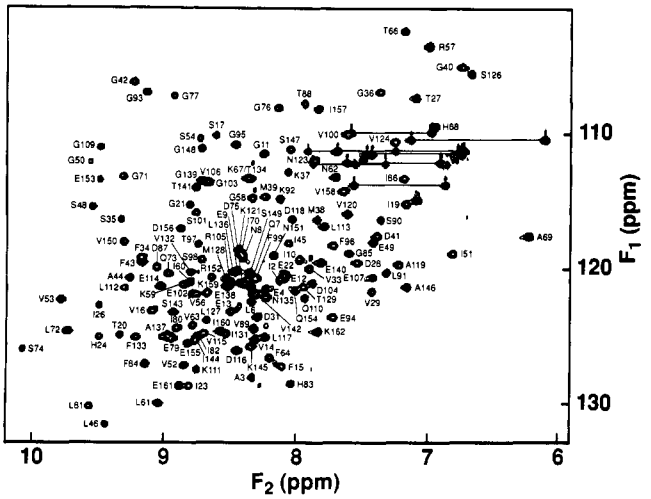

FIGURE 3: ^{1}H–^{15}N HSQC spectrum of ^{15}N-labeled III^{glc} recorded under the same conditions as the spectra in Figure 2. The assignments obtained in this study are indicated.

of the ^{1}H–^{15}N cross peaks are indicated. The spectrum is quite well resolved and contains a peak for each of the expected 150 secondary amide ^{15}NH groups (162 residues less the N-terminal methionine and 11 prolines) and 20 peaks corresponding to the $^{15}NH_2$ groups of the 10 Asn and Gln residues. The latter peaks are easily recognized by the weak resonances observed about 0.5–0.6 ppm upfield from them in the ^{15}N dimension, which originate from the ~8% semideuterated NHD moieties (Bax et al., 1990a). In cases where resonance overlap occurs, the peaks were resolved by recording a spectrum at 298 K.

The utility of the ^{1}H–^{15}N HSQC-NOESY spectrum in resolving degenerate amide protons is illustrated in Figure 4. In this example, the amide ^{1}H resonances of His 24 and Ile 26 are degenerate at 9.48 ppm, and as a result the two C$^\alpha$H–NH cross peaks observed in the ^{1}H–^{1}H NOESY spectrum (Figure 4A) cannot be unambiguously assigned. The two amide ^{15}N resonances are, however, resolved in the ^{15}N dimension of the ^{1}H–^{15}N HSQC-NOESY spectrum (Figure 4B), and the sequential C$^\alpha$H–NH connectivities were therefore assigned as indicated (Figure 4C). As illustrated in Figure 5, the ^{1}H–^{15}N HSQC–NOESY and HSQC-TOCSY experi-

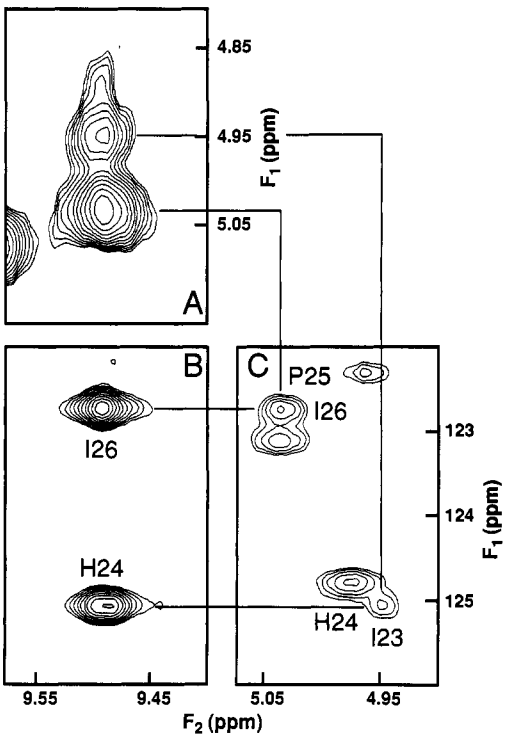

FIGURE 4: Assignment of sequential $d_{\alpha N}(i,i+1)$ NOEs involving the degenerate amide NHs of His 24 and Ile 26. (A) An expanded section of the NH–C$^\alpha$H region of the 2D ^{1}H NOESY spectrum in Figure 2B. (B) An expanded section of the NH–^{15}N region of the ^{1}H–^{15}N HSQC-NOESY spectrum recorded under the same conditions as the spectra in Figures 2 and 3, showing the autocorrelation peaks of His 24 and Ile 26. (C) An expanded section of the C$^\alpha$H–^{15}N region of the same ^{1}H–^{15}N HSQC-NOESY spectrum showing the sequential $d_{\alpha N}(i,i+1)$ NOE connectivities.

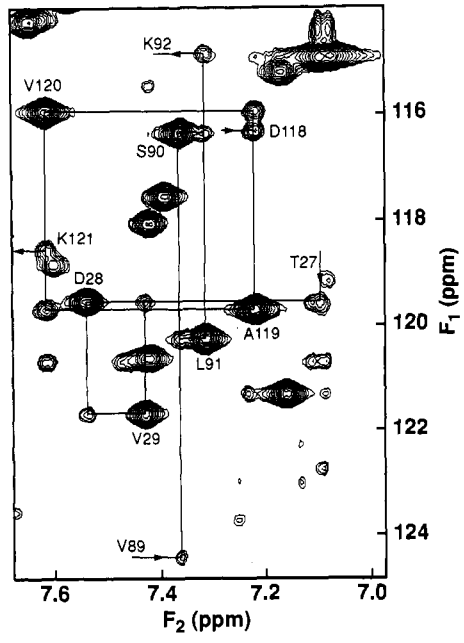

FIGURE 5: Expanded section of the amide region of the same ^1H–^{15}N HSQC-NOESY spectrum shown in Figure 4B,C. Assignments and NOE connectivities are indicated.

ments offer other advantages over the corresponding homonuclear experiments: ambiguities present because of overlapping amide and aromatic resonances are eliminated, and NH–NH NOEs can be observed between resonances with similar (although not identical) chemical shifts (such correlations would be obscured by the diagonal in homonuclear 2D spectra). The sequential $d_{NN}(i,i+1)$ NOE between residues Ser 90 and Leu 91 is unobservable in a 2D ^1H–^1H NOESY spectrum due to the presence of the intense diagonal. Additionally, all the peaks in this region of the spectrum can potentially be confused with aromatic resonances. A further advantage of the heteronuclear experiments is that assignments already made with the 2D homonuclear spectra can be confirmed without having to change the sample conditions (i.e., pH or temperature).

Heteronuclear 3D Spectra. Sequential assignment of ~90% of the ^{15}N and ^1H backbone resonances of III^{glc} was possible using the above 2D heteronuclear spectra. A disadvantage of the 2D heteronuclear experiments, however, is that the $C^\alpha H$ resonances close to the H_2O signal are difficult to observe, due to the unavoidable band of t_1 noise centered about this peak. III^{glc} has several such $C^\alpha H$ resonances, and the $C^\alpha H$–NH_{i+1} sequential NOEs involving these resonances could not be observed (even at lower temperatures in most cases). This problem is effectively eliminated by the 3D heteronuclear experiments. In the spectra acquired with the pulse sequences in Figure 1, the $F_1(^1H)$–$F_3(^1H)$ planes correspond to the $F_1(^1H)$–F_2(NH) region of conventional ^1H–^1H TOCSY or NOESY spectra but are edited with respect to the amide ^{15}N chemical shifts in F_2. The use of spin-lock purge pulses for solvent suppression, rather than solvent presaturation, ensures that $C^\alpha H$–NH connectivities will be observed close to the water resonance.

As has been noted by others (Zuiderweg & Fesik, 1989; Marion et al., 1989a,b; Fesik & Zuiderweg, 1990; Driscoll et al., 1990a), the improvement in resolution in the 3D spectra relative to the 2D spectra is dramatic. Lys 67 and Thr 134 were the only pair of residues with ^{15}N and ^1H amide chemical shifts that were insufficiently resolved to allow NOEs to be distinguished in the 3D NOESY-HMQC spectrum. In this case the sequential connectivities were obtained from a 2D ^1H–^{15}N HSQC-NOESY spectrum acquired at 298 K.

Assignment of ^1H and ^{15}N resonances from the 3D spectra was achieved by using short-range NOEs to connect spin systems, in an analogous way to the sequential assignment procedure now commonly used for 2D homonuclear spectra of smaller proteins (Billeter et al., 1982; Wüthrich, 1986). In the representation of the 3D spectra shown in Figure 6, strips taken parallel to the F_1 axis for different ^{15}N planes have been ordered according to their sequential assignment. Figure 6A,B illustrates the sequential assignment of residues 57–64, while Figure 6C,D shows the assignment of residues 109–117. The 3D TOCSY-HMQC spectrum (Figure 6A,C) was used to identify intraresidue NOEs in the NOESY-HMQC spectrum (Figure 6B,D). The strongest cross peaks observed in NOESY spectra are usually sequential $d_{\alpha N}(i,i+1)$ or $d_{NN}(i,i+1)$. These connectivities are indicated by arrows in Figure 6B,D. Given the large proportion of sequential assignments that had already been established by using the homonuclear and heteronuclear 2D spectra, confirming and extending the assignments was relatively simple. In the case of two amides being connected by a strong $d_{NN}(i,i+1)$ NOE, the assignment is unambiguous since the NOE usually appears in different ^{15}N planes of the 3D NOESY spectrum and is symmetrical about the $F_1(^1H) = F_3(^1H)$ line. This is illustrated in Figure 6B for the $d_{NN}(i,i+1)$ connectivity between Leu 61 and Asn 62 and in Figure 6D for Leu 112 and Leu 113. Ambiguity may remain, however, in the case where spin systems are connected via $d_{\alpha N}(i,i+1)$ NOEs if the $C^\alpha H$ resonance is degenerate (or nearly degenerate) with another. This problem is worse in the 3D spectra because cross peaks to $C^\alpha H$ resonances must be resolved in the F_1 dimension, where the digitial resolution is low. For III^{glc} the problem of $C^\alpha H$ degeneracy was alleviated either by the observation of $d_{\beta N}(i,i+1)$ or $d_{\gamma N}(i,i+1)$ NOEs or by using the 2D ^1H relayed-NOESY spectrum, in which a $d_{\alpha N}(i,i+1)$ NOE connectivity results in a $d_{NN}(i,i+1)$ cross peak at $F_1 = NH(i)$, $F_2 = NH(i+1)$.

A summary of the sequential NOE data obtained from the 3D NOESY-HMQC and 2D HSQC-NOESY specta is given in Figure 7. The NOEs were characterized as strong, medium, or weak on the basis of their total cross-peak volumes in the 3D NOESY-HMQC spectrum. Total volumes were obtained by adding volumes measured for a given cross peak in two or three adjacent ^{15}N planes. Table I lists the ^{15}N and ^1H assignments obtained to date. Assignment of proline residues has proven difficult due to the absence of an amide proton and overlap in the homonuclear 2D spectra and must await the application of 3D heteronuclear triple-resonance spectroscopy (Ikura et al., 1990) to $^{13}C/^{15}N$ doubly labeled III^{glc}.

Amide NH Exchange Rates. The observation of slowly exchanging amide protons together with certain repeating patterns of short-, medium-, and long-range NOE connectivities can be used to identify secondary structural elements (Wüthrich et al., 1984). We have identified 78 slowly exchanging amide NH groups in III^{glc} and have determined their exchange rate constants by fitting a single-exponential decay to the exchange data obtained for each cross peak.[2] Figure 8A shows plots of the volume integrals of the resonances in ^1H–^{15}N HSQC spectra of several slowly exchanging amide protons as a function of time following dissolution of the

[2] The data for Ile 70 and Lys 121 could not be fit to single-exponential functions due to overlap of these resonances. In these two cases double-exponential functions were used yielding rate constants of 1.20 ± 0.03 and 57.0 ± 2.7 × 10^{-5} s^{-1} respectively.

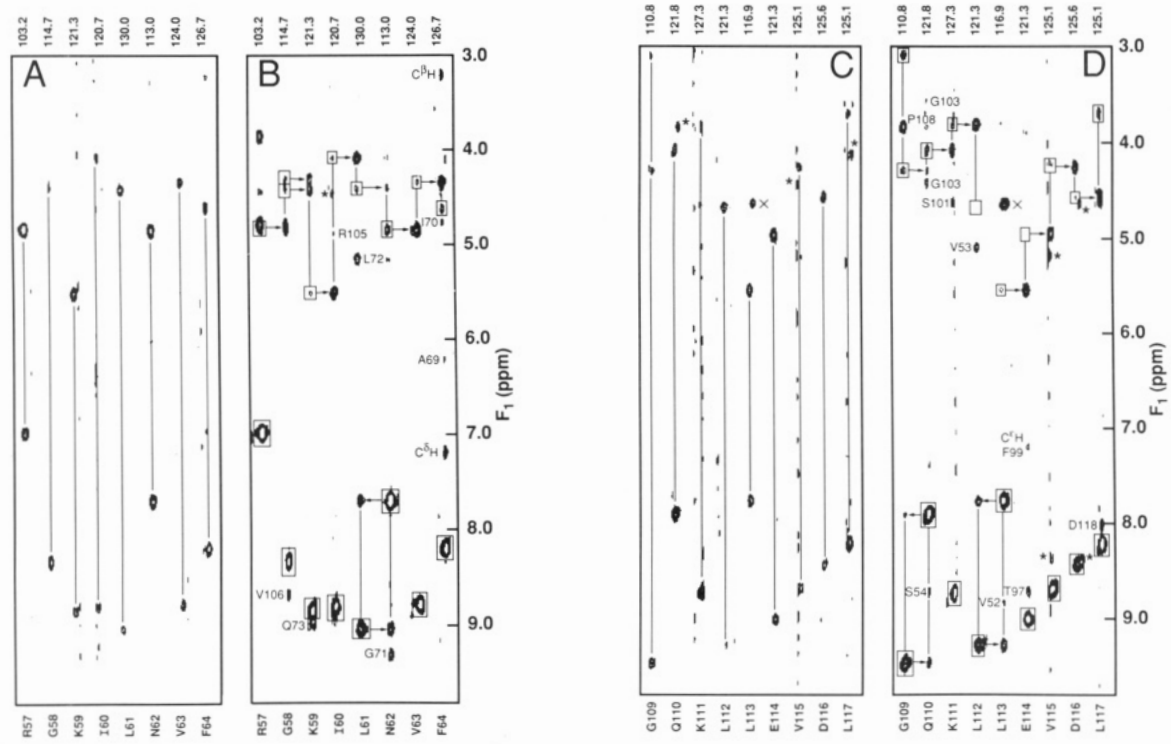

FIGURE 6: (A) Strips taken from the amide region of different ^{15}N planes of the 3D TOCSY-HMQC spectrum of III^{glc} arranged in sequential order for residues 57–64. The ^{15}N chemical shift is shown at the top, and the sequential assignment is indicated at the bottom. A line is drawn between the diagonal NH peaks and the corresponding NH–C$^{\alpha}$H cross peak. (B) The corresponding amide strips from the 3D NOESY-HMQC spectrum of III^{glc}. The diagonal NH peaks and the intraresidue NH–C$^{\alpha}$H cross peaks identified from the TOCSY experiment are boxed. Sequential $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ NOEs are indicated by arrows. Long-range NOE cross peaks are labeled. NOEs from amides other than 57–64 are marked with an asterisk. Strips from the (C) 3D TOCSY-HMQC and (D) NOESY-HMQC spectra are also shown in sequential order for residues 109–117. The slices are labelled in a similar fashion to panels A and B. The cross peaks marked with a cross correspond to solvent-exchange peaks from an arginine side-chain amide that has been folded in the F_2 (^{15}N) dimension.

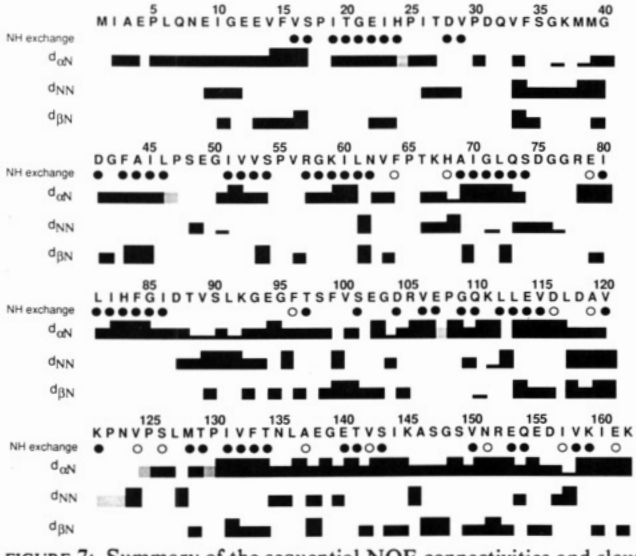

FIGURE 7: Summary of the sequential NOE connectivities and slow amide NH exchange rates observed for III^{glc}. The one-letter code for the amino acid sequence is given at the top. Characteristic $d_{\alpha N}(i,i+1)$, $d_{NN}(i,i+1)$, and $d_{\beta N}(i,i+1)$ NOE connectivities are indicated by bars between the two residues. The height of the bars gives a qualitative measure of the relative strength of the NOE in the 100-ms 3D NOESY-HMQC spectrum acquired at pH 6.6 and 308 K. Sequential NOEs involving the C$^{\delta}$Hs of proline residues were observed in a 100-ms 2D ^1H NOESY spectrum and are indicated by grey bars. Backbone amide protons with measurable solvent-exchange rate constants 10^{-3} s^{-1} < k_m < 10^{-2} s^{-1} at pH* 6.6 and 308 K are indicated by open circles. Those protons with solvent exchange rate constants <10^{-3} s^{-1} at pH* 6.6 and 308 K are indicated by solid circles.

protein in D_2O. The first-order rate constants of the slowly exchanging amide protons determined in this way are plotted against the protein sequence in Figure 8B. Those amide protons that were not detectable in the first ^1H–^{15}N HSQC spectrum recorded after dissolution in D_2O were assumed to have exchange rate constants $k_m > 10^{-2}$ s^{-1}, the upper limit for measurement by the method used. The amide protons with $k_m < 3 \times 10^{-3}$ s^{-1} are assumed to be involved in hydrogen bonds or to be buried in a solvent inaccessible region of the protein; the presence of specific hydrogen bonds can be deduced from the pattern of observed NOEs.

Secondary Structure. The observed groupings of sequential $d_{\alpha N}(i,i+1)$ NOE connectivities (Figure 7) suggests that a significant proportion of the polypeptide chain of III^{glc} is in an extended conformation. No regular helical regions could be identified from the sequential NOE data. During the sequential assignment process many long-range backbone–backbone NOE connectivities [$d_{NN}(i,j)$, $d_{\alpha N}(i,j)$, and $d_{\alpha\alpha}(i,j)$] characteristic of antiparallel β-sheet secondary structure were noted. For example, the long-range $d_{NN}(i,j)$ and $d_{\alpha N}(i,j)$ NOEs observed for the amide NH resonances of residues 57–64 are indicated in Figure 6B. The long-range $d_{\alpha\alpha}(i,j)$ NOEs observed in the C$^{\alpha}$H region of the 2D ^1H NOESY spectrum of III^{glc} in D_2O (Figure 9) also indicate that the β-strands are arranged in anti-parallel β-sheets. Acquisition of a D_2O NOESY spectrum was necessary in order to observe $d_{\alpha\alpha}(i,j)$ NOEs involving C$^{\alpha}$H resonances close to the solvent peak.

Figure 10 shows a schematic representation of the secondary structure elements of III^{glc} as determined from the long-range NOEs. The position of hydrogen bonds has been predicted on the basis of the indicated NOEs and the observation of slow amide proton exchange (Figure 8). The structure is seen to consist of a well-characterized eight-stranded β-sheet and two smaller sheets. Five β-bulges have also been identified involving residues 61–62, 99–100, 112–113, 145–146, and

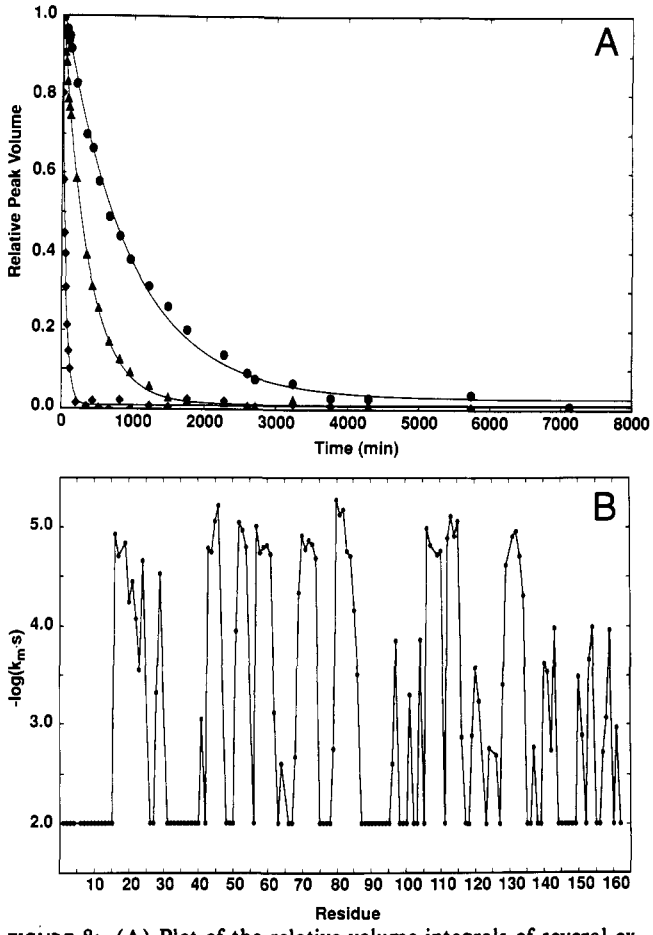

FIGURE 8: (A) Plot of the relative volume integrals of several exchanging amide protons of III^{glc} versus time following dissolution in D_2O: (●) His 83; (▲) Ala 69; (◆) Val 150. Rate constants ($k_m \times 10^{-5}$ s^{-1}) were calculated to be 1.75 ± 0.07 (His 83), 4.52 ± 0.09 (Ala 69), and 32.0 ± 2.1 (Val 150). (B) Plot of the negative logarithm of the amide proton exchange rate constants (k_m) versus the residue number of III^{glc}. Amide protons that were not detected in the first ^1H–^{15}N HSQC spectrum acquired after dissolution in D_2O have been plotted at 2.0, which represents the lower limit for k_m. Proline residues have not been plotted.

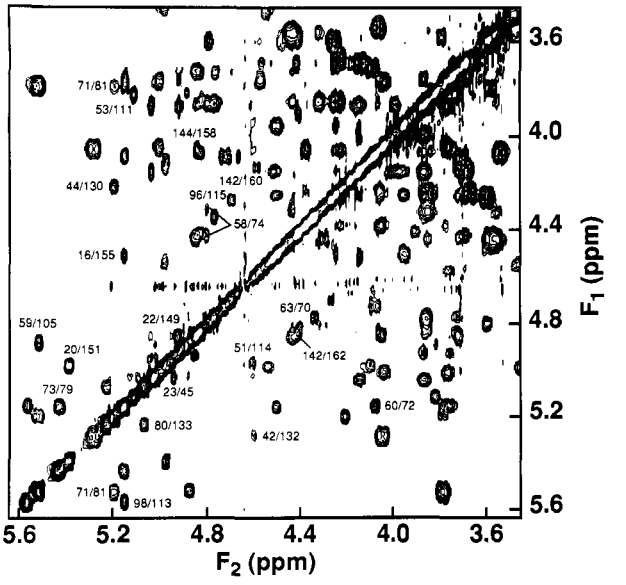

FIGURE 9: The C$^\alpha$H region of the 100-ms 2D ^1H NOESY spectrum of III^{glc} in D_2O at pH* 6.6 and 308 K. Long-range $d_{\alpha\alpha}(i,j)$ NOEs are indicated.

157–158. The β-bulges are characterized by strong $d_{NN}(i,i+1)$ NOEs, as shown in Figure 6, panels B and D for Leu 61–Asn 62 and Leu 112–Leu 113, respectively.

An important consequence of the secondary structure is that His 68 and His 83 are in close proximity to each other. Indeed, the side-chain C$^\delta$H and C$^\epsilon$H protons of His 83 have weak NOE connectivities with the C$^\delta$H of His 68 in homonuclear NOESY spectra acquired with 100-ms mixing times (data not shown).

Discussion

Nearly complete backbone ^1H and ^{15}N assignments of the 162-residue enzyme III^{glc}-like domain of *B. subtilis* II^{glc} have been determined by using a combination of 2D ^1H–^1H and ^1H–^{15}N and 3D ^1H–^{15}N NMR spectroscopy. The only backbone ^1H resonances that remain unassigned are those of Met 1, Gln 32, and the C$^\alpha$Hs of Ser 54 and 5 of the 11 prolines. The sequential assignments made are listed in Table I. This work demonstrates that it is possible to assign the NMR spectra of a protein of >17 kDa, and of unknown three-dimensional structure, by using a combination of uniform ^{15}N isotopic labeling and 2D and 3D heteronuclear NMR techniques. The use of ^1H–^{15}N spectroscopy was essential to resolve ambiguities present in the 2D ^1H–^1H spectra due to chemical shift degeneracy and overlap of amide proton resonances. Heteronuclear techniques have the advantage that peaks in the "fingerprint" region of a 2D ^1H spectrum are separated according to the ^{15}N chemical shift of the amide group, thereby improving resolution. The 3D ^1H–^{15}N spectra acquired without presaturation of the H_2O resonance were particularly advantageous, since they not only provided improved resolution of the amide proton resonances, relative to both the 2D homonuclear and heteronuclear experiments, but also allowed observation of C$^\alpha$H–NH connectivities for C$^\alpha$H resonances close to the solvent. Only two samples, one unlabeled and one uniformly ^{15}N-labeled, were used (although only the ^{15}N-labeled sample was required). This is clearly more efficient than alternative assignment procedures requiring multiple samples with amino acid specific isotope labeling and, therefore, the acquisition and analysis of a large number of spectra (Torchia et al., 1989; McIntosh et al., 1990).

The large number of sequential $d_{\alpha N}(i,i+1)$ NOE connectivities observed (Figure 7) indicates that III^{glc} is predominantly composed of β-strands connected by loops and turns. Analysis of the medium- and long-range NOEs reveals that all the strands are arranged in an antiparallel manner. The major secondary structural unit is an eight-stranded antiparallel β-sheet, shown schematically in Figure 10A. Two smaller β-sheets are illustrated in Figure 10B,C. The 13 strands are labeled with roman numerals in order of the polypeptide sequence. Strong $d_{NN}(i,i+1)$ and several weaker $d_{\alpha N}(i,i+2)$ and $d_{\alpha N}(i,i+3)$ NOE connectivities have been observed between residues 88–92 and 117–123 that may indicate the presence of an irregular "helix-like" structure or helical turns.

The assignment of the backbone ^1H and ^{15}N resonances, coupled with the ability to acquire 2D ^1H–^{15}N HSQC spectra in as little as 10 min, facilitated the determination of proton solvent exchange rate constants for the slowly exchanging amide protons (Figure 8). These data were a valuable aid in the characterization of the secondary structure present in III^{glc}. The locations of the β-strands shown in Figure 10 and the slowly exchanging amide protons (Figure 8) are strongly correlated. Additionally, rapid amide proton exchange was generally observed for those residues that lie in parts of the amino acid sequence outside the β-strands. These regions of the protein sequence are also rich in proline and glycine residues, consistent with the presence of loops and turns connecting the β-strands.

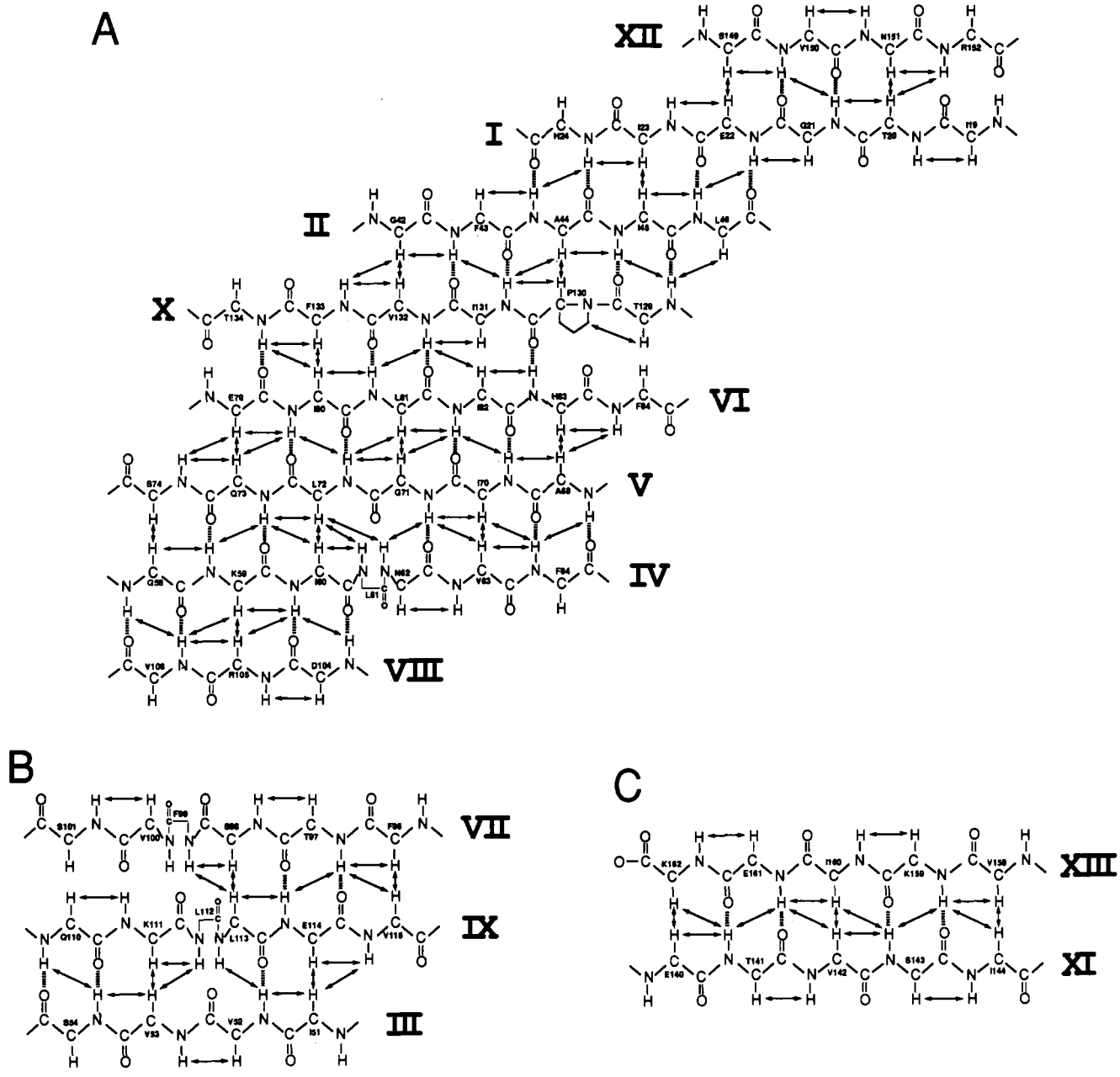

FIGURE 10: Secondary structure elements of III^{glc} as determined by analysis of NOE connectivities and amide NH exchange. Observed NOEs are indicated by arrows, and hydrogen bonds predicted from amide proton/solvent exchange and NOE data are indicated by broken lines. The β-strands are labeled with roman numerals in the order that they occur in the polypeptide sequence.

In addition to the β-sheets shown in Figure 10, a small antiparallel β-structure has been identified between residues 16–17 and 154–155, by observation of a $d_{\alpha\alpha}(i,j)$ NOE between residues Val 16 and Glu 155 (Figure 9), a $d_{\alpha N}(i,j)$ NOE between residues Glu 155 and Ser 17, and a $d_{NN}(i,j)$ NOE between Ser 17 and Gln 154. This small sheet is immediately adjacent to turns involving residues Ile 19 and Pro 18, and Arg 152 and Glu 153 [a $d_{\alpha N}(i,i+2)$ NOE connectivity is observed between Arg 152 and Gln 154]. As a consequence of this turn, the two-stranded β-sheet illustrated in Figure 10C packs against the main sheet shown in Figure 10A. Support for this conclusion comes from long-range side-chain–side-chain NOEs observed between Phe 43, Phe 133, Val 142, Val 158, and Ile 160. From a topological point of view, it is also apparent that the three-stranded β-sheet illustrated in Figure 10B must "cross over" the main sheet in order to connect strand II with strand IV, strand VI with strand VIII, and strand VIII with strand X. The conformations of the polypeptide backbone in the loop regions have not yet been defined.

The N-terminal 15 residues of III^{glc} appear to be disordered in solution. Only sequential NOEs were observed for these residues, and their amide protons could not be detected in the amide proton solvent exchange experiments (Figure 8B), indicating that their exchange rate constants are greater than $\sim 10^{-2}$ s^{-1}. In addition, the chemical shifts of the backbone C$^{\alpha}$H and side-chain ^{1}H resonances assigned to these residues are all close to the "random coil" values (Wüthrich, 1986). Flexibility of the N-terminal region of III^{glc} is to be expected, because the first 13 residues have been identified as being part of the Q-linker, which connects the C-terminal III^{glc}-like domain studied here to the membrane-bound II^{glc} domain (Sutrina et al., 1990). Such Q-linkers are typically found in proteins of prokaryotic two-component regulatory and signal-transduction systems and are usually 15–25 residues long and rich in glutamine, glutamate, arginine, serine, and proline (Wootton & Drummond, 1989). These flexible elements presumably allow the independent domains to interact during phosphoryl transfer.

Table I: ^{15}N and ^{1}H Chemical Shifts for III^{glc} at pH 6.6 and 308 K, with 10 mM Potassium Phosphate

residue	^{15}N	NH	$C^{\alpha}H$	$C^{\beta}H$	other
Met-1					
Ile-2	120.9	8.12	4.19		$C^{\gamma}H_3$ 0.93
Ala-3	128.1	8.33	4.35	1.36	
Glu-4	121.6	8.24	4.58	1.86, 2.04	
Pro-5			4.42		
Leu-6	122.5	8.34	4.36	1.60, 1.69	$C^{\gamma}H$ 1.62; $C^{\delta}H_3$ 0.90, 0.95
Gln-7	121.3	8.34	4.51	1.98, 2.05	$C^{\gamma}H$ 2.26, 2.32
Asn-8	119.5	8.41	4.68	2.83, 2.90	
Glu-9	120.1	8.45	4.28	1.96, 2.06	$C^{\gamma}H$ 2.25
Ile-10	119.3	7.97	4.25	1.95	$C^{\gamma}H_3$ 0.92; $C^{\gamma}H$ 1.18, 1.45; $C^{\delta}H_3$ 0.84
Gly-11	111.5	8.25	3.97, 3.97		
Glu-12	120.8	8.09	4.39	1.94, 2.04	$C^{\gamma}H$ 2.27, 2.30
Glu-13	123.2	8.49	4.51	2.08	
Val-14	125.3	8.31	4.03	2.03	$C^{\gamma}H_3$ 0.94, 1.02
Phe-15	127.3	8.10	5.03	2.87, 3.26	$C^{\delta}H$ 7.18
Val-16	123.2	9.09	5.16	1.61	$C^{\gamma}H_3$ 0.57, 0.75
Ser-17	110.2	8.61	4.37	3.76, 4.58	
Pro-18					
Ile-19	115.3	7.17	4.62	1.37	$C^{\gamma}H_3$ 1.13
Thr-20	125.0	9.32	4.97	4.10	$C^{\gamma}H_3$ 1.44
Gly-21	115.3	8.80	4.09, 4.72		
Glu-22	120.4	8.07	4.85	2.10, 2.18	
Ile-23	128.7	8.80	4.95	1.66	C_2H_3 0.86
His-24	125.1	9.48	5.01	2.64, 3.41	$C^{\delta}H$ 7.12, $C^{\epsilon}H$ 8.53
Pro-25			5.03		$C^{\delta}H$ 3.78, 4.04
Ile-26	122.8	9.48	3.82	1.59	$C^{\gamma}H_3$ 1.16
Thr-27	107.4	7.09	4.15	4.52	$C^{\gamma}H_3$ 1.43
Asp-28	119.5	7.53	4.74	2.74, 2.81	
Val-29	121.7	7.41	3.80	2.09	$C^{\gamma}H_3$ 0.74, 0.82
Pro-30					
Asp-31	123.5	8.29	4.69	2.58, 2.96	
Gln-32					
Val-33	120.0	7.90	3.73	1.65	$C^{\gamma}H_3$ 0.33, 0.79
Phe-34	119.2	9.18	4.36	2.65, 2.90	$C^{\delta}H$ 7.07; $C^{\epsilon}H$ 6.71; $C^{\zeta}H$ 6.51
Ser-35	116.4	9.32	4.26		
Gly-36	107.0	7.36	3.60, 4.26		
Lys-37	112.9	8.07	4.12	2.08	
Met-38	116.4	7.84	4.15		
Met-39	114.7	8.24	4.58	1.94, 2.06	
Gly-40	105.1	6.74	3.73, 4.65		
Asp-41	117.6	7.37	5.00	2.72, 2.88	
Gly-42	106.3	9.22	4.05, 5.28		
Phe-43	119.6	9.16	4.89	3.15, 3.3	$C^{\delta}H$ 7.10
Ala-44	120.8	9.26	5.20	0.81	
Ile-45	118.1	8.05	5.03	1.21	$C^{\gamma}H_3$ 0.40
Leu-46	131.6	9.44	4.85	2.22	
Pro-47					$C^{\delta}H$ 3.28, 4.43
Ser-48	115.5	9.54	4.60	3.88, 4.00	
Glu-49	118.1	7.42	4.70	1.95	
Gly-50	112.2	9.57	4.17, 4.23		
Ile-51	118.9	6.81	4.61	1.61	$C^{\gamma}H_3$ 0.78; $C^{\gamma}H$ 1.60, 1.10
Val-52	127.2	8.85	4.96	2.08	$C^{\gamma}H_3$ 0.98, 1.01
Val-53	122.4	9.78	5.12	1.58	$C^{\gamma}H_3$ 0.69, 0.73
Ser-54	110.4	8.73			
Pro-55					
Val-56	122.0	8.76	4.84	2.30	$C^{\gamma}H_3$ 0.30, 1.02
Arg-57	103.6	7.00	4.85	1.23	
Gly-58	114.8	8.33	4.31, 4.42		
Lys-59	121.2	8.85	5.52	1.58, 1.66	
Ile-60	121.0	8.79	4.08	2.39	$C^{\gamma}H_3$ 0.68
Leu-61	130.0	9.04	4.42	1.43	
Asn-62	113.2	7.70	4.85	2.44	
Val-63	124.2	8.78	4.34	2.06	$C^{\gamma}H_3$ 0.79, 0.96
Phe-64	126.6	8.20	4.60	2.62, 3.22	$C^{\delta}H$ 7.20
Pro-65					
Thr-66	102.5	7.18	4.07	4.83	$C^{\gamma}H$ 1.22
Lys-67	113.3	8.37	3.49	2.04	
His-68	109.5	6.95	3.96	1.20, 2.87	$C^{\delta}H$ 5.42; $C^{\epsilon}H$ 7.27
Ala-69	117.6	6.22	5.02	−0.02	
Ile-70	118.9	8.41	4.77	1.51	$C^{\gamma}H_3$ 0.89
Gly-71	113.2	9.31	3.79, 5.52		
Leu-72	124.7	9.73	5.16	1.46, 1.47	$C^{\gamma}H$ 1.52; $C^{\delta}H_3$ 0.67
Gln-73	120.4	8.96	5.43	1.98, 1.86	$C^{\gamma}H$ 2.07, 2.45
Ser-74	126.1	10.07	4.85	3.86, 4.78	
Asp-75	121.1	8.39	4.57	1.85, 2.15	
Gly-76	108.2	8.14	3.58, 4.44		
Gly-77	107.3	8.93	3.68, 4.15		
Arg-78	122.1	8.23	4.59		
Glu-79	125.2	8.93	5.15	1.71, 1.87	$C^{\gamma}H$ 2.07, 2.36
Ile-80	124.4	8.90	5.07	1.33	$C^{\gamma}H_3$ 0.49
Leu-81	130.3	9.56	5.20	1.73, 1.51	

Table I (Continued)

residue	^{15}N	NH	$C^{\alpha}H$	$C^{\beta}H$	other
Ile-82	125.0	8.72	4.45	1.65	$C^{\gamma}H_3$ 0.67
His-83	128.5	8.04	4.61	2.32, 2.75	$C^{\delta}H$ 6.71; $C^{\epsilon}H$ 7.87
Phe-84	127.1	9.14	4.49	2.44, 3.60	$C^{\delta}H$ 7.01
Gly-85	118.9	7.60	3.35, 3.71		
Ile-86	113.3	7.18	4.68	1.89	$C^{\gamma}H_3$ 0.99
Asp-87	119.9	9.05	4.53	2.71, 3.17	
Thr-88	107.8	7.94	3.79	4.04	$C^{\gamma}H_3$ 1.55
Val-89	124.5	8.32	3.38	1.80	$C^{\gamma}H_3$ 0.81, 0.99
Ser-90	116.4	7.36	4.16	4.01, 3.98	
Leu-91	120.3	7.31	4.35		
Lys-92	114.9	8.12	3.74		
Gly-93	107.0	9.13	3.86, 4.25		
Glu-94	123.6	7.71	4.16	2.08	$C^{\gamma}H$ 2.28
Gly-95	110.9	8.45	3.54, 4.07		
Phe-96	118.3	7.72	4.70	2.76, 2.87	$C^{\delta}H$ 6.87; $C^{\epsilon}H$ 7.07; $C^{\zeta}H$ 7.00
Thr-97	118.2	8.74	4.41	3.91	$C^{\gamma}H_3$ 0.77
Ser-98	119.3	8.71	5.15	3.76, 3.77	
Phe-99	120.8	8.30	4.56	2.42, 3.47	$C^{\delta}H$ 7.08; $C^{\epsilon}H$ 7.23
Val-100	110.1	7.61	4.90	2.28	$C^{\gamma}H_3$ 0.83, 1.05
Ser-101	115.9	8.76	4.66	3.59, 3.81	
Glu-102	121.9	8.79	3.43	1.86, 1.97	$C^{\gamma}H$ 2.13, 2.28
Gly-103	113.7	8.66	3.57, 4.45		
Asp-104	121.1	7.87	4.51	2.55, 2.73	
Arg-105	121.0	8.50	4.88	1.83, 2.01	
Val-106	113.6	8.68	5.21	1.89	$C^{\gamma}H_3$ 0.53, 0.72
Glu-107	120.7	7.42	4.78		
Pro-108			3.85		$C^{\delta}H$ 3.36, 3.72
Gly-109	111.1	9.48	3.09, 4.30		
Gln-110	122.2	7.93	4.09	1.80	$C^{\gamma}H$ 2.26
Lys-111	127.6	8.75	3.82		
Leu-112	121.5	9.30	4.68	1.58	
Leu-113	116.9	7.79	5.56	1.38, 1.48	$C^{\delta}H_3$ 0.86, 0.88
Glu-114	121.3	9.02	4.97	1.99, 1.99	$C^{\gamma}H$ 2.17
Val-115	124.9	8.69	4.27	1.07	$C^{\gamma}H_3$ −0.18, −0.27
Asp-116	126.1	8.45	4.58	2.42, 2.82	
Leu-117	125.2	8.23	3.70	1.42, 1.89	$C^{\gamma}H$ 1.74; $C^{\delta}H_3$ 0.73, 0.80
Asp-118	116.4	8.03	4.35	2.58, 2.70	
Ala-119	119.7	7.21	4.19	1.34	
Val-120	116.0	7.61	3.88	1.74	$C^{\gamma}H_3$ 0.73, 0.87
Lys-121	118.6	8.43	3.95		
Pro-122					$C^{\delta}H$ 3.22, 3.69
Asn-123	112.0	7.85	4.91	2.58, 3.08	
Val-124	110.6	7.25	5.05	2.84	$C^{\gamma}H_3$ 0.72, 0.90
Pro-125			4.54		$C^{\delta}H$ 3.87, 4.16
Ser-126	105.6	6.67	4.39	3.04, 4.06	
Leu-127	123.8	8.68	4.26	1.58, 1.52	
Met-128	120.8	8.52	3.73		
Thr-129	121.4	7.95	4.98	4.14	$C^{\gamma}H_3$ 0.88
Pro-130			4.21		$C^{\delta}H$ 3.46, 4.55
Ile-131	124.9	8.53	4.46	0.44	$C^{\gamma}H_3$ 0.42
Val-132	120.3	8.80	4.59	1.89	$C^{\gamma}H_3$ 0.71, 0.86
Phe-133	125.1	9.21	5.23	2.98, 3.22	$C^{\delta}H$ 7.30; $C^{\epsilon}H$ 6.91; $C^{\zeta}H$ 6.23
Thr-134	113.4	8.36	4.61	4.59	$C^{\gamma}H_3$ 1.19
Asn-135	122.1	8.22	5.18	2.65, 3.11	
Leu-136	120.3	8.46	4.18	1.87, 1.70	
Ala-137	125.1	8.97	4.38	1.34	
Glu-138	121.4	8.52	4.04		
Gly-139	113.5	8.73	3.68, 4.25		
Glu-140	119.6	7.81	4.83	1.65, 2.35	$C^{\gamma}H$ 2.18, 1.97
Thr-141	114.1	8.76	4.66	4.08	$C^{\gamma}H_3$ 1.12
Val-142	121.9	8.30	4.13	1.97	$C^{\gamma}H_3$ 0.66, 0.66
Ser-143	123.3	8.90	4.80	3.61, 3.61	
Ile-144	125.3	8.76	4.05	1.82	$C^{\gamma}H_3$ 0.93
Lys-145	125.8	8.34	4.35		
Ala-146	121.4	7.15	4.54	1.33	
Ser-147	111.2	8.04	4.43	3.70, 3.70	
Gly-148	111.1	8.71	3.71, 4.10		
Ser-149	120.3	8.33	4.92	3.76, 3.88	
Val-150	118.1	9.31	5.04	1.96	$C^{\gamma}H_3$ 0.81, 0.91
Asn-151	119.1	8.16	5.39	2.57, 2.68	
Arg-152	120.7	8.63	3.60		
Glu-153	113.5	9.49	3.23		
Gln-154	121.7	8.01	4.12	1.82	
Glu-155	125.6	8.82	4.51	1.99, 2.11	$C^{\gamma}H$ 2.28, 2.40
Asp-156	117.1	8.87	4.49	2.56, 2.96	
Ile-157	108.2	7.84	4.16		
Val-158	114.3	7.64	4.84	1.51	$C^{\gamma}H_3$ 0.13, 0.26
Lys-159	121.8	8.67	4.50	1.52, 1.66	
Ile-160	124.7	8.57	4.59	1.62	$C^{\gamma}H_3$ 0.58; $C^{\gamma}H$ 0.57, 0.89; $C^{\delta}H_3$ −0.07
Glu-161	128.7	8.88	4.62	1.76, 2.02	$C^{\gamma}H$ 2.10
Lys-162	124.7	7.83	4.41	1.67, 1.77	

Other regions of the protein for which defined secondary structure has not yet been detected include the long loops between β-strands I and II (Pro 30–Gly 41) and between β-strands VI and VII (Asp 87–Gly 95). All of the residues in these loop regions have amide proton solvent exchange rates that are too fast to measure by the technique employed here, suggesting that they are exposed to the solvent and therefore exchange rapidly. By contrast, regular secondary structure is maintained right to the C-terminal residue, Lys 162. The small antiparallel β-sheet from Glu 140–Ile 144 and Val 158–Lys 162 (Figure 10C) appears to be quite regular, according to the NOE connectivities, and a series of slowly exchanging amide protons imply that the sheet is firmly hydrogen bonded.

Another notable feature of the secondary structure is the identification of five β-bulges. β-bulges have the effect of changing the register of the strands in a β-sheet (Richardson et al., 1978). Consequently the observation of two pairs of β-bulges opposite each other in adjacent strands within a β-sheet is particularly interesting, i.e., the β-bulges at residues 99–100 and 112–113 in strands VII and IX, respectively (Figure 10B), and those at 145–146 and 156–157 situated at the right-hand side of strands XI and XIII, respectively, in Figure 10C. The latter pair of bulges may be considered pseudo-β-bulges since they do not exhibit β-structure on both sides but rather are situated at one end of a β-sheet. The fifth β-bulge, between residues 61 and 62 in β-strand IV (Figure 10A) appears to be accommodated by Gly 71 in β-strand V.

Apart from the high content of antiparallel β-sheet (and the absence of parallel β-sheet or regular helical structure), the most interesting result of the secondary structure analysis is the finding that His 68 and His 83 are in close proximity. His 83 has been identified as the phosphorylation site of the *B. subtilis* III^{glc}-like domain by site-directed mutagenesis. In addition, mutation of His 68 to alanine results in a protein that can accept but cannot transfer a phosphoryl group (Reizer and Saier, unpublished). Similar results have also been reported for *E. coli* III^{glc} (Presper et al., 1989). Observation of NOE connectivities between the side-chain protons of the two histidines indicates that they are within 5 Å of each other and supports the earlier suggestion that His 68, as well as His 83, may play a critical role in III^{glc} function. Further discussion of the role of His 68 must, however, await the determination of the high-resolution solution structures of both III^{glc} and phospho-III^{glc}.

The backbone ^1H and ^{15}N assignments and secondary structure analysis of III^{glc} reported in this paper represent the first step in the 3D solution structure determination of this 162-residue protein and provide a basis for studies of the protein backbone dynamics, both of which are currently under way in this laboratory. Preliminary distance geometry calculations, using data obtained from the 3D ^1H–^{15}N NOESY-HMQC spectra, show a global fold consistent with the β-sheet structures shown in Figure 10 but indicate that more extensive side-chain assignments and identification of side-chain–side-chain NOE connectivities will be required before an accurate 3D structure can be calculated. Work on assigning the aliphatic ^1H and ^{13}C resonances of III^{glc} using double- and triple-resonance heteronuclear 3D NMR techniques (Ikura et al., 1990; Kay et al., 1990; Bax et al., 1990b; Clore et al., 1990) is currently in progress.

Added in Proof

After submission of the manuscript, a paper appeared describing the secondary structure of III^{glc} from *E. coli*, determined on the basis of heteronuclear 3D NMR spectroscopy (Pelton et al., 1991). The secondary structure of *E. coli* III^{glc} is similar to that determined here for *B. subtilis* III^{glc}. Medium-range $d_{\alpha N}(i,i+2)$ and $d_{\alpha N}(i,i+3)$ NOEs analogous to those reported for *E. coli* III^{glc} have also been observed by us for *B. subtilis* III^{glc}. The comparative weakness of the $d_{\alpha N}(i,i+3)$ NOEs, however, leads us to believe that regular helix (either α-helix or 3_{10}-helix) is absent from *B. subtilis* III^{glc}.

Acknowledgments

We thank Drs. Walter Chazin and Nicholas Skelton for many helpful discussions. We are grateful to Dr. Mark Rance for useful discussions and assistance with spectrometer hardware.

References

Bax, A. (1989) *Annu. Rev. Biochem. 58*, 223–256.
Bax, A., & Subramanian, S. J. (1986) *J. Magn. Reson. 67*, 565–569.
Bax, A., Sklenar, V., Clore, G. M., & Gronenborn, A. M. (1987) *J. Am. Chem. Soc. 109*, 7188–7190.
Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., & Tschudin, R. (1990a) *J. Magn. Reson. 86*, 304–318.
Bax, A., Clore, G. M., & Gronenborn, A. M. (1990b) *J. Magn. Reson. 88*, 425–431.
Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.
Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93–106.
Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.
Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1984) *Mol. Phys. 48*, 535–560.
Cavanagh, J., & Rance, M. (1990) *J. Magn. Reson. 88*, 72–85.
Cavanagh, J., Palmer, A. G., III, Wright, P. E., & Rance, M. (1991) *J. Magn. Reson. 91*, 429–436.
Clore, G. M., & Gronenborn, A. M. (1989) *Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.
Clore, G. M., Bax, A., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 8172–8184.
Davis, D. G. (1989) *J. Magn. Reson. 81*, 603–607.
Dean, D. A., Reizer, J., Nikaido, H., & Saier, M. H., Jr. (1990) *J. Biol. Chem. 265*, 21005–21010.
de Boer, M., Broekhuizen, C. P., & Postma, P. W. (1986) *J. Bacteriol. 167*, 393–395.
Dörschug, M., Frank, R., Kalbitzer, H. R., Hengstenberg, W., & Deutscher, J. (1984) *Eur. J. Biochem. 144*, 113–119.
Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T., & Gronenborn, A. M. (1990a) *Biochemistry 29*, 3542–3556.
Driscoll, P. C., Gronenborn, A. M., Wingfield, P. T., & Clore, G. M. (1990b) *Biochemistry 29*, 4668–4682.
Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.
Fesik, S. W., & Zuiderweg, E. R. P. (1990) *Q. Rev. Biophys. 23*, 97–131.
Ikura, M., Kay, L. E., & Bax, A. (1990) *Biochemistry 29*, 4659–4667.
Kalbitzer, H. R., Deutscher, J., Hengstenberg, W., & Rösch, P. (1981) *Biochemistry 20*, 6178–6185.
Kay, L. E., Marion, D., & Bax, A. (1989) *J. Magn. Reson. 84*, 72–84.
Kay, L. E., Ikura, M., & Bax, A. (1990) *J. Am. Chem. Soc. 112*, 888–889.
Kessler, H., Gemmecker, G., & Steuernagel, S. (1988) *Angew. Chem., Int. Ed. Engl. 27*, 564–566.
Live, D. H., Davis, D. G., Agosta, W. C., & Cowburn, D. (1984) *J. Am. Chem. Soc. 106*, 1939–1941.

McIntosh, L. P., & Dahlquist, F. W. (1990) *Q. Rev. Biophys. 23*, 1–38.
McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G., & Dahlquist, F. W. (1990) *Biochemistry 29*, 6341–6362.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., & Bax, A. (1989a) *J. Am. Chem. Soc. 111*, 1515–1517.
Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989b) *Biochemistry 28*, 6150–6156.
Marion, D., Ikura, M., & Bax, A. (1989c) *J. Magn. Reson. 84*, 425–430.
Meadow, N. D., & Roseman, S. (1982) *J. Biol. Chem. 257*, 14526–14537.
Meadow, N. D., Fox, D. K., & Roseman, S. (1990) *Annu. Rev. Biochem. 59*, 497–542.
Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989) *J. Magn. Reson. 85*, 608–613.
Morris, G. A., & Freeman, R. (1979) *J. Am. Chem. Soc. 101*, 760–762.
Nelson, S. O., Wright, J. K., & Postma, P. W. (1983) *EMBO J. 2*, 715–720.
Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., & Campbell, I. D. (1990) *J. Magn. Reson. 87*, 488–501.
Novotny, M. J., Frederickson, W. L., Waygood, E. B., & Saier, M. H., Jr. (1985) *J. Bacteriol. 162*, 810–816.
Osumi, T., & Saier, M. H., Jr. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1457–1461.
Pelton et al. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3479–3483.
Presper, K. A., Wong, C.-Y., Liu, L., Meadow, N. D., & Roseman, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4052–4055.
Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1986) *Numerical Recipes*, Cambridge University Press, Cambridge.
Rance, M. (1987) *J. Magn. Reson. 74*, 557–564.
Rance, M., & Byrd, R. A. (1983) *J. Magn. Reson. 54*, 221–240.
Rance, M., & Wright, P. E. (1986) *J. Magn. Reson. 66*, 372–378.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.
Reizer, J., Saier, M. H., Jr., Deutscher, J., Grenier, F., Thompson, J., & Hengstenberg, W. (1988) *CRC Crit. Rev. Microbiol. 15*, 297–338.
Richardson, J. S., Getzoff, E. D., & Richardson, D. C. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2574–2578.
Saier, M. H., Jr. (1989) *Microbiol. Rev. 53*, 109–120.
Saier, M. H., Jr., Novotny, M. J., Comeau-Fuhrman, D., Osumi, T., & Desai, J. D. (1983) *J. Bacteriol. 155*, 1351–1357.
Shaka, A. J., Barker, P. B., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547–552.
Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274–293.
Sutrina, S. L., Reddy, P., Saier, M. H., Jr., & Reizer, J. (1990) *J. Biol. Chem. 265*, 18581–18589.
Torchia, D. A., Sparks, S. W., & Bax, A. (1989) *Biochemistry 28*, 5509–5524.
Wagner, G. (1983) *J. Magn. Reson. 55*, 151–156.
Wootton, J. C., & Drummond, M. H. (1989) *Protein Eng. 2*, 535–543.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, J. Wiley and Sons, New York.
Wüthrich, K. (1989) *Acc. Chem. Res. 22*, 36–44.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.
Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387–2391.