Solution Structure of the 30 kDa N-Terminal Domain of Enzyme I of the *Escherichia coli* Phosphoenolpyruvate:Sugar Phosphotransferase System by Multidimensional NMR†

Daniel S. Garrett,‡ Yeong-Jae Seok,§ Der-Ing Liao,||,⊥ Alan Peterkofsky,§ Angela M. Gronenborn,*,‡ and G. Marius Clore*,‡

Laboratory of Chemical Physics and Laboratory of Molecular Biology, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, and Laboratory of Biochemical Genetics, Building 36, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892

Received November 26, 1996; Revised Manuscript Received January 13, 1997⊗

ABSTRACT: The three-dimensional solution structure of the 259-residue 30 kDa N-terminal domain of enzyme I (EIN) of the phosphoenolpyruvate:sugar phosphotransferase system of *Escherichia coli* has been determined by multidimensional nuclear magnetic resonance spectroscopy. Enzyme I, which is autophosphorylated by phosphoenolpyruvate, reversibly phosphorylates the phosphocarrier protein HPr, which in turn phosphorylates a group of membrane-associated proteins, known as enzymes II. To facilitate and confirm NH, ^{15}N, and ^{13}C assignments, extensive use was made of perdeuterated ^{15}N- and ^{15}N/^{13}C-labeled protein to narrow line widths. Ninety-eight percent of the ^{1}H, ^{15}N, and ^{13}C assignments for the backbone and first side chain atoms of protonated EIN were obtained using a combination of double and triple resonance correlation experiments. The structure determination was based on a total of 4251 experimental NMR restraints, and the precision of the coordinates for the final 50 simulated annealing structures is 0.79 ± 0.18 Å for the backbone atoms and 1.06 ± 0.15 Å for all atoms. The structure is ellipsoidal in shape, approximately 78 Å long and 32 Å wide, and comprises two domains: an α/β domain (residues 1−20 and 148−230) consisting of six strands and three helices and an α-domain (residues 33−143) consisting of four helices. The two domains are connected by two linkers (residues 21−32 and 144−147), and in addition, at the C-terminus there is another helix which serves as a linker between the N- and C-terminal domains of intact enzyme I. A comparison with the recently solved X-ray structure of EIN [Liao, D.-I., Silverton, E., Seok, Y.-J., Lee, B. R., Peterkofsky, A., & Davies, D. R. (1996) *Structure 4*, 861−872] indicates that there are no significant differences between the solution and crystal structures within the errors of the coordinates. The active site His189 is located in a cleft at the junction of the α and α/β domains and has a pK_a of ∼6.3. His189 has a trans conformation about χ_1, a g^+ conformation about χ_2, and its Nϵ2 atom accepts a hydrogen bond from the hydroxyl proton of Thr168. Since His189 is thought to be phosphorylated at the Nϵ2 position, its side chain conformation would have to change upon phosphorylation.

The phosphoenolpyruvate:sugar phosphotransferase system (PTS),[1] which is found throughout the bacterial kingdom, is responsible for the coupled phosphorylation and translocation of numerous sugars across the cytoplasmic membrane [see Postma et al. (1993) and Herzberg and Klevit (1994) for reviews]. The PTS system comprises a cascade of proteins. The first step in the pathway involves the autophosphorylation of enzyme I (EI) by phosphoenolpyruvate (PEP) to generate phosphorylated EI (P-EI) and pyruvate. P-EI is then responsible for the phosphorylation of a small phosphocarrier protein known as HPr, which in turn phosphorylates a number of membrane-associated proteins, collectively known as enzymes II (EII), which effect the sugar-specific/translocation reactions. EI itself is a 64 kDa protein consisting of an N-terminal and a C-terminal domain (Licalsi et al., 1991; Lee et al., 1994). The N-terminal domain of EI (EIN) extends from residues 1 to 259 and can be phosphorylated in a fully reversible manner by phosphorylated HPr. EIN, however, cannot be autophosphorylated by PEP. X-ray and NMR structures have been determined for HPr (Klevit & Waygood, 1986; Hammen et al., 1991; Wittekind et al., 1992;

† This work was supported by the AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G.M.C. and A.M.G.). The PDB accession codes are 1EZA for the restrained regularized mean structure, 1EZB and 1EZC for the ensemble of 50 simulated annealing structures, and R1EZAMR for the experimental restraints and ^{1}H, ^{15}N, and ^{13}C resonance assignments.

* Authors to whom correspondence should be addressed (e-mail: G.M.C., clore@vger.niddk.nih.gov; A.M.G., gronenborn@vger.niddk.nih.gov).

‡ Laboratory of Chemical Physics.

§ Laboratory of Biochemical Genetics.

|| Laboratory of Molecular Biology.

⊥ Present address: DuPont Central Research and Development, Building E228, Wilmington, DE 19880-0228.

⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1997.

[1] Abbreviations: PTS, Phophoenolpyruvate:sugar phosphotranferase system of *Escherichia coli*; EI, enzyme I of the PTS; P-EI, phosphorylated form of EI; EIN, N-terminal domain (residues 1−259) of EI; EII, enzymes II of the PTS; HPr, histidine phosphocarrier protein of the PTS; PEP, phosphoenolpyruvate; PPDK, pyruvate phosphate dikinase; NMR, nuclear magnetic resonance spectroscopy; NOE, nuclear Overhauser enhancement; ROE, rotating frame Overhauser enhancement; HSQC, heteronuclear single-quantum coherence; HOHAHA, homonuclear Hartmann−Hahn; TOCSY, total correlation spectroscopy.

S0006-2960(96)02924-8

Herzberg et al., 1992; Jia et al., 1993; van Nuland et al., 1994, 1995) and EIIA^{glc} (Liao et al., 1991; Worthylake et al., 1991; Fairbrother et al., 1992). Intact EI is too large for NMR and, to date, has proved resistant to crystallization. In this paper, we present the determination of the solution structure of EIN by multidimensional NMR. While this work was in progress, and after we had completed the resonance assignments and elucidation of the secondary structure, the crystal structure of EIN became available (Liao et al., 1996).

Recent progress in NMR methodology, in particular the use of perdeuteration (Torchia et al., 1988; Grzesiek et al., 1993, 1995; Yamazaki et al., 1994; Venters et al., 1995; Farmer & Venters, 1995; Metzler et al., 1996), has permitted backbone assignments to be obtained for a system as large as 64 kDa, namely, the complex of two tandem dimers of Trp repressor (108 residues per subunit) with a 22 base pair oligonucleotide (Shan et al., 1996). Backbone assignments using conventional triple resonance experiments on $^{15}N/^{13}C$-labeled samples have also been obtained for single chain monomeric proteins up to 269 residues, in particular on two ~27 kDa members of the subtilisin family of proteases (Fogh et al., 1994; Remerowski et al., 1994). No three-dimensional solution NMR structure, however, of a protein much larger than about 200 residues has yet been published, and thus EIN, at 259 residues in length, is clearly among the largest monomeric proteins whose three-dimensional structure is being solved at the present time by NMR.

EXPERIMENTAL PROCEDURES

Sample Preparation. EIN was expressed from the plasmid pLP2 in *Escherichia coli* strain GI698, as described previously (Seok et al., 1996). For the expression of EIN at natural isotopic abundance, the same medium was used as described previously (Seok et al., 1996). For the expression of isotopically labeled EIN, a modified M9 minimal medium was employed. To make [U-$^{15}N/^{13}C$]EIN containing Phe and Tyr at natural ($^{12}C/^{14}N$) isotope abundance, 20 μg/mL L-phenylalanine and 20 μg/mL L-tyrosine were added to the minimal medium. The cells were grown until mid-log phase ($A_{600} = 0.5$), and protein expression was induced as described previously (Seok et al., 1996), except that the cells were incubated for 2 days after the addition of tryptophan when growth was carried out in D_2O. Protein purification was carried out as described previously (Seok et al., 1996). Samples for NMR contained 1.4−1.5 mM EIN in 40 mM sodium phosphate buffer, pH 7−8. The following samples were used for NMR: EIN at natural isotopic abundance ($^1H/^{14}N/^{12}C$) in 90% H_2O/10% D_2O and in 99.996% D_2O; [U-^{15}N]EIN in 90% H_2O/10% D_2O; [U-$^{15}N/^{13}C$]EIN in 90% H_2O/10% D_2O and in 99.996% D_2O; [U-$^2H/^{15}N$]EIN in 90% H_2O/10% D_2O; [U-$^2H/^{15}N/^{13}C$]EIN in 90% H_2O/10% D_2O; and [U-(^{12}C-Tyr/Phe)/$^{13}C/^{15}N$]EIN in 99.996% D_2O. ^{15}N and ^{13}C labeling was >99%, while ^2H labeling was >90%. (Note that, in the perdeuterated samples, only the nonexchangeable protons are deuterated.)

NMR Spectroscopy. All NMR experiments were carried out at 40 °C on Bruker AMX600 and AMX500 spectrometers equipped with *z*-shielded gradient triple resonance probes. Spectra were processed with the NmrPipe package (Delaglio et al., 1995) and analyzed using the programs PIPP, CAPP, and STAPP (Garrett et al., 1991). The sequential assignment of the ^1H, ^{13}C, and ^{15}N chemical shifts was achieved by means of through-bond heteronuclear correlations along the backbone and side chain; $^3J_{HN\alpha}$, $^3J_{C'C\gamma(aromatic)}$, and $^3J_{NC\gamma(aromatic)}$ couplings were obtained by quantitative *J* correlation spectroscopy; and interproton distances were derived from NOE experiments. Details of most of the experiments, together with the original references, are provided in the following reviews: Bax and Grzesiek (1993), Bax et al. (1994), Bax (1994), and Clore and Gronenborn (1991, 1994). A summary of the through-bond and through-space (NOE) correlation experiments employed is provided in Tables 1 and 2, respectively.

Structure Calculations. Approximate interproton distance restraints were derived from the multidimensional NOE spectra, essentially as described previously (Clore & Gronenborn, 1991). NOEs were grouped into four distance ranges, 1.8−2.7 Å (1.8−2.9 Å for NOEs involving NH protons), 1.8−3.3 Å (1.8−3.5 Å for NOEs involving NH protons), 1.8−5.0 Å, and 1.8−6.0 Å, corresponding to strong, medium, weak, and very weak NOEs. Distances involving methyl groups, aromatic ring protons, and non-stereospecifically assigned methylene protons were represented as a $(\sum r^{-6})^{-1/6}$ sum (Nilges, 1993). Protein backbone hydrogen-bonding restraints (two per hydrogen bond, $r_{NH-O} = 1.5-2.8$ Å, $r_{N-O} = 2.4-3.5$ Å) within areas of regular secondary structure were introduced during the final stages of refinement using standard NMR criteria (Wüthrich, 1986; Clore & Gronenborn, 1989) based on backbone NOEs and $^3J_{HN\alpha}$ coupling constants, supplemented by secondary ^{13}C shifts (Spera & Bax, 1991). ϕ, ψ, χ_1, and χ_2 torsion angle restraints were derived from the NOE/ROE and coupling constant data, and the minimum ranges employed were ±15°, ±40°, ±20°, and ±30°, respectively (Nilges et al., 1990; Powers et al., 1993). χ_1 angles for aromatic residues were derived from $^3J_{C'C\gamma(aromatic)}$ and $^3J_{NC\gamma(aromatic)}$ coupling constants (Hu et al., 1996). The structures were calculated using the program X-PLOR 3.1 (Brünger, 1993), adapted to incorporate pseudopotentials for $^3J_{HN\alpha}$ coupling constant (Garrett et al., 1994) and secondary $^{13}C\alpha/^{13}C\beta$ chemical shift (Kuszewski et al., 1995) restraints, and a conformational database potential (Kuszewski et al., 1996, 1997). Two different protocols were employed with identical end results: a modified version of the hybrid distance geometry−simulated annealing protocol (Nilges et al., 1988a) and a torsion angle dynamics protocol starting from random initial coordinates (Rice & Brunger, 1994; Stein et al., 1996), followed by conventional simulated annealing in Cartesian coordinate space (Nilges et al., 1988a,b). The target function that is minimized during simulated annealing and restrained regularization comprises only quadratic harmonic potential terms for covalent geometry, $^3J_{HN\alpha}$ coupling constant and secondary $^{13}C\alpha$ and $^{13}C\beta$ chemical shift restraints, square-well quadratic potentials for the experimental distance and torsion angle restraints, a quartic van der Waals repulsion term for the nonbonded contacts, and a conformational database potential. There were *no* hydrogen-bonding, electrostatic, or 6-12 Lennard-Jones empirical potential energy terms in the target function. As in previous structure determinations from this laboratory (Clore & Gronenborn, 1991), an iterative refinement strategy was employed, incorporating more experimental restraints at each successive stage as the quality of the structures improves. This includes taking care that predicted NOEs corresponding to short interproton distances are present in the spectra. In the context of iterative refinement, the use of the confor-

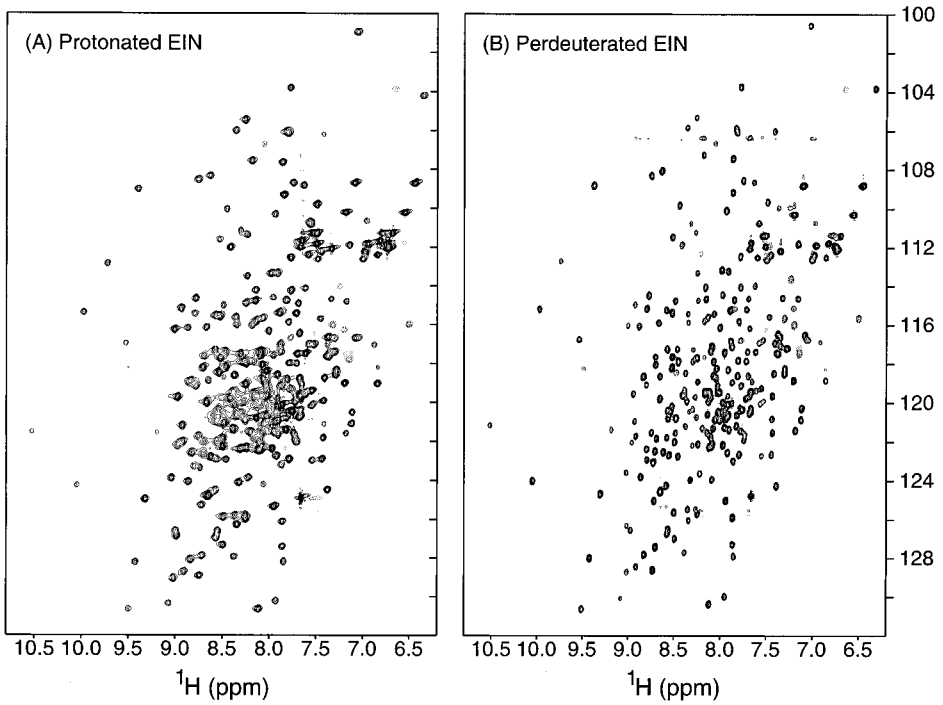

FIGURE 1: Comparison of the ^{1}H−^{15}N HSQC spectrum of protonated and perdeuterated ^{15}N-labeled EIN.

mational database potential was particularly valuable for defining rotamers of internal side chains. There are several reasons for this. First, the use of the conformational database ensures that the side chain torsion angles lie in the vicinity of their rotamers. Hence, the ensemble of structures can be readily examined to identify which rotamer, if several are possible for a given side chain, is consistent with nonsequential interresidue NOE data. Second, with the exception of Ser and Thr, the $\chi_1 = +60°$ rotamer is very rarely populated for much of ϕ,ψ space, and consequently, the relative intensities of the intraresidue NH(i)−Cβ_aH(i) and NH(i)−Cβ_bH(i) in the 3D ^{15}N-separated ROE spectrum (which is essentially free of spin diffusion effects) can be used to distinguish between the −60° and 180° rotamers (Wagner et al., 1986). Naturally, this procedure must be used with caution, and care must be taken to ensure that the chosen rotamer is fully consistent with the interresidue NOE data, while the alternate rotamers are inconsistent with these data. In all cases, the side chain rotamers deduced in this manner were consistent with those observed in the recently determined crystal structure of EIN (Liao et al., 1996). The same procedure in conjuction with the conformational database can also be used to define χ_2 rotamers of many internal side chains. For example, the conformational database potential (which is based on analysis of 1.7 Å or better resolution X-ray structures) reveals that in the case of Leu a χ_1 of 180° implies a χ_2 of +60°, a χ_1 of −60° implies a χ_2 of 180°, and a χ_1 of +60° is essentially unpopulated (Kuszewski et al., 1997). For Ile, Met, Arg, and Lys, a χ_1 of −60° implies a χ_2 of 180° or −60°, a χ_1 of 180° implies a χ_2 of 180° or +60°, and a χ_1 of +60° implies a χ_2 of 180° (Kuszewski et al., 1997).

The coordinates of the final 50 simulated annealing structures (accession codes 1EZB and 1EZC), together with the coordinates of the restrained regularized mean structure, $\overline{(SA)}$r (accession code 1EZA), and the complete list of experimental NMR restraints and ^{1}H, ^{15}N, ^{13}C assignments (accession code R1YZAMR) have been deposited in the Brookhaven Protein Data Bank.

RESULTS AND DISCUSSION

Resonance Assignment. EIN presents a difficult problem for NMR. First, the spectral dispersion is limited on account of the large proportion of helix in the structure, adding to the inherent complexity of the spectrum due to the length of the polypeptide chain (259 residues). Second, the line widths are relatively large due to the high molecular mass (∼30 kDa) and unusual shape (see below) of EIN. In the absence of specific isotope labeling, little can be done about the spectral complexity. The line widths, however, can be significantly narrowed by perdeuteration (Torchia et al., 1988; Grzesiek et al., 1993, 1995; Yamazaki et al., 1994; Venters et al., 1995; Farmer & Venters, 1995; Metzler et al., 1996; Shan et al., 1996). Thus, the average T_2 for the backbone amide protons is increased about 2-fold, from ∼13 ms for U-^{1}H/^{15}N-labeled EIN to ∼28 ms for [U-^{2}H/^{15}N]EIN. This improvement is readily apparent from a comparison of the 2D ^{1}H−^{15}N HSQC spectra of [U-^{1}H/^{15}N]- and [U-^{2}H/^{15}N]EIN (Figure 1). As a result, it is possible to obtain high-quality triple resonance spectra on [U-^{2}H/^{15}N/^{13}C]EIN. A comparison of the CBCA(CO)NH and d-CBCA(CO)NH, CBCANH and d-HNCACB, and C(CO)NH and d-C(CO)NH experiments recorded on U-^{1}H/^{15}N/^{13}C]- and [U-^{2}H/^{15}N/^{13}C]EIN is provided in Figure 2. These examples demonstrate that perdeuteration results in a significant improvement in spectral quality. A dramatic example is provided by the Val203 and G204 strips in the d-C(CO)NH/C(CO)NH experiments in which the complete carbon spin system of the preceding residues, Ile202 and Val203, respectively, can be delineated for [U-^{2}H/^{15}N/^{13}C]EIN but is completely absent for [U-^{1}H/^{15}N/^{13}C]EIN. Nevertheless, for many stretches of the sequence, we were still able to obtain a considerable number of high-quality sequential through-bond correlations for the U-^{1}H/^{15}N/^{13}C sample (Figure 3). Overall, analysis of the d-CBCA(CO)NH, d-HNCACB, and d-C(CO)NH

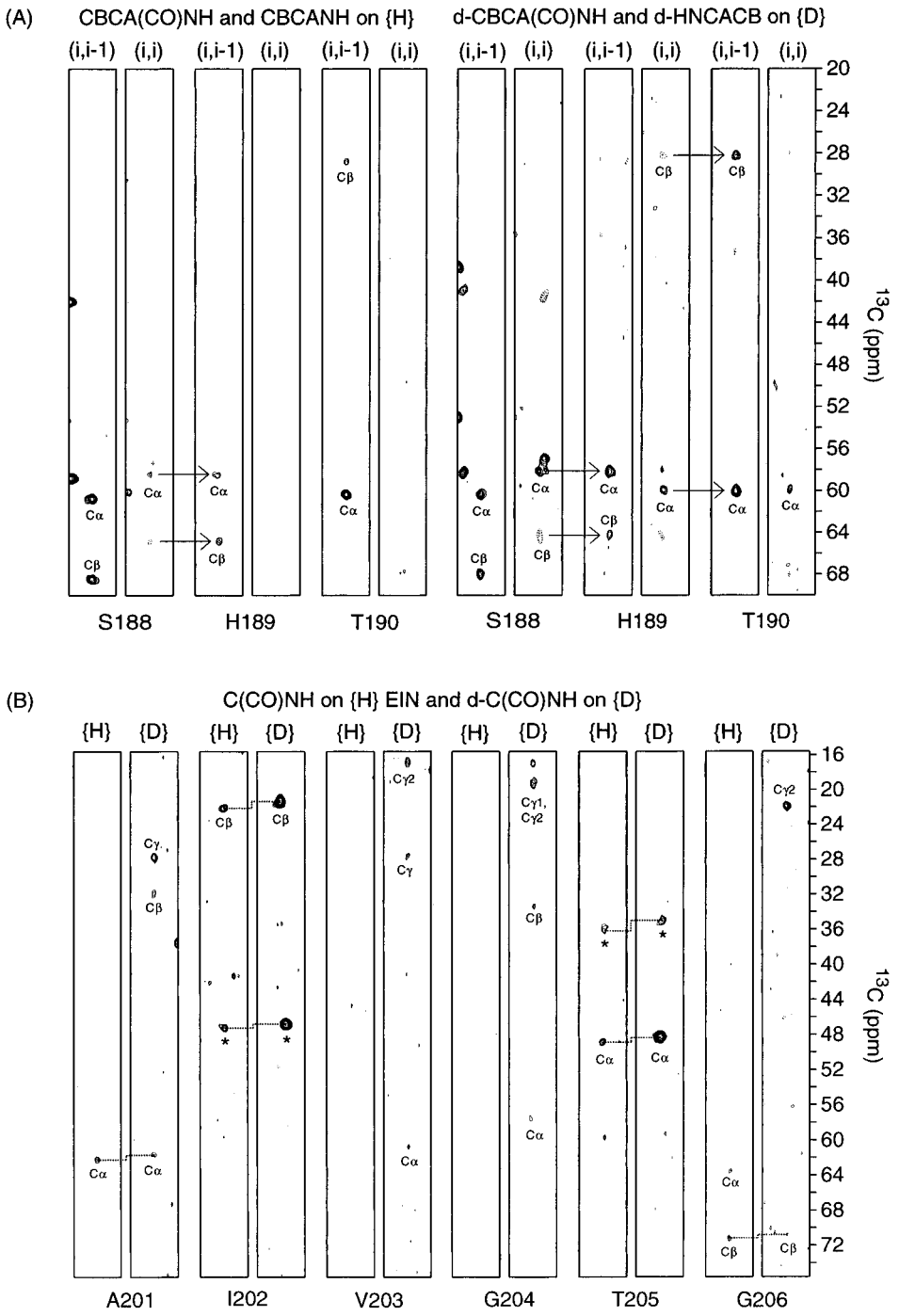

FIGURE 2: Selected strips taken from several 3D triple resonance experiments comparing the results on uniformly protonated and perdeuterated [U-^{15}N/^{13}C]EIN. (A) CBCANH and CBCA(CO)NH experiments on protonated EIN versus the d-HNCACB and d-CBCA(CO)NH experiments on perdeuterated EIN. (B) C(CO)NH and d-C(CO)NH experiments on protonated and perdeuterated EIN, respectively (peaks labeled with an asterisk in B arise from resonances that have their maximal intensities on an adjacent slice). The CBCANH and d-HNCACB experiments correlate the Cα/Cβ resonances of both the (i − 1) and i residues with the ^{15}N−^{1}H resonances of residue i. The CBCA(CO)NH and d-CBCA(CO)NH experiments correlate only the Cα/Cβ resonances of the (i − 1) residue with the ^{15}N−^{1}H resonances of residue i, while theC(CO)NH and d-C(CO)NH experiments correlate the side chain and Cα resonances of residue (i − 1) with the ^{15}N−^{1}H resonances of residue i.

experiments was very helpful in confirming assignments previously determined from the protonated samples and in some cases correcting assignments. Indeed, using the perdeuterated samples permitted the backbone resonances of an additional eight residues, including the active site His189, to be obtained. Using the double and triple resonance experiments summarized in Table 1, we were able to obtain 98% of the ^{1}H, ^{15}N, and ^{13}C assignments for the backbone (N, NH, Cα, CαH) and first adjacent side chain (Cβ, CβH) atoms of protonated EIN.

While beneficial for the purposes of line-width reduction, replacement of protons attached to carbon by deuterons does result in significant isotope shifts for both the ^{13}C and ^{15}N resonances. The deuterium isotope shift for ^{15}N, which arises from the two-bond effect involving the proton or deuteron attached to the Cα carbon, is relatively independent of amino acid type and has an average value of 0.29 ± 0.08 ppm. The deuterium isotope shift for the ^{13}Cα carbons also varies little with amino acid type but, as expected, is somewhat larger, with an average value of 0.48 ± 0.07 ppm. No

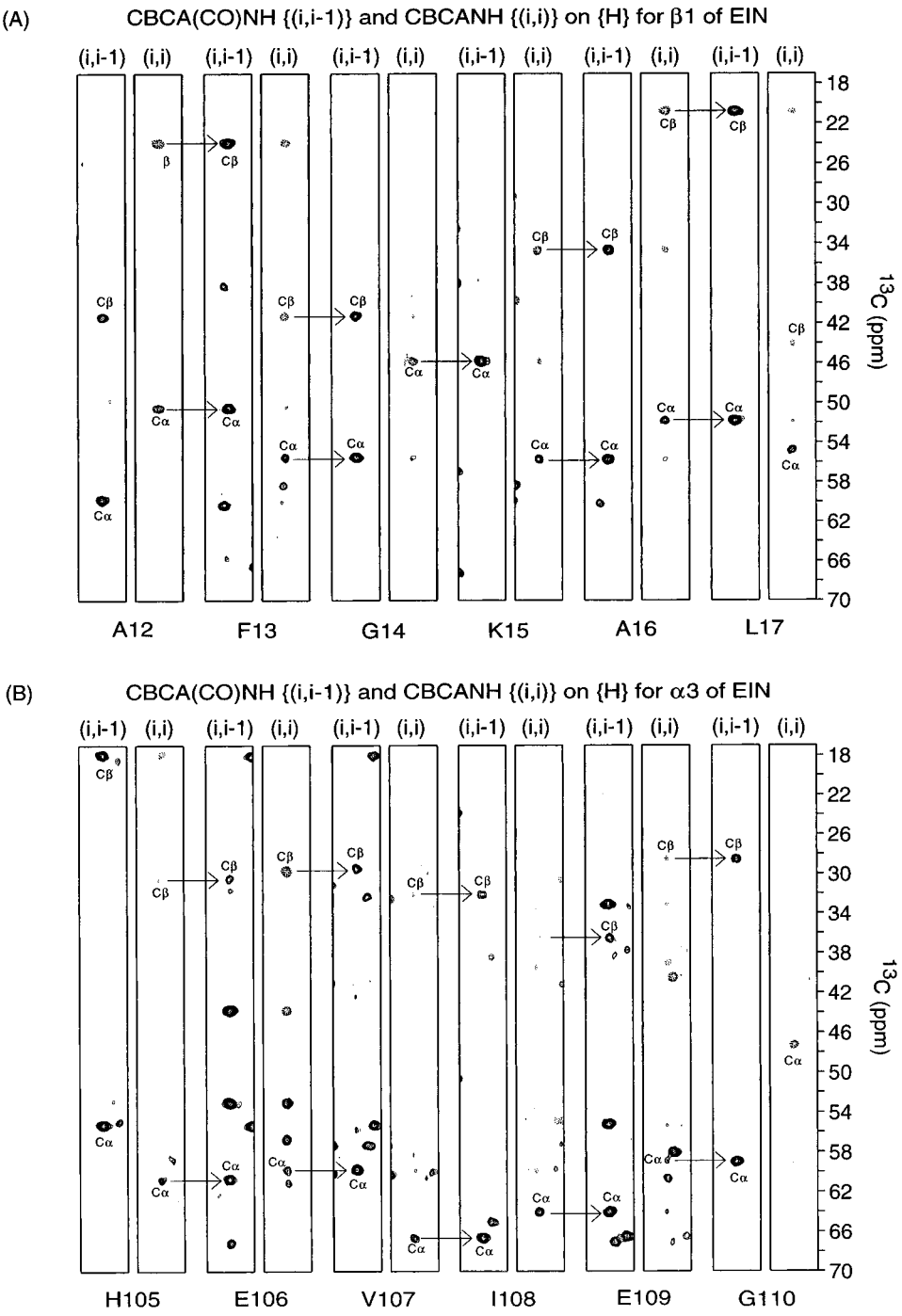

FIGURE 3: Strips taken from (A) strand β1 (residues 12−17) and (B) helix α3 (residues 105−110) for the CBCANH and CBCA(CO)NH experiments recorded on [U-^{1}H/^{15}N/^{13}C]EIN.

significant effect of secondary structure on the deuterium isotope shift of the ^{15}N resonances could be detected (0.29 ± 0.07 and 0.32 ± 0.07 ppm for α-helix and β-sheet, respectively). For the ^{13}Cα resonances, on the other hand, there is a small effect of secondary structure on the deuterium isotope effect (0.50 ± 0.08 and 0.44 ± 0.08 ppm for α-helix and β-sheet, respectively). In contrast, large variations in isotope shifts are seen for the side chains. For example, the average deuterium isotope shifts for the ^{13}Cβ atoms range from ∼0.46 ppm for Ser to ∼1.1 ppm for Leu. Similar observations relating to deuterium isotope shifts have recently been made by Venters et al. (1996). Since there is virtually no deuterium isotope shift for the backbone amide protons, the transfer of assignments determined from the perdeuterated spectra to the protonated spectra is relatively straightforward.

A summary of the observed sequential and medium-range backbone NOEs, together with the $^{3}J_{HN\alpha}$ coupling constants, and secondary carbon chemical shifts is presented in Figure 4. From these data, as well as a qualitative interpretation of long-range backbone−backbone NOEs, it is possible to delineate the secondary structure of EIN, which is also summarized in this figure.

Structure Determination. The structure calculations, using simulated annealing, were based on a total of 4251 experimental NMR restraints. The various NOE and ROE spectra used to obtain approximate interproton distance restraints are summarized in Table 1. An example of the quality of the NOE data is provided by several selected planes taken from the 4D ^{15}N/^{15}N-separated NOE spectrum recorded on [U-^{2}H/^{15}N]EIN shown in Figure 5. The latter spectrum permitted

Table 1: Summary of NMR Experiments Recorded on EIN[a]

experiment	correlation
(A) Through-Bond Correlation Experiments	
[U-^{1}H/^{15}N]EIN	
2D ^{1}H−^{15}N HSQC	^{15}N(i)−NH(i)
3D ^{15}N-separated HOHAHA (34 ms mixing time)	H(i)−^{15}N(i)−NH(i)
3D HNHA[b]	Cα H(i)−^{15}N(i)−NH(i)
[U-^{1}H/^{15}N/^{13}C]EIN	
2D ^{1}H−^{13}C HSQC	^{13}C$_j$−H$_j$
3D HNCO	^{13}CO($i-1$)−^{15}N(i)−NH(i)
3D HNCA	^{13}Cα(i,$i-1$)−^{15}N(i)−NH(i)
3D CBCANH	^{13}Cα/^{13}Cβ(i,$i-1$)−^{15}N(i)−NH(i)
3D HBHANH	CαH/CβH(i,$i-1$)−^{15}N(i)−NH(i)
3D CBCA(CO)NH	^{13}Cα/^{13}Cβ($i-1$)−^{15}N(i)−NH(i)
3D HBHA(CBCACO)NH	CαH/CβH($i-1$)−^{15}N(i)−NH(i)
3D C(CO)NH	^{13}C$_j$($i-1$)−^{15}N(i)−NH(i)
3D HCCH-COSY	H$_j$−^{13}C$_j$−^{13}C$_{j\pm1}$−H$_{j\pm1}$
3D HCCH-TOCSY (21 ms mixing time)	H$_j$−^{13}C$_j$···^{13}C$_{j\pm n}$−H$_{j\pm n}$
2D ^{13}C′−{^{13}Cγ} spin−echo difference ^{1}H−^{15}N HSQC[b]	^{15}N(i)−NH(i) (residue $i+1$ from aromatic)
2D ^{15}N−{^{13}Cγ} spin−echo difference ^{1}H−^{15}N HSQC[b]	^{15}N(i)−NH(i) (aromatic)
[U-^{2}H/^{15}N]EIN[c]	
2D ^{1}H−^{15}N HSQC	^{15}N(i)−NH(i)
[U-^{2}H/^{15}N/^{13}C]EIN[c,d]	
3D d-HNCACB	^{13}Cα/^{13}Cβ(i,$i-1$)−^{15}N(i)−NH(i)
3D d-CBCA(CO)NH	^{13}Cα/^{13}Cβ($i-1$)−^{15}N(i)−NH(i)
3D d-C(CO)NH	^{13}C$_j$($i-1$)−^{15}N(i)−NH(i)
[U-^{12}C(Tyr,Phe)/^{15}N/^{13}C]EIN[e]	
^{12}C-filtered HOHAHA	H$_j$(^{12}C)···H$_{j\pm n}$(^{12}C)
(B) Through-Space Correlation Experiments	
[U-^{15}N]EIN	
3D ^{15}N-separated NOE (100 ms mixing time)	H- - -NH(^{15}N)
3D ^{15}N-separated ROE (30 ms mixing time)	H- - -NH(^{15}N)
[U-^{15}N/^{13}C]EIN	
3D ^{13}C-separated NOE (75 ms mixing time)	H- - -H(^{13}C)
4D ^{13}C/^{13}C-separated NOE (75 ms mixing time)	H(^{13}C)- - -H(^{13}C)
4D ^{13}C/^{15}N-separated NOE (75 ms mixing time)	H(^{13}C)- - -H(^{15}N)
[U-^{2}H/^{15}N/^{13}C]EIN[c]	
4D ^{15}N/^{15}N-separated NOE[f] (170 ms mixing time)	H(^{15}N)- - -H(^{15}N)
[U-^{12}C(Tyr,Phe)/^{15}N/^{13}C]EIN[e]	
2D ^{12}C-filtered NOE (75 ms mixing time)	H(^{12}C)- - -H(^{12}C)
3D ^{13}C-separated/^{12}C-filtered NOE (75 ms mixing time)	H(^{13}C)- - -H(^{12}C)

[a] All experiments, except where noted, are described in the following reviews (which also provide the original references): Bax and Grzesiek (1993); Bax et al. (1994); Clore and Gronenborn (1991, 1994). [b] The 3D HNHA, 2D ^{13}C′−{^{13}Cγ} spin−echo difference ^{1}H−^{15}N HSQC, and 2D ^{15}N−{^{13}Cγ} spin−echo difference ^{1}H−^{15}N HSQC spectra represent quantitative J correlation spectra which permit one to determine $^{3}J_{\mathrm{HN}\alpha}$, $^{3}J_{\mathrm{C'C}\gamma\mathrm{(aromatic)}}$ and $^{3}J_{\mathrm{NC}\gamma\mathrm{(aromatic)}}$ coupling constants, respectively (Bax et al., 1994; Hu et al., 1996). [c] Only the carbon-attached protons are deuterated. [d] The d-HNCACB experiment is the same as the HNCACB experiment described by Wittekind and Mueller (1993) except that ^{2}H decoupling is employed throughout. The d-CBCA(CO)NH and d-C(CO)NH experiments are the same as the CBCA(CO)NH and C(CO)NH experiments, except that the experiment starts off on ^{13}C atoms instead of ^{13}C-attached protons (i.e., the first INEPT step transferring magnetization from protons to carbon atoms is omitted) and deuterium decoupling is employed throughout. [e] The carbon atoms of Tyr and Phe are at natural isotopic abundance; carbon atoms of all other residues are ^{13}C-labeled. [f] This experiment is similar to the ^{13}C/^{13}C-separated NOE experiment except that it involves a semiconstant evolution period for NH protons and selective water flip-back pulses to preserve water magnetization along the z axis (Grzesiek et al., 1995).

the unambiguous assignment of several long-range NOEs which rapidly allowed the β-sheets to be aligned. In fact, strands β2, β3, and β4 could not be aligned prior to analyzing the 4D ^{15}N/^{15}N-separated NOE spectrum, which also provided unique assignments for several short- and medium-range NOEs which better defined the α-helices. Another feature of the 4D ^{15}N/^{15}N-separated NOE experiment is that the huge reduction in proton density obtained by using a U-^{2}H/^{15}N sample permits one to employ long mixing times without significant spin diffusion and hence provides the possibility of detecting interproton distance contacts greater than 6 Å between backbone amide protons (Grzesiek et al., 1995). In the case of EIN we were able to observe 10 very weak NOEs in the 4D ^{15}N/^{15}N-separated NOE which were left out of the structure calculations and were found to correspond to NH−NH distances in the 6.5−7.5 Å range in both the ensemble of calculated simulated annealing structures and the X-ray structure. The NOE spectra recorded on the [U-^{12}C(Tyr,Phe)/^{15}N/^{13}C]EIN sample were also very helpful for the assignment of intra- and interresidue NOEs involving aromatic residues.

The availability of the crystal structure (Liao et al., 1996), subsequent to the elucidation of the secondary structure from a qualitative interpretation of the NOE, coupling constant, and secondary carbon shift data, significantly expedited the determination of an initial global fold by helping us to rapidly pinpoint some inevitable errors in the initial NOE restraints file. Once the fold was established, however, knowledge of the X-ray structure was not used to speed up the iterative refinement process (which took about 9 months) and hence did not impact the final outcome in terms of precision and accuracy of the coordinates. A summary of the structural statistics and distribution of NOEs is provided in Table 2 and Figure 6, respectively, and a best fit superposition of the final 50 simulated annealing structures is shown in Figure 7A. The precision of the coordinates (that is, the atomic RMS distribution of the 50 simulated annealing structures about their mean coordinate positions) for residues 1−246

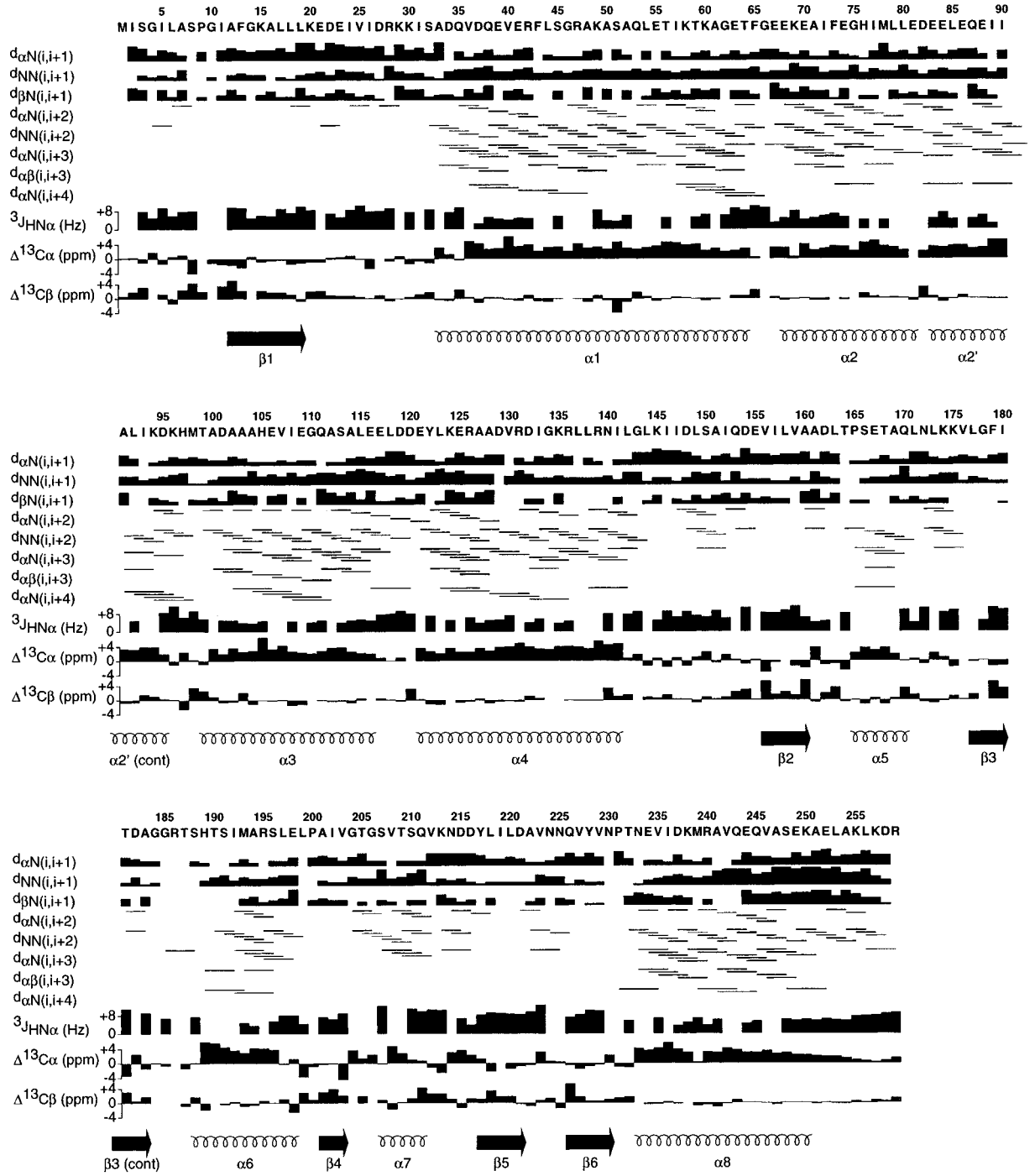

FIGURE 4: Summary of the observed sequential backbone NOEs, $^3J_{HN\alpha}$ coupling constants, and secondary $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts, together with the derived secondary structure of EIN.

is 0.75 ± 0.16 Å for the backbone (N, Cα, C, O) atoms, 1.02 ± 0.13 Å for all non-hydrogen atoms, and 0.77 ± 0.16 Å for all ordered non-hydrogen atoms (84% of the total number of non-hydrogen atoms). While the C-terminal helix extends up to residue 250, the C-terminus begins to become disordered from residue 247 onward (Figure 7A). Indeed, no long-range NOEs are observed for residues 247−259 (Figure 6).

Description of the Structure. Two views showing ribbon diagrams of EIN are displayed in Figure 7B. The structure of EIN is elongated in shape, approximately 78 × 32 × 32 Å, which is reflected in a ratio of 1.0:3.1:3.0 for the three principal components of the inertia tensor (calculated excluding the disordered C-terminal tail from residues 250−259). It is this elongated shape that accounts for the line widths being larger than one might expect for a molecule of ~30 kDa. In particular, resonances arising from $^{15}N-H$ and $^{13}C-H$ vectors that are perpendicular to the long axis of EIN will be approximately 1.5 times as broad as those oriented parallel to this axis. Thus, the backbone ^{15}N $T_{1\rho}$ values range from ~68 ± 6 ms for residues parallel to the long axis to ~48 ± 2 ms for residues perpendicular to the long axis (unpublished data). EIN comprises two domains, an α/β domain (residues 1−20 and 148−230) and an α domain (resides 33−143) connected by two linkers (residues 21−32 and 144−147). The α/β and α domains are positioned at an angle of about 130° relative to each other, giving EIN an L-shaped appearance. In addition, there is a C-terminal

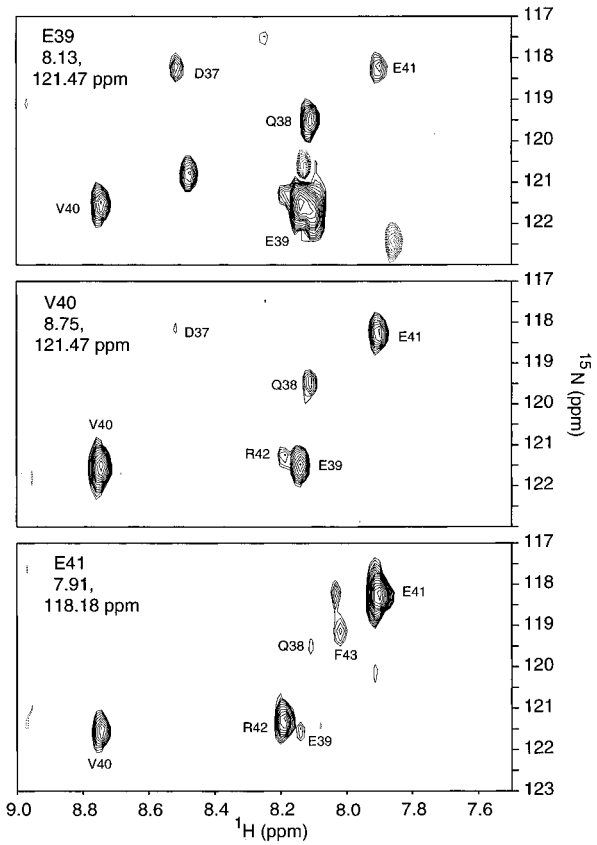

FIGURE 5: Examples of selected $^{15}N(F_3)-^{1}H(F_4)$ planes of the 4D $^{15}N/^{15}N$-separated NOE spectrum recorded with a 170 ms mixing time on perdeuterated ^{15}N-labeled EIN. The absence of spin diffusion pathways involving carbon-attached protons permits a long mixing time to be employed, resulting in a high signal-to-noise ratio for the 4D spectrum.

helix (α8, residues 233−250) which serves as a linker between the N- and C-terminal domains of EI.

The α/*β* domain comprises six *β*-strands (*β*1, 12−19; *β*2, 156−160; *β*3, 177−183; *β*4, 201−203; *β*5, 217−221; and *β*6, 226−231) arranged in a +1*x*, +1*x*, +1*x*, −4, −1 topology and three α-helices (α5, residues 165−170; α6, residues 188−198; and α7, residues 207−211) which link strands *β*2 and *β*3, *β*3 and *β*4, and *β*4 and *β*5. *β*-Strands 1 and 2 are linked via the α domain. Strands *β*1−*β*4 comprise a parallel *β*-sheet, whereas strands *β*1, *β*5, and *β*6 comprise an antiparallel *β*-sheet. This topology is identical to that found for the phosphohistidine domain (residues 390−504) of pyruvate phosphate dikinase (PPDK) with the exception that α7 of EIN is replaced by a three-stranded antiparallel *β*-sheet in PPDK (Herzberg et al., 1996). The two protein domains can be aligned with a Cα RMS difference of 1.8 Å for 83 residues (residues 2−18, 156−167, 168−171, 175−185, 188−210, 213−223, and 224−228 of EIN superimposed on residues 393−409, 424−435, 432−435, 441−451, 454−476, 490−500, and 500−504, respectively, of PPDK).

The α domain comprises four helices (α1, 33−64; α2/α2′, 68−81/83−95; α3, 99−116; and α4, 121−141) arranged in a four-helix bundle with an up−down−up−down topology. The second helix displays a large kink between residues 81 and 83 such that the angle between the long axes of α2 (residues 68−81) and α2′ (residues 83−95) is ∼40°. This kink is stabilized by hydrogen bonds from the carboxylate of Asp82 to the backbone amides of Glu84 and Leu85. Helices α1 and α2 are approximately antiparallel (with interhelical angles of ∼170° and ∼150° between α1 and α2 and between α1 and α2′, respectively), as are helices α3 and α4 (interhelical angle ∼170°). These two pairs of helices are oriented at approximately 50° to each other. The interface between the four helices forms a tightly packed hydrophobic core. In addition, the four-helix bundle is stabilized by five potential electrostatic interactions: between Asp37(α1) and Lys94(α2′), Arg42(α1) and Asp101(α3), Arg47(α1) and Glu86(α2′), Ser51(α1) and Glu81(α2), and Gln53(α1) and Asn140(α4).

The interface between the α and α/*β* domains comprises helices α1, α2, and α4 of the α domain and helices α5 and α6, as well as some loops of the α/*β* domain. It is mainly stabilized by hydrophobic interactions: in particular, between helix α1 (Ile57, Lys60, Thr64, and Phe65), the loop preceding strand *β*2 (Ser150), helix α5 (Ser166, Ala169, and Gln170), the loop connecting helix α5 and strand *β*3 (Leu173), and helix α6 (Met193 and Leu197); between helix α2 (Phe73) and helix α5 (Pro165, Ser166, and Ala169); and between helix α4 (Asp129, Asp132, and Arg136) and helix α5 (Thr164, Pro165, and Ser166). There are three potential electrostatic interactions: between Lys60 and Ser150, Asp132 and Thr164, and Arg136 and Ser166.

Comparison with the X-ray Structure of EIN. Best fit superpositions of the restrained regularized mean NMR structure and the X-ray structure (Liao et al., 1996) of EIN are provided in Figure 8. Only residues 2−249 are visible in the electron density map. It is apparent from Figure 8 that the two structures are very similar. The RMS difference between the backbone (N, Cα, C, O) atoms of the two structures is 1.6 Å for residues 2−249 (1.5 Å for residues 2−246), 1.4 Å for the α/*β* and α domains together (residues 2−230), 1.2 Å for the α/*β* domain (residues 2−20 and 148−230; Figure 8C), and 1.1 Å for the α domain (residues 33−143; Figure 8B). Given that the precision of the backbone coordinates for the NMR structure is ∼0.8 Å, the expected accuracy lies between 1.2 and 1.6 Å (Clore et al., 1993). In the case of the X-ray data, the backbone RMS difference between the two monomers in the assymetric unit is 0.5 Å. Hence, one can conclude that, overall, there is no significant difference between the NMR and X-ray structures within the errors of the coordinates.

The essential agreement between the NMR and X-ray structures is further supported by an examination of angular RMS differences. The average ϕ and ψ angular RMS differences between the two structures are 17° and 19°, respectively [excluding seven residues with deviations greater than 100° in either ϕ or ψ, specifically Ser3 (ψ), Gly4 (ϕ), Ile147 (ψ), Asp148 (ϕ), Gly185 (ϕ), Thr205 (ψ), and Gly206 (ϕ), all of which are located in loops]. The backbone coordinates for these seven residues have generally low precision in the ensemble of NMR structures (>0.8 Å) and high *B*-factors (>50 Å^2) in the X-ray structure. The differences in backbone torsion angles for these seven residues arise from a change in the backbone ϕ angle from negative values in the NMR structures (for Gly4, Asp148, Gly185, and G206) to positive values in the X-ray structure, with compensatory changes in ψ for the $i-1$ residue. The average interhelical angular RMS difference observed in the α domain between the two structures is 2.9 ± 2.0°.

Two features of the comparison of the two structures are noteworthy. First, the backbone RMS difference between the NMR and X-ray structures for the two domains combined

Table 2: Structural Statistics[a]

	⟨SA⟩	(SA)r
Structural Statistics		
RMS deviations from exptl distance restraints (Å)[b]		
all (3048)	0.037 ± 0.001	0.028
interresidue sequential ($\|i-j\|=1$) (952)[c]	0.034 ± 0.001	0.026
interresidue short range ($1<\|i-j\|\leq 5$) (809)	0.048 ± 0.001	0.033
interresidue long range ($\|i-j\|>5$) (586)	0.037 ± 0.004	0.031
intraresidue (471)	0.019 ± 0.004	0.019
H-bonds (230)	0.038 ± 0.004	0.029
RMS deviations from exptl dihedral restraints (deg) (543)[b]	0.252 ± 0.054	0.238
RMS deviations from exptl $^3J_{HN\alpha}$ coupling constants (Hz) (163)[b]	0.97 ± 0.02	0.67
RMS deviations from exptl ^{13}C shifts		
$^{13}C\alpha$ (ppm) (257)	1.10 ± 0.02	1.00
$^{13}C\beta$ (ppm) (241)	1.00 ± 0.02	0.99
deviations from idealized covalent geometry		
bonds (Å) (4045)	0.004 ± 0.0004	0.005
angles (deg) (7373)	0.539 ± 0.015	0.564
impropers (deg) (1954)	0.515 ± 0.023	0.612
Measures of Structure Quality		
E_{L-J} (kcal·mol^{-1})[d]	−1093 ± 13	−1075
PROCHECK[e]		
% residues in most favorable region of Ramachandran plot	93.2 ± 0.9	90.8
no. of bad contacts/100 residues	2.9 ± 0.8	5.4
H-bond energy	0.74 ± 0.05	0.70
overall *G*-factor	0.20 ± 0.03	0.15
WHATIF[e]		
packing score	−0.25 ± 0.03	−0.36
torsion angle score	0.35 ± 0.02	0.02
position-specific rotamer score	0.68 ± 0.005	0.52
Coordinate Precision[f]		
backbone (Å)	0.75 ± 0.16	
all atoms (Å)	1.02 ± 0.13	
all ordered atoms (Å)	0.77 ± 0.16	

[a] The notation of the NMR structures is as follows: ⟨SA⟩ are the final 50 simulated annealing structures; $\overline{SA}$ is the mean structure obtained by averaging the coordinates of the individual SA structures (residues 1−246) best fitted to each other; ($\overline{SA}$)r is the restrained minimized mean structure obtained by restrained regularization of the mean structure $\overline{SA}$. The number of terms for the various restraints is given in parentheses. The final force constants employed for the various terms in the target function used for simulated annealing are as follows: 1000 kcal·mol^{-1}·Å^{-2} for bond lengths, 500 kcal·mol^{-1}·rad^{-2} for angles and improper torsions (which serve to maintain planarity and chirality), 4 kcal·mol^{-1}·Å^{-4} for the quartic van der Waals repulsion term (with the hard sphere effective van der Waals radii set to 0.8 times their value used in the CHARMM PARAM19/20 parameters), 30 kcal·mol^{-1}·Å^{-2} for the experimental distance restraints (interproton distances and hydrogen bonds), 200 kcal·mol^{-1}·rad^{-2} for the torsion angle restraints, 1 kcal·mol^{-1}·Hz^{-2} for the coupling constant restraints, 0.5 kcal·mol^{-1}·ppm^{-2} for the carbon chemical shift restraints, and 1.0 for the conformational database potential. [b] None of the structures exhibited distance violations greater than 0.5 Å, dihedral angle violations greater than 5°, or $^3J_{HN\alpha}$ coupling constant violations greater than 2 Hz. The torsion angle restraints comprise 255 ϕ, 6 ψ, 168 χ_1, and 114 χ_2, angles. [c] Only structurally useful intraresidue NOEs are included in the intraresidue interproton distance restraints. Thus, intraresidue NOEs between protons separated by two bonds or between non-stereospecifically assigned protons separated by three bonds are not incorporated in the restraints. [d] E_{L-J} is the Lennard-Jones van der Waals energy calculated with the CHARMM PARAM19/20 protein parameters (Brooks et al., 1983) and is *not* included in the target function for simulated annealing or restrained minimization. [e] The programs PROCHECK (Laskowski et al., 1993) and WHATIF (Vriend & Sander, 1993) were used to assess the overall quality of the structures. For the PROCHECK statistics, less than 10 bad contacts per 100 residues, a hydrogen bond energy of 0.6−1.0, and an overall *G*-factor of greater than −0.5 are expected for a good quality structure. For the WHATIF statistics, a packing score greater than −0.5 is indicative of a high quality structure; a torsion angle score of less than −2 for any residue is poor; a position-specific rotamer score of 1.0 indicates that all rotamers are in their preferred conformations, and a score of 0.0 indicates that no rotamers are in preferred orientations. [f] The precision of the atomic coordinates is defined as the average RMS difference between the 50 final simulated annealing structures and the mean coordinates, $\overline{SA}$. The values given relate to residues 1−246, since the C-terminus begins to become disordered from residue 247 onward, and no long-range NOEs are observed for residues 247−259. The values given for the backbone atoms relate to the N, Cα, C, and O atoms; those given for all atoms and for all ordered atoms refer only to non-hydrogen atoms. The ordered atoms comprise 84% of the total number of non-hydrogen atoms.

is about 20% larger than the backbone RMS difference for the individual domains. This is due to a ∼0.5° difference in the relative orientations of the two domains. As the Cα−Cα distance from the tip of the α domain (residue 33) to the tip of the α/β domain (residue 223) is ∼63 Å, this very small difference in relative orientations translates to larger than average RMS differences at the two ends of EIN. Second, the orientation of the C-terminal helix (α8) with respect to the α/β domain differs by ∼7° in the two structures (Figure 8A). The C-terminal helix interacts with the undersurface of the α/β domain and is stabilized by hydrophobic interactions in both structures between Val235 (α8) and Ile219 (β5), Val235(α8) and Tyr228(β6), Met239 (α8) and Ile219 (β5), and Val246 (α8) and Val223 (turn between β5 and β6). There are no further interactions beyond residue 246. In addition, there is a hydrogen bond between the side chains of Asn224 and Gln243 in the X-ray structure which serves to position the C-terminal end of α8 with regard to the α/β domain. We failed to observe any NOEs between these two residues. The absence of any restraints between these two residues is responsible for the displacement of the C-terminal end of helix α8 in the NMR structure. Whether indeed this difference is real or not cannot be ascertained from our available data.

The Active Site. EIN catalyzes reversible phosphotransfer between His15 of HPr and His189 of EIN (Postma et al.,

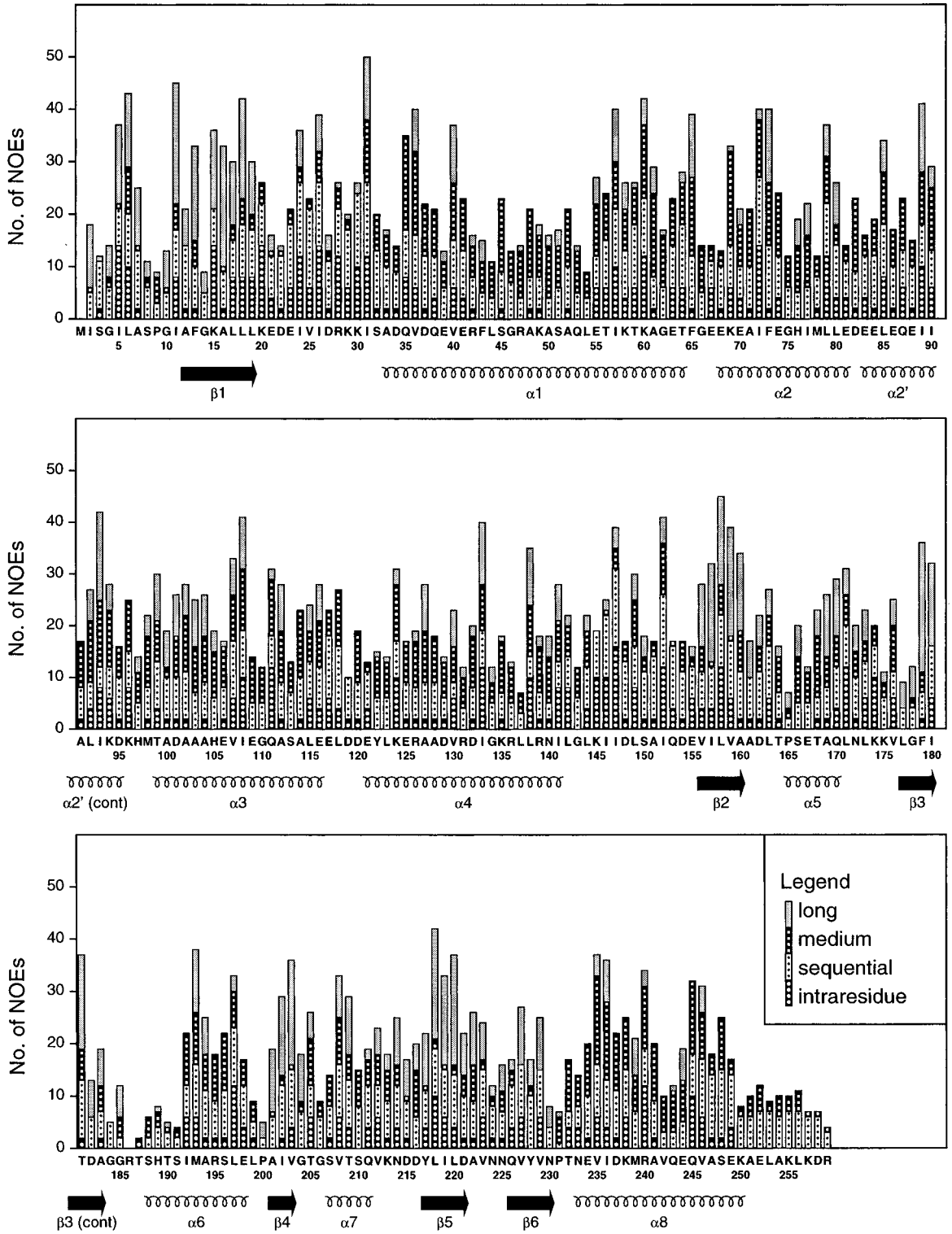

FIGURE 6: Summary of the distribution of NOEs as a function of residue number observed for EIN. The intraresidue restraints shown represent only those that are structurally useful (i.e., not redundant with the covalent geometry; see footnote *c* to Table 2).

1993). The location of His189, at the C-terminal end of helix α6, is indicated in the ribbon diagrams shown in Figure 7B, and a stereoview of the active site and surrounding residues is shown in Figure 9. His189 is located in a cleft formed by helices α2, α4, α5, and α6, at the junction of the α and α/β domains.

On the basis of the values of the $^{3}J_{C'C\gamma(\text{aromatic})}$ (2.1 ± 0.2 Hz) and $^{3}J_{NC\gamma(\text{aromatic})}$ (2.8 ± 0.1 Hz) coupling constants measured from 2D $^{13}C'-\{^{13}C\gamma\}$ and $^{15}N-\{^{13}C\gamma\}$ spin−echo difference $^{1}H-^{15}N$ HSQC spectra (Hu et al., 1996), we conclude that the C′−Cα−Cβ−Cγ and N−Cα−Cβ−Cγ torsion angles of His189 are gauche and trans, respectively, indicating that the side chain χ_1 angle of His189 is in the trans conformation. (For reference, the trans $^{3}J_{C'C\gamma(\text{aromatic})}$ and $^{3}J_{NC\gamma(\text{aromatic})}$ couplings have values of 3.5−4.5 and 2.5−3.0 Hz, respectively; Hu et al., 1996.) The strong NOE observed between the Hδ2 atom of His189 and the methyl group of Thr168 orients the Nε2 atom of His189 toward the hydroxyl group of Thr168, in agreement with the postulated hydrogen bond between His189(Nε2) and Thr168(Oγ)

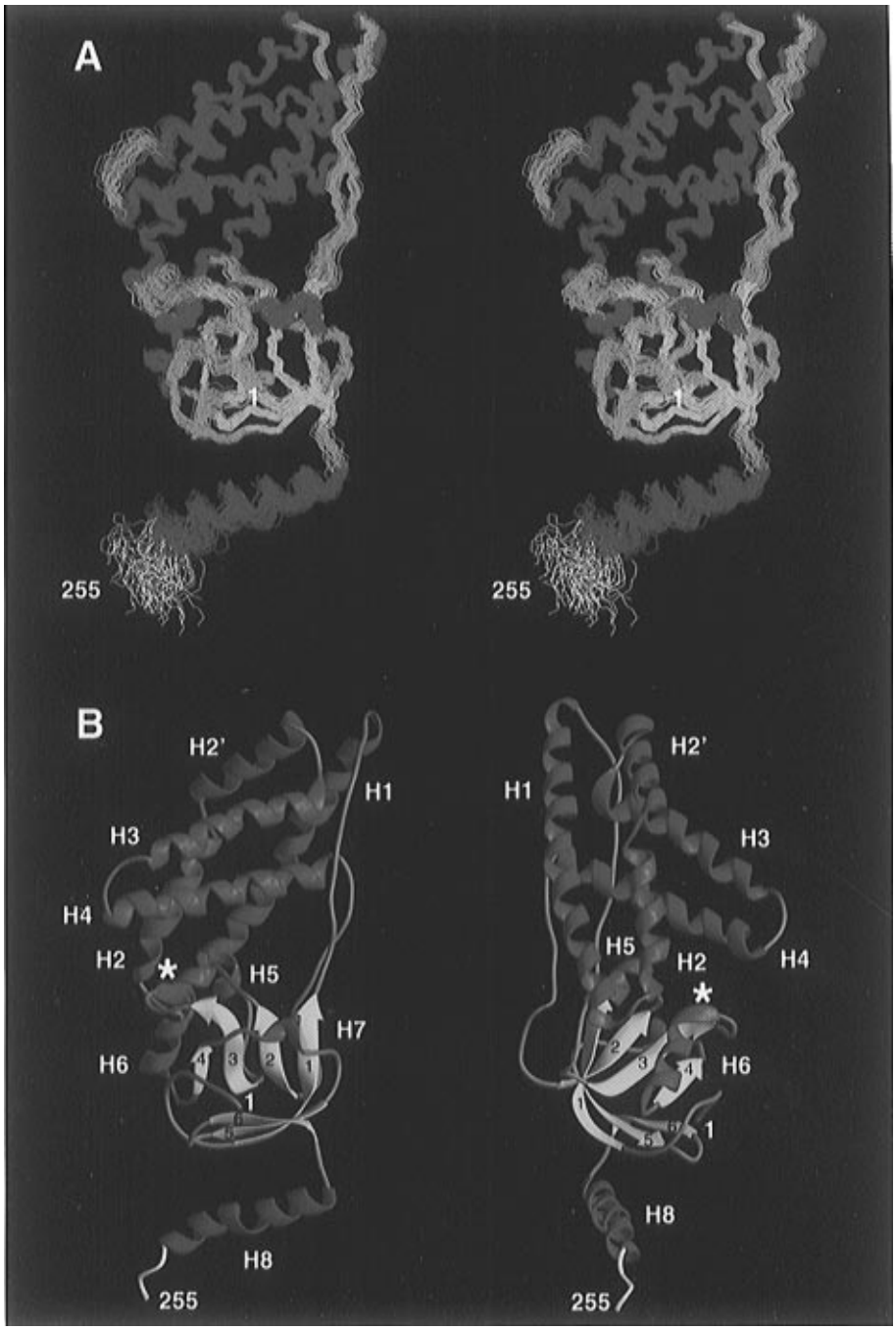

FIGURE 7: (A) Stereoview showing a superposition of the backbone (N, Cα, C) atoms of the 50 final simulated annealing structures of EIN. (B) Ribbon diagrams illustrating two views of the backbone of EIN. Helices are shown in red, strands are in yellow, loops are in blue, and the disordered C-terminus is in white. The asterisk in (B) indicates the location of the active site histidine (His189). (A) was generated with the program MolMol (Koradi et al., 1996) and (B) with the program RIBBONS (Carson, 1987).

observed in the crystal structure (Liao et al., 1996). In this light, it is worth noting that nitrogen and carbon cannot be distinguished crystallographically and that at 2.5 Å resolution it would not be possible to distinguish *a priori* between a g^+ χ_2 conformation with the Nϵ2 atom hydrogen bonded to the Oγ of Thr168 and a g^- χ_2 conformation with the Cϵ1 atom directed toward the Oγ of Thr168. However, these two possibilities can be differentiated on chemical grounds since the measured distance of 2.77 Å between the Nϵ2 atom of His189 and the Oγ atom of Thr168 in the crystal structure is short enough to preclude the alternate conformation. There are two types of potential hydrogen bonds between His189-(Nϵ2) and Thr168(Oγ): either the Hϵ2 proton of His189 donates a hydrogen bond to the Oγ of Thr168 or the unprotonated Nϵ2 atom of His189 accepts a hydrogen bond from the OγH hydroxyl proton of Thr168. Since no nitrogen-bonded imidazole ring proton is observed in the NMR spectrum, this suggested that the second possibility is the more likely.

Confirmation that the Nϵ2 atom of His189 accepts a hydrogen bond from the OγH hydroxyl proton of Thr168 was obtained from the pattern of cross-peaks and ^{15}N chemical shifts of the Nϵ2 and Nδ1 atoms observed in the long-range ^{1}H$-^{15}$N correlation spectrum shown in Figure 10. In such a spectrum, cross-peaks of approximately equal intensity are observed for the Nϵ2$-$Hϵ1, Nϵ2$-$Hδ2 and Nδ1$-$Hϵ1 two-bond correlations ($^2J_{NH} \sim 6-10$ Hz; Blomberg et al., 1977), while either a very weak or absent cross-peak

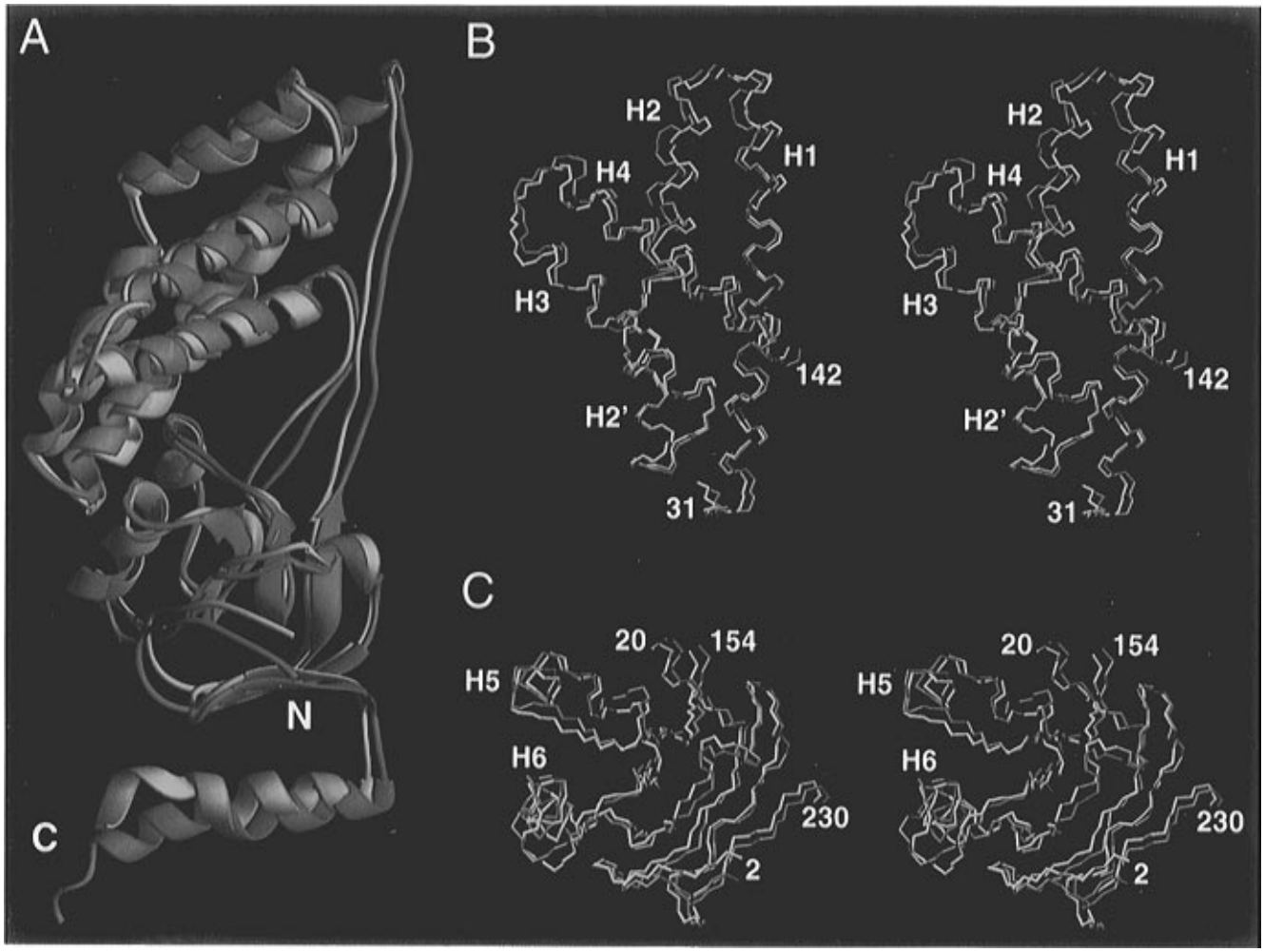

FIGURE 8: Comparison of the solution (red) and X-ray (blue) structures of EIN. (A) Ribbon diagram showing a superposition of the two structures best fitted to residues 2−246. (B and C) Backbone (N, Cα, C) traces showing superpositions of the α (best fitted to residues 31−142) and α/β (best fitted to residues 2−20 and 154−230) domains, respectively. (A) was generated with the program RIBBONS (Carson, 1987) and (B) and (C) with the program MolMol (Koradi et al., 1996). The coordinates of the X-ray structure (PDB accession code 1ZYM) were taken from Liao et al. (1996).

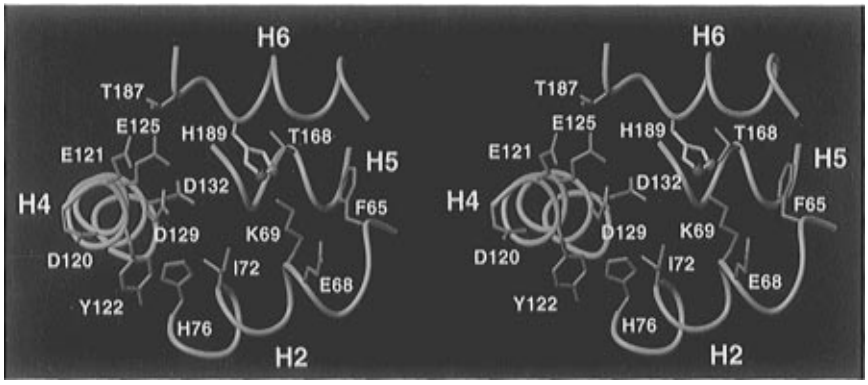

FIGURE 9: Stereoview of the active site of EIN. The backbone is shown as a blue tube, side chains are in red, and the active site histidine (His189) is in yellow. The Nε2 and Cε1 atoms of His189 are represented as a green and purple sphere, respectively. The figure was generated with the program MolMol (Koradi et al., 1996).

is observed for the Nδ1−Hδ2 three-bond correlation ($^3J_{NH}$ ∼ 2−3 Hz; Blomberg et al., 1977). For a neutral histidine, the protonated nitrogen resonates at ∼168 ppm, while the unprotonated nitrogen resonates at ∼250 ppm (Bachovchin, 1986). In general, the Nε2−H tautomer is the more stable species (Blomberg et al., 1977). For a positively charged fully protonated histidine, the two nitrogens resonate around 175 ppm, with the Nδ1 atom generally resonating about 1 ppm to lower field than the Nε2 atom (Bachovchin, 1986). At pH 7, the Nδ1 atoms of His76, His97, and His105 resonate at 248.4, 195.5, and 219.7 ppm, respectively, while the corresponding Nε2 atoms resonate at 165.5, 178.2, and 183.3 ppm, respectively (Figure 10). His76 is thus completely neutral, while the shifts for His97 and His105 indicate the presence of a rapid equilibrium between neutral and positively charged states in which the neutral form is the common Nε2−H tautomer. In contrast, at pH 7 the Nδ1 and Nε2 atoms of His189 resonate at 190.6 and 217.6 ppm,

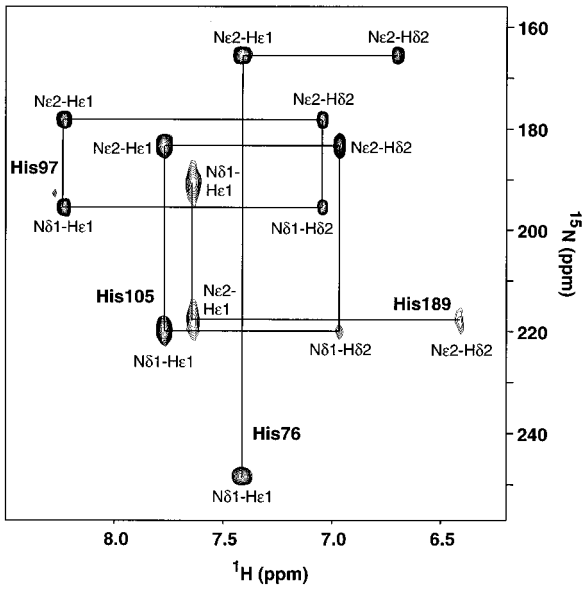

FIGURE 10: Long-range ^{1}H−^{15}N correlation spectrum of EIN at pH 7. From the chemical shifts of the Nδ1 and Nε2 atoms, one can conclude that the neutral species of His189 is the Nδ1−H tautomer, while for His76, His97, and His105 the neutral species is the Nε2−H tautomer (see text). The spectrum was recorded with a 22 ms dephasing delay during which time the ^{1}H and ^{15}N signals become antiphase (Pelton et al., 1993).

respectively, indicating that the neutral species is the Nδ1−H tautomer (Figure 10).

The pK_a of His189, obtained by recording a series of ^{1}H−^{13}C correlation spectra to follow the pH dependence of the chemical shifts of the Hδ2 and Hε1 resonances, is 6.3 ± 0.2 (uncorrected for the deuterium isotope effect on the glass electrode). This pK_a value is comparable to that of a free histidine (pK_a ∼ 6.5; Creighton, 1993) and is slightly higher than that for His15 of HPr (pK_a ∼ 5.6−6.1; Gassner et al., 1977; Kalbitzer et al., 1982; Dooijewaard et al., 1979). (For reference, the pK_a values of the other three histidine residues in EIN, namely, His76, His97, and His105, are <6, 7.3 ± 0.2, and 6.4 ± 0.1, respectively; a more accurate estimate for the pK_a of His76 could not be obtained since the protein aggregates below pH 5.8. The depressed pK_a of His76 is due to the fact that His76 is buried and has a solvent surface accessibility of only 5% relative to that of a Gly-His-Gly tripeptide segment.)

In the phosphorylated form of EIN, the phosphoryl group is thought to be bonded to the Nε2 atom of His189 (Weigel et al., 1982). In both the NMR and X-ray structures of free EIN, only the Nδ1 atom of His189 is solvent accessible. To permit the Nε2 atom to become accessible to an incoming phosphate, the conformation of His189 must change relative to that in free EIN. As discussed by Liao et al. (1996), this can be accomplished in two ways: either by leaving the trans χ_1 rotamer of His189 unaltered and changing the χ_2 angle from the g^+ to the g^- conformation or by changing the χ_1 angle of His189 to the g^- rotamer. In the former case His189 is located in the shallow depression, while in the latter case it is located in the deep depression at the interface of the two subdomains. Alternatively, a more substantial conformational change of the polypeptide chain around His189 may occur. At the present time, we do not have any data to distinguish between these possibilities.

There are a large number of negatively charged Asp and Glu residues in the active site region (Figure 9). It is possible, by analogy with the complex between HPr and enzyme IIA^{glc}, that one or more of these could provide a potential acceptor for a salt bridge with the invariant Arg17 of HPr which is essential for the phosphoryl transfer activity of HPr (Sharma et al., 1991; Andersen et al., 1993).

CONCLUDING REMARKS

In this paper we have demonstrated that it is possible to determine the three-dimensional solution structure of a 259-residue protein, namely, EIN, by means of multidimensional NMR spectroscopy. A key feature of this structure determination was the use of perdeuterated samples to permit virtually complete backbone assignments to be obtained and to record a high-quality 4D ^{15}N/^{15}N-separated NOE spectrum which provided many useful NH−NH NOEs with which to define both the helices and sheets. Also useful was a reversed-labeled U-^{12}C(Tyr/Phe)/^{13}C sample of EIN to facilitate both the assignment and the observation of NOEs involving aromatic protons. Given the precision of the backbone coordinates of the NMR (∼0.8 Å) and X-ray (∼0.5 Å) structures, there is no significant overall difference in the structure of EIN in the solution and crystal states. The conformation of the active site His189 is locked in a single trans χ_1 rotamer conformation with the Nε2 atom directed toward the Oγ1 atom of Thr168 in both the NMR and crystal structures. In addition, the NMR data indicate that the neutral form of His189 is the Nδ1−H tautomer with the Nε2 atom of His189 accepting a hydrogen bond from the OγH hydroxyl proton of Thr168. This is functionally important since a conformational change in the side chain of His189 must take place if the histidine becomes phosphorylated at the Nε2 position.

ACKNOWLEDGMENT

We thank Frank Delaglio for software support; Rolf Tschudin for technical support; Jin-Shan Hu, Stephan Grzesiek, and Ad Bax for providing us with the pulse sequences for the 2D ^{13}C′−{^{13}Cγ} spin−echo difference and ^{15}N−{^{13}Cγ} spin−echo difference ^{1}H−^{15}N HSQC spectra; Ad Bax, John Kuszewski, Nico Tjandra, and Andy Wang for numerous stimulating discussions; Osnat Herzberg for providing us with the coordinates of PPDK; and David Davies for providing us with the X-ray coordinates of EIN prior to publication.

REFERENCES

Anderson, J. W., Pullen, K., Georges, F., Klevit, R. E., & Waygood, E. B. (1993) *J. Biol. Chem. 268*, 12325−12333.

Bachovchin, W. W. (1986) *Biochemistry 25*, 7751−7759.

Bax, A. (1994) *Curr. Opin. Struct. Biol. 4*, 738−744.

Bax A., & Grzesiek, S. (1993) *Acc. Chem. Res. 26*, 131−138.

Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., & Zhu, G. (1994) *Methods Enzymol. 239*, 79−106.

Blomberg, F., Mauzer, W., & Rüterjans, H. (1977) *J. Am. Chem. Soc. 99*, 8149−8159.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Brünger, A. T. (1993) *X-PLOR Version 3.1 Manual*, Yale University, New Haven, CT.

Carson, M. (1987) *J. Mol. Graphics 5*, 103−106.

Clore, G. M., & Gronenborn, A. M. (1989) *CRC Crit. Rev. Biochem. Mol. Biol. 24*, 479−564.

Clore, G. M., & Gronenborn, A. M. (1991) *Science 252*, 1390−1399.
Clore, G. M., & Gronenborn, A. M. (1994) *Protein Sci. 3*, 372−390.
Clore, G. M., Robien, M. A., & Gronenborn, A. M. (1993) *J. Mol. Biol. 231*, 82−102.
Creighton, T. E. (1993) *Proteins: Structure and Molecular Properties*, Freeman, New York.
Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
Dooijewaard, G., Roosien, F. F., & Robillard, G. T. (1979) *Biochemisry 18*, 2996−3001.
Fairbrother, W. J., Gippert, G. P., Reizer, J., Saier, M. H., & Wright, P. E. (1992) *FEBS Lett. 296*, 148−152.
Farmer, B. T., II, & Venters, R. A. (1995) *J. Am. Chem. Soc. 117*, 4187−4188.
Fogh, R. H., Schipper, D., Boelens, R., & Kaptein, R. (1994) *J. Biomol. NMR 4*, 123−128.
Garrett, D. S., Powers, R., Gronenborn, A. M., & Clore, G. M. (1991) *J. Magn. Reson. 95*, 214−220.
Garrett, D. S., Kuszewski, J., Hancock, T. J., Lodi, P. J., Vuister, G. W., Gronenborn, A. M., & Clore, G. M. (1994) *J. Magn. Reson., Ser. B 104*, 99−103.
Gassner, M., Stehlik, D., Schreckner, O., Hengstenberg, W., Maurer, W., & Rüterjans, H. (1977) *Biochemistry 75*, 287−296.
Grzesiek, S., Anglister, J., Ren, H., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 4369−4370.
Grzesiek, S., Wingfield, P. T., Stahl, S., Kaufman, J. D., & Bax, A. (1995) *J. Am. Chem. Soc. 117*, 9594−9595.
Hammen, P. K., Waygood, E. B., & Klevit, R. E. (1991) *Biochemistry 30*, 11842−11850.
Herzberg, O., & Klevit, R. E. (1994) *Curr. Opin. Struct. Biol. 4*, 814−822.
Herzberg, O., Reddy, P., Sutrina, S., Saier, M. H., Reizer, J., & Kapadia, G. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2499−2503.
Herzberg, O., Chen, C. C. H., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., & Dunaway-Mariano, D. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2652−2657.
Hu, J.-S., Grzesiek, S., & Bax, A. (1996) *J. Magn. Reson.* (in press).
Jia, Z., Quail, J. W., Waygood, E. B., & Delbaere, L. T. J. (1993) *J. Biol. Chem. 268*, 22490−22501.
Kalbitzer, H. R., Hengstenberg, W., Rösch, P., Muss, P., Bernsmann, P., Engelmann, R., Dörschung, M., & Deutscher, J. (1982) *Biochemistry 21*, 2879−2825.
Klevit, R. E., & Waygood, E. B. (1986) *Biochemistry 25*, 7760−7769.
Koradi, R., Billeter, M., & Wüthrich, K. (1996) *J. Mol. Graphics 14*, 52−55.
Kuszewski, J., Qin, J., Gronenborn, A. M., & Clore, G. M. (1995) *J. Magn. Reson., Ser. B 106*, 92−96.
Kuszewski, J., Gronenborn, A. M., & Clore, G. M. (1996) *Protein Sci. 6*, 1067−1080.
Kuszewski, J., Gronenborn, A. M., & Clore, G. M. (1997) *J. Magn. Reson.* (in press).
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Lee, B. R., Lecchi, P., Lannell, L., Jaffe, H., & Peterkofsky, A. (1994) *Arch. Biochem. Biophys. 312*, 121−124.
Liao, D.-I., Kapadia, G., Reddy, P., Saier, M. H., Reizer, J., & Herzberg, O. (1991) *Biochemistry 30*, 9583−9594.
Liao, D.-I., Silverton, E., Seok, Y.-J., Lee, B. R., Peterkofsky, A., & Davies, D. R. (1996) *Structure 4*, 861−872.
Licalsi, C., Crocenzi, T. S., Freire, E., & Roseman, S. (1991) *J. Biol. Chem. 266*, 19519−19527.
Metzler, W. J., Wittekind, M., Goldfarb, V., Mueller, L., & Farmer, B. T. (1996) *J. Am. Chem. Soc. 118*, 6800−6801.
Nilges, M. (1993) *Proteins: Struct., Funct., Genet. 17*, 295−309.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988a) *FEBS Lett. 229*, 129−136.
Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988b) *Protein Eng. 2*, 27−38.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1990) *Biopolymers 29*, 813−822.
Pelton, J. G., Torchia, D. A., Meadow, N. D., & Roseman, S. (1993) *Protein Sci. 2*, 543−558.
Postma, P. W., Lengeler, J. W., & Jacobson, G. R. (1993) *Microbiol. Rev. 57*, 543−594.
Remerowski, M. L., Domke, T., Groenewegen, A., Pepermans, H. A. M., Hilbers, C. W., & van de Ven, F. J. M. (1994) *J. Biomol. NMR 4*, 257−278.
Reynolds W. F., Peat, I. R., Freeman, M. H., & Lyerla, J. R. (1973) *J. Am. Chem. Soc. 95*, 338−331.
Rice, L. M., & Brünger, A. T. (1994) *Proteins: Struct., Funct., Genet. 19*, 277−290.
Seok, Y.-J., Lee, B. R., Zhu, P.-P., & Peterkofsky, A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 347−351.
Shan, X., Hardner, K. H., Muhandiram, D. R., Rao, N. S., Arrowsmith, C. H., & Kay, L. E. (1996) *J. Am. Chem. Soc. 118*, 6570−6579.
Sharma, S., Georges, F., Delbaere, L. T. J., Lee, J. S., Klevit, R. E., & Waygood, E. B. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4877−4881.
Spera, S., & Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490−5492.
Stein, E. G., Rice, L. M., & Brünger, A. T. (1997) *J. Magn. Reson., Ser. B* (in press).
Torchia, D. A., Sparks, S. W., & Bax, A. (1988) *J. Am. Chem. Soc. 110*, 2320−2321.
van Nuland, N. A., Hangyi, I. W., Schaik, R. C., Berendsen, H. J. C., van Gunsteren, W. F., Scheek, R. M., & Robillard, G. T. (1994) *J. Mol. Biol. 237*, 544−559.
van Nuland, N. A. J., Boelens, R., Scheek, R. M., & Robillard, G. T. (1995) *J. Mol. Biol. 246*, 180−193.
Venters, R. A., Metzler, W. J., Spicer, L. D., Mueller, L., & Farmer, B. T., II (1995) *J. Am. Chem. Soc. 117*, 9592−9593.
Venters, R. A., Farmer, B. T., Fierke, C. A., & Spicer, L. D. (1996) *J. Mol. Biol. 264*, 1101−1116.
Vriend, G., & Sander, C. (1993) *J. Appl. Crystallogr. 26*, 47−60.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1986) *J. Mol. Biol. 196*, 611−639.
Weigel, N., Kukuzurinska, M. A., Nakazawa, A., Waygood, E. B., & Roseman, S. (1982) *J. Biol. Chem. 257*, 14477−14491.
Wittekind, M., & Mueller, L. (1993) *J. Magn. Reson., Ser. B 101*, 201−205.
Wittekind, M., Rajagopal, P., Branchini, B. R., Reizer, J., Saier, M. H., & Klevit, R. E. (1992) *Protein Sci. 1*, 1363−1376.
Worthylake, D., Meadow, N. D., Roseman, S., Liao, D.-I., Herzberg, O., & Remington, S. J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10382−10386.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., & Kay, L. E. (1994) *J. Am. Chem. Soc. 116*, 11655−11666.

BI962924Y