Accelerated Publications

Sequence-Specific ^1H NMR Assignments and Secondary Structure in Solution of *Escherichia coli trp* Repressor†

C. H. Arrowsmith,* R. Pachter, R. B. Altman, S. B. Iyer, and O. Jardetzky
Stanford Magnetic Resonance Laboratory, Stanford University, Stanford, California 94305-5055
Received April 20, 1990

ABSTRACT: Sequence-specific ^1H NMR assignments are reported for the active L-tryptophan-bound form of *Escherichia coli trp* repressor. The repressor is a symmetric dimer of 107 residues per monomer; thus at 25 kDa, this is the largest protein for which such detailed sequence-specific assignments have been made. At this molecular mass the broad line widths of the NMR resonances preclude the use of assignment methods based on ^1H–^1H scalar coupling. Our assignment strategy centers on two-dimensional nuclear Overhauser spectroscopy (NOESY) of a series of selectively deuterated repressor analogues. A new methodology was developed for analysis of the spectra on the basis of the effects of selective deuteration on cross-peak intensities in the NOESY spectra. A total of 90% of the backbone amide protons have been assigned, and 70% of the α and side-chain proton resonances are assigned. The local secondary structure was calculated from sequential and medium-range backbone NOEs with the double-iterated Kalman filter method [Altman, R. B., & Jardetzky, O. (1989) *Methods Enzymol. 177*, 218–246]. The secondary structure agrees with that of the crystal structure [Schevitz, R., Otwinowski, Z., Joachimiak, A., Lawson, C. L., & Sigler, P. B. (1985) *Nature 317*, 782], except that the solution state is somewhat more disordered in the DNA binding region and in the N-terminal region of the first α-helix. Since the repressor is a symmetric dimer, long-range intersubunit NOEs were distinguished from intrasubunit interactions by formation of heterodimers between two appropriate selectively deuterated proteins and comparison of the resulting NOESY spectrum with that of each selectively deuterated homodimer. Thus, from spectra of three heterodimers, long-range NOEs between eight pairs of residues were identified as intersubunit NOEs, and two additional long-range intrasubunits NOEs were assigned.

The *trp* repressor from *Escherichia coli* regulates the initiation of transcription by binding to operator DNA sequences in three operons involved in tryptophan biosynthesis (Klig et al., 1988). The affinity of the repressor for its operator targets is allosterically controlled by L-tryptophan. The repressor is a symmetric dimer of 25 kDa with 107 amino acid residues per subunit and is a member of the helix–turn–helix family of DNA binding proteins. Although much is known from the X-ray crystal structures of the apo and active forms of the repressor (Schevitz et al., 1985; Zhang et al., 1987) as well as its complex with operator DNA (Otwinowski et al., 1988), it is desirable to study this system in solution. The crystal data indicate the presence of "flexible" regions of the protein (Lawson et al., 1988), and the solution structure and dynamics will provide important insights into the mechanisms of both specific and nonspecific DNA binding. We are studying the *trp* repressor/operator system by nuclear magnetic resonance (NMR) spectroscopy in an effort to understand the processes of molecular recognition involved in DNA binding and allosteric regulation. A prerequisite for the study of the system by NMR is the assignment of the resonances in the NMR spectrum to specific protons in the macromolecules. This step has already been carried out for an operator DNA fragment, and a solution structure has been determined (Lefèvre et al., 1987). Here we report the sequence-specific assignments for the *trp* repressor together with the regions of regular secondary structure.

The *trp* repressor presents a number of challenges to the NMR spectroscopist. With a molecular mass of 25 kDa the *trp* repressor has a relatively long rotational correlation time, which gives rise to short relaxation times and broad spectral lines. Due to the antiphase nature of the cross-peaks in two-dimensional (2D) *J*-correlated spectra (COSY spectra), broad line widths result in cancellation of the positive and negative components of the cross-peak. This is a particular problem in the fingerprint region of the mostly α-helical repressor where the αCH–NH coupling constant is relatively small and the NH resonances are quite broad. Since these cross-peaks are a vital link in the sequential assignment process (Wüthrich, 1986), this method could not be used. A second consequence of large molecular size (by NMR standards) is the increased importance of spin diffusion (Madrid & Jardetzky, 1988; Madrid et al., 1989). The transfer of magnetization by cross-relaxation throughout the spin network can prevent significant buildup of magnetization between two adjacent protons, thereby reducing the sensitivity of the 2D nuclear Overhauser experiment (NOESY) and at the same time complicating the interpretation of NOE data in terms of internuclear distances. Because *trp* repressor is mostly α-helical, the αCH protons resonate over a relatively narrow range of chemical shifts (Szilágyi & Jardetzky, 1989). Thus, with 107 residues there is severe overlap in many regions of the 2D spectrum, especially the αCH–NH region. To overcome some of these problems, we have used selective deuteration to render the 2D NMR spectra more interpretable. Selective deuteration simplifies the ^1H NMR spectra because deuterium does not resonate at the proton frequency (Markley et al., 1968). In addition, if the level of deuterium incorporation is high enough, the remaining protons will be surrounded

† This work was supported by NIH Grants GM33385 and RR02300.
* Corresponding author.

0006-2960/90/0429-6332$02.50/0

Table I: Isotopic Composition of Selectivity Deuterated *trp* Repressor Analogues

analogue[a]	deuterium content of amino acids[b]		
	<10% deuterium	partially deuterated	>80% deuterium
GIT	Gly, Ile, Thr, Trp	Glx, Ala, Asx, Ser	Pro, Met, Leu, Val, Phe, Tyr, His, Lys, Arg
HLW	His, Leu, Trp	Glx, Ala, Asx, Gly, Ser	Pro, Met, Val, Ile, Phe, Lys, Arg, Tyr, Thr
GMAPS	Gly, Met, Ala, Pro, Ser, Trp	Glx, Asx	Leu, Ile, Val, Phe, Tyr, His, Thr, Lys, Arg
GRIFSTV	Gly, Arg, Ile, Phe, Ser, Thr, Bal, Trp	Glx, Asx, Ala	Leu, Met, Pro, His, Tyr, Lys
SHAGED	Ser, His, Ala, Gly, Glx, Asx, Trp	Pro	Met, Leu, Ile, Val, Lys, Arg, Thr, Phe, Tyr
VFT	Val, Phe, Thr	Glx, Asx, Ala, Ser, Gly	Pro, Met, Leu, Ile, Lys, Arg, Tyr, His

[a] The names of the analogues are based on the protonated amino acids used in the media from which the protein was isolated. [b] Glx = Glu and Gln; Asx = Asp and Asn.

by fewer other protons and will have fewer pathways for magnetization transfer. This affects the relaxation times and NOE intensities (Pachter et al., unpublished work) of the remaining protons.

Materials and Methods

Protein Preparation. *trp* repressor was isolated from *E. coli* strains CY15070 and CY15071 containing the overproducing plasmids pJPR2 and both pJPR2 and pMS421, respectively. The growth and isolation are essentially that described by Paluh and Yanofsky (1986), except in the case of selectively deuterated proteins where a mixture of deuterated amino acids (algal lysate from MSD Isotopes) and a 10-fold excess of the appropriate natural abundance amino acids was used rather than yeast extract and casein hydrolysate. Because there are a number of protonated species in the media (glucose, H_2O), there is some dilution of the deuterium label due to amino acid biosynthesis by the bacteria. Thus, the metabolically active amino acids such as Glx, Ala, and Asx and to a lesser extent Gly and Ser were at least partially protonated in all analogues. Table I lists the six different selectively deuterated analogues prepared in this study. The isotopic composition of each analogue was verified by checking for the appropriate spin systems in the COSY spectrum and by GC–MS analysis of an acid hydrolysis of several of the proteins.

Heterodimer Formation. Three heterodimers were prepared as follows: *VFT/SHAGED*: A total of 200 μL of a solution that was 2.2 mM in SHAGED and 6.6 mM in L-Trp in NMR H_2O buffer (see below) was mixed with 220 μL of a 2.0 mM solution of VFT in the same buffer (no tryptophan). The mixture was incubated at 37 °C overnight and then heated for 2 h at 60 °C. Before the NOESY spectrum was started, the L-Trp concentration was brought up to 6.6 mM by addition to a lyophilized 30-μL aliquot of a 10 μg/mL solution of L-Trp. *VFT/Natural Abundance*: A total of 200 μL of a 2.0 mM solution of VFT with 6.0 mM L-tryptophan in H_2O NMR buffer was mixed with 333 μL of a 1.2 mM solution of natural abundance repressor in D_2O NMR buffer. The mixture was concentrated to 400 μL in a Centricon 10 microconcentrator (Amicon) and then incubated overnight at 37 °C. The L-Trp concentration was brought up to 6 mM as above, and the solution was then lyophilized and taken up in 400 μL of "100%" D_2O (Cambridge Isotope Labs). *VFT/GIT*: A total of 160 μL of a 2.0 mM solution of VFT with 6.0 mM L-Trp in H_2O buffer was mixed with 246 μL of a solution of 1.3 mM GIT and 6.0 mM L-Trp in H_2O buffer. The solution was heated for 2 h at 60 °C before an NMR spectrum was taken. According to a study by Graddis et al. (1988), at these temperatures and relative concentrations of protein and L-Trp, heterodimer formation should occur very rapidly (i.e., <3 min at 60 °C).

NMR Spectroscopy. NMR samples were 1.5–2.5 mM in protein monomer concentration and 4.5–7.5 mM in L-tryptophan in a 500 mM potassium chloride/50 mM sodium phosphate buffer. Spectra were recorded at pH 5.6 in 90%H_2O/10%D_2O buffer and at pH 5.7 and 7.6 in D_2O buffer. For samples in D_2O and for HOHAHA spectra in H_2O the solvent resonance was suppressed with presaturation. For NOESY spectra in H_2O a 200-ms mixing time was used, and the solvent peak was suppressed by use of a NOESY sequence with a jump–return (Plateau & Gueron, 1982) observe pulse combined with very low-power presaturation. Spectra were referenced to internal 3-(trimethylsilyl)tetradeuteriopropionate. All 2D data sets were recorded at 35 or 45 °C at 500 MHz on a Bruker AM-500 using TPPI (Marion & Wüthrich, 1983) for quadrature detection in F_1. Data sizes were 1K, zero filled to 2K or 4K in the F_2 dimension and 512 t_1 slices zero filled to 1K or 2K in F_1. The data were processed on the Bruker Aspect 3000 using sine bell or sine bell squared window functions shifted by $\pi/3$ or $\pi/4$. A base-line correction was applied in F_2 in some cases for NOESY spectra in H_2O.

Secondary Structure Calculations. The secondary structure of individual peptide sequences was calculated with the double-iterated Kalman filter method (Altman & Jardetzky, 1989). This method uses principles of Bayesian probability to sequentially refine estimates of the mean position of the protein atoms and their associated variance and covariance, thus providing a set of atomic coordinates consistent with the applied constraints, as well as an explicit quantification of the uncertainty in atomic position for any segment of a protein. Five calculations were performed involving residues 24–42, 43–65, 66–76, 77–93, and 94–105 with full atomic representation and with XH_n groups represented as pseudoatoms (Wüthrich et al., 1983). A total of 156, 193, 81, 113, and 107 atoms/pseudoatoms were used for regions 1–5, respectively, with initial random mean positions having a variance of 100 Å^2. These coordinates were subsequently refined with all of the available experimental NOE information, as well as the covalent distance and stereochemical data.

Covalent chemical distance information of bond lengths and all possible distances implied by bond angles were used, i.e., a total of 382, 471, 197, 271, and 278 covalent constraints for regions 1–5, respectively. A variance of 0.1 Å^2 was assigned to covalent distances in the backbone while a larger uncertainty (up to 0.5 Å^2) was assumed for distances between pseudoatoms within side chains. Experimental NMR distance information consists of the sequential and medium-range NOEs reported in Figure 3. A total of 42, 47, 8, 34, and 34 NOE distance constraints were used for calculations 1–5, respectively. The N^i–N^{i+1} distances were assigned a mean value of 3.1 with a variance of 0.5 Å^2. A mean value of 4.0–4.6 and a variance of 1.1 Å^2 were used for the remaining distances in order to account for the pseudoatom representation and the inherent uncertainty in distance evaluations from NOE data. Dihedral (ϕ, ψ) angles were introduced, allowing the full Ramachandran range of angles by use of values of (−90°, 60°) with large variances (999°^{2}, 8000°^{2}) instead of the helical values. In

addition, dihedral angles with means of 180° and variances of 25°[2] were used for peptide bonds and 0°/180° for all planar aromatic residues. All computations of the double-iterated Kalman filter procedure were carried out with the vectorized parallelized program PROTEAN-Part II (Altman et al., 1990) on a STARDENT minisupercomputer and a Cray Y-MP.

RESULTS

A preliminary report of our assignment strategy based on selective deuteration was given by Arrowsmith et al. (1990). Here we report in detail the assignment process for the active L-tryptophan-bound *trp* repressor. It was necessary to rely heavily on NOE spectra for the assignment of many side-chain resonances to amino acid type because of the high degree of overlap and large number of missing cross-peaks in the COSY and HOHAHA spectra. Although selective deuteration was extremely useful for identifying certain COSY spin systems, it did not help resolve, for example, the 19 mostly overlapping Glx spin systems. In such cases, intraresidue NOEs from NH to side-chain protons were used to match side-chain proton chemical shifts with particular spin systems.

Assignment of Side-Chain Protons to Amino Acid Type. The four Thr residues were distinguished from the ten Ala residues by the former's relay peaks ($\alpha \rightarrow$ methyl) in relay COSY and HOHAHA spectra of GRIFSTV and GIT in D_2O and the absence of their β-methyl cross-peaks in COSY spectra of analogues in which Thr is deuterated. The five Val spin systems were identified from their well-resolved patterns in the COSY spectrum of GRIFSTV and VFT as well as the absence of these cross-peaks in spectra of all other analogues. The three Ile spin systems were deduced from their COSY cross-peaks and intraresidue NOEs in the 200-ms NOESY of GIT in D_2O. The nine Asx residues were assigned to nine complete or partial AMX spin systems which had at least moderate intensity in all analogues but were strongest in COSY spectra of SHAGED. COSY spin systems of three of the five Gly residues could be identified. The α,α' protons of Gly 64 have the same chemical shift (see below) and therefore have no COSY cross-peak. In COSY spectra of natural abundance repressor, five of the six Ser spin systems are observable, and an additional AMX system is seen in the HOHAHA of GIT. None of these resonances gave rise to NOEs to amide protons, however. The Ser and Asx AMX spin systems were distinguished from those of the aromatic residues because the aromatic α and β protons have strong NOEs to aromatic ring protons. The chemical shifts of the aromatic residues are essentially the same as those of the aporepressor, which have been assigned by Hyde et al. (1989). Except for Gly, the number of spin systems for each amino acid type discussed above corresponds exactly to the number of those amino acid residues in the repressor monomer. This confirms that we are dealing with a symmetric dimer in which corresponding protons in each monomer are in equivalent environments.

Except for a few residues [see Arrowsmith et al. (1989)], the remaining spin systems, Leu (19), Glx (19), Arg (9), Pro (4), Met (3), and Lys (3), had either incomplete or unresolved COSY spin systems, and therefore, assignments for these residue types were based on intraresidue NOEs from side-chain protons to the amide proton. The identification of 16 of the Glx side chains from intraresidue NOEs is shown in Figure 1.

Sequence-Specific Assignments. The first step in the sequential assignment process was the identification of short stretches of sequential NH^i–NH^{i+1} NOEs in this highly α-helical protein (Lane & Jardetzky, 1985; Schevitz et al., 1985).

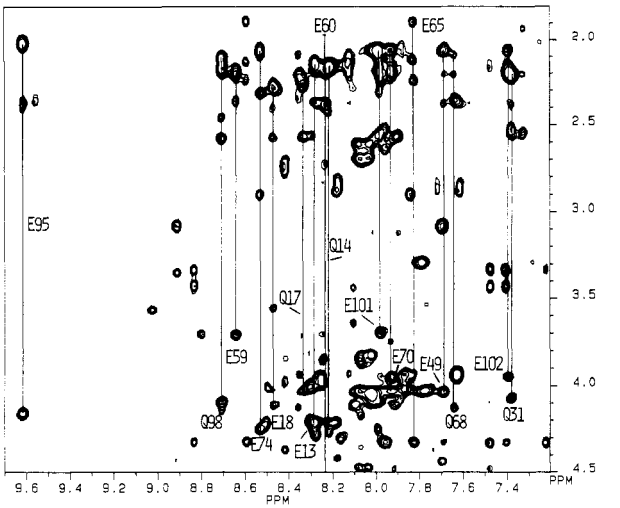

FIGURE 1: Portion of the 500-MHz NOESY spectrum of selectively deuterated analogue SHAGED in 90% H_2O/10% D_2O buffer at 45 °C and a mixing time of 200 ms. 16 of the 19 Glx residues are indicated with vertical lines at the amide proton frequency and intersecting the α-, β-, and γ-proton cross-peaks.

Next, intraresidue NOEs from each amide proton are identified, and the amino acid is classified as to residue type by matching the side-chain resonances with a COSY/HOHAHA spin system, when possible, or by matching to random coil chemical shifts. Most amide protons also have NOEs to the side-chain protons of the preceding residue and/or the α proton of the preceding third residue if these residues are protonated. Intraresidue NOEs may be distinguished from NOEs to the preceding residue because for protonated residues the former are more intense (Pachter et al., unpublished data) even within longer side chains such as Leu and Ile. The weaker NOEs from amide protons to the preceding residue help establish the directionality of the NH–NH connectivities. Since all the amide protons remain protonated in all the deuterated analogues, the NH–NH NOESY cross-peak region is the same for all analogues except for variations in intensity of the cross-peaks with deuteration, a given cross-peak having a greater intensity when one or both of the residues are deuterated (see Figure 3). This variation in cross-peak intensity, combined with a knowledge of the isotopic composition of each analogue, was used as confirmatory evidence for the assignment of an amide proton to residue type. Weak NH^i–NH^{i+2} NOEs characteristic of helical structures were also observed and helped confirm the primary amide–amide connectivities. Once a string of NH–NH connectivities is established and residue types are assigned, the sequence of the string is matched to the appropriate section of the protein sequence (Gunsalus & Yanofsky, 1980). This process is illustrated in Figure 2 for residues 53–59 in the NOESY spectrum of VFT. Amino acid types whose spin systems could be identified in COSY spectra (Ala, Asx, Val, Ile, Thr) and aromatic residues were used as points of reference within a sequence of NH–NH connectivities. Sequence-specific assignments for the *trp* repressor monomer are given in Table II and are discussed in detail below.

Residues 2–7. The side-chain resonances of residues 2–7 have been reported (Arrowsmith et al., 1989), and the amide protons of these residues were assigned in this study from HOHAHA spectra of the natural abundance repressor in H_2O at 44- and 66-ms mixing times. In these spectra, correlations between amide protons and side-chain protons are only observed for approximately 20% of the expected residues, and these appear to be from the more mobile residues in the

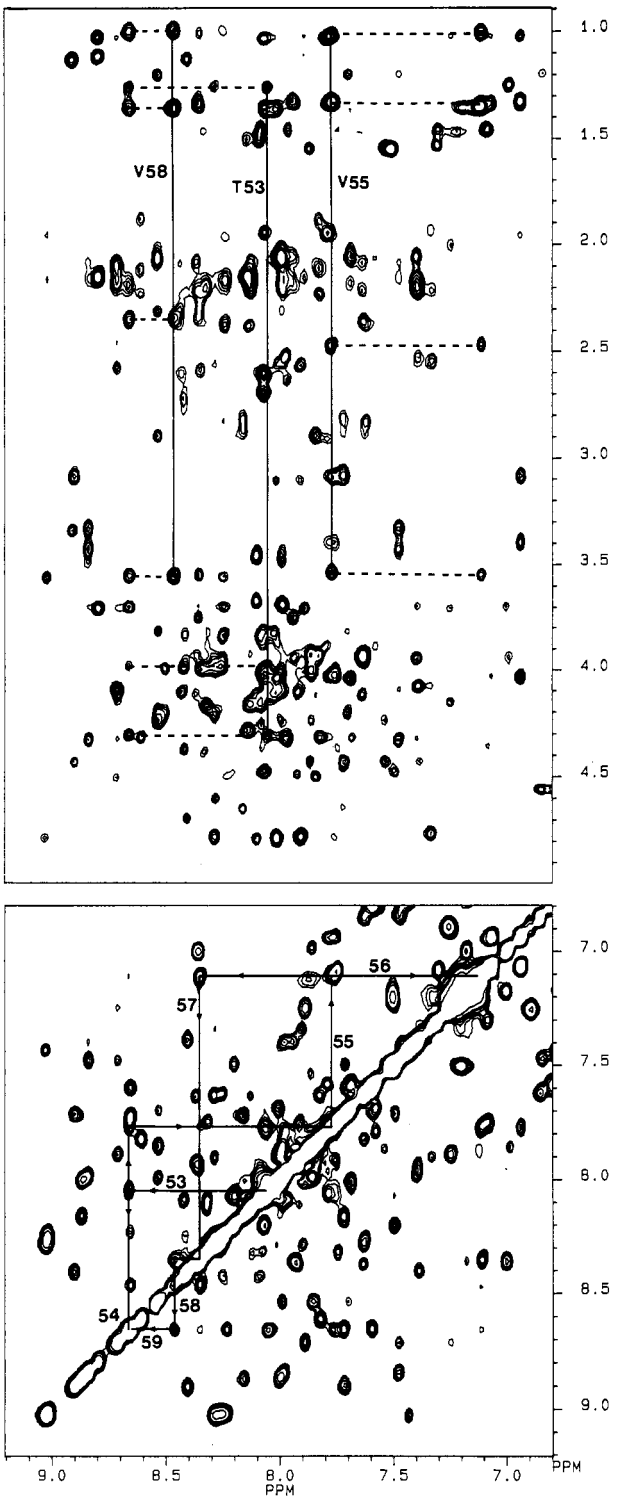

FIGURE 2: Amide cross-peak region of the 500-MHz NOESY spectrum of selectively deuterated analogue VFT in 90% H_2O/10% D_2O at 45 °C and a mixing time of 200 ms. Sequential NH^i–NH^{i+1} connectivities for residues 53–59 are indicated in the lower section, and intraresidue NOEs from the amides of Thr53, Val55, and Val58 are connected by vertical lines in the top section. Dotted horizontal lines show the connectivities from side chain (i) to NH (i+1).

protein. Thus, cross-peaks from the side-chain resonances of the mobile N-terminal "arm" to the respective amide protons can be readily identified as some of the strongest cross-peaks in this spectrum. However, no correlation is observed for the N-terminal Ala residue due to fast exchange of the amino protons. The encoded Met1 is removed in vivo (Gunsalus & Yanofsky, 1980).

Residues 12–14. There is a single NH–NH NOE between an Ala and Glx residue as well as interresidue side-chain–NH NOEs from this Ala to the Glx. This dipeptide was assigned to Ala12-Glu13 because it is the only Ala-Glx pair in the sequence. Glu14 was assigned to the only remaining set of Glx intraresidue cross-peaks after all others had been assigned (see Figure 1).

Residues 16–20. The side-chain protons of His16 and Trp19 have been assigned for the aporepressor (Hyde et al., 1989), and the chemical shifts of these residues remain the same upon the binding of L-tryptophan. NOEs are observed from α, β, and H4 protons of His16 to an amide proton assigned to His16. These intraresidue NOEs are weak and only observable at 35 °C in the NOESY spectra of HLW and SHAGED. NOEs from the α and β protons of Trp 19 have strong NOEs to two NH protons that are connected by an NOE. One of the NHs has no further NH connectivities; the other connects to the NHs of two successive Glx residues. Thus, this string of NHs is assigned to Gln17-Glu18-Trp19 and Leu20. Residues 17 and 18 have intraresidue NOEs from their side chains to NHs in spectra of all analogues, and Leu20 has NOEs from its NH to its α, γ, and δ protons in spectra of HLW. The chemical shifts of the amide protons of residues 16 and 17 are too similar in value to produce a resolved NH–NH cross-peak.

Residues 22–36. The side-chain resonances of the unique Phe22 are the same as those reported for the aporepressor (Hyde et al., 1989). The NH of Phe22 was identified by intraresidue NOEs to the side-chain protons. A string of NH^i–NH^{i+1} connectivities starting with Val23 [assigned in Hyde et al. (1989)] and ending 13 residues later defines the amide backbone of residues 23–36. The chemical shifts of the NHs of Phe22 and Val23 are too close in value to result in a resolved NOESY cross-peak. This connectivity chain stops at NH36 because residue 37 is a Pro and has no amide proton. Strong, well-resolved intraresidue side-chain–amide NOEs are observed for Asp24, Asn28, Ala29, Glu31, Asn32, and Asp33 in all analogues and those for Val23 were observed in NOESYs of VFT and GRIFSTV. Each of these residues also has NOEs from their α and/or β protons to the NH of the next residue in the sequence. Residues 23, 24, 28, 29, 32, and 33 all have resolved COSY spin systems. The α- and δ,δ'-methyl protons of Leu26, Leu34, and Leu36 were identified by unique intraresidue NOEs to their respective amide protons in spectra of HLW. A weak intraresidue NOE from residual α and β protons of Tyr30 was observed in the NOESY spectrum of GMAPS. Lysine residues are deuterated in all analogues, and thus, no side-chain assignments were identified for Lys27. Weak intraresidue NH–side-chain NOEs are observed for His35 in NOESY spectra of SHAGED and HLW. In addition to the intraresidue NOEs used to assign each residue, at least one or more sequential NOEs were identified for each residue as shown in Figure 3.

Residues 38–43. A string of NH–NH connectivities was identified involving six residues and in which several of the NH–NH cross-peaks are much weaker in spectra of HLW compared to other analogues, indicating that this sequence is rich in leucines. On the basis of intraresidue NOEs to the amides, this sequence was assigned to Leu38-Leu39-Asn40-Leu41-Met42-Leu43. Intraresidue side-chain–NH NOEs for Leu39, Asn40, and Met42 are well resolved in spectra in which these residues are protonated. The side-chain protons of Asn40 match a COSY spin system that has been assigned to Asx, and those of Met42 match a partial spin system assigned to a Met in the COSY spectrum of GMAPS.

Residues 44–52. Thr44 was identified from the missing Thr β–γCH_3 cross-peak in the COSY spectrum of the mutant Thr → Met44 at natural abundance. The α proton of Thr44 was

Table II: ^{1}H Chemical Shifts (ppm) of *trp* Repressor at 47 °C and pH 5.7[a]

no.	amino acid	α	β	β′	γ	γ′	δ	δ′	NH	other
2	Ala	4.16	1.56							
3	Gln	4.42	2.02	2.12	2.39				8.43	
4	Gln	4.42	2.02	2.12	2.39				8.43	
5	Ser	4.80	3.88	3.95					8.34	
6	Pro	4.40	2.20	1.72	1.95	1.82	3.64	3.75		
7	Tyr	4.57	2.97	3.33					7.91	7.12 (2,6), 6.84 (3,5)
8	Ser									
9	Ala									
10	Ala									
11	Met									
12	Ala	4.22	1.46						8.19	
13	Glu	4.22	2.12	2.17	2.38				8.29	
14	Gln	4.22	2.15	2.17	2.38	2.42			8.22	
15	Arg									
16	His	4.70	3.45	3.58					8.34	8.50 (H2), 7.35 (H4)
17	Gln	4.00	2.24	2.37	2.57				8.34	
18	Glu	4.11	2.29	2.38	2.59				8.43	
19	Trp	4.79	3.47	3.65					8.09	7.08 (H2), 9.98 (NH), 7.78 (H4), 6.94 (H5), 7.06 (H6), 7.33 (H7)
20	Leu	3.00			1.60		0.78	0.85	7.99	
21	Arg									
22	Phe	4.03	3.40	3.07					7.78	6.95 (2,6), 7.08 (3,4,5)
23	Val	2.92	1.93		0.03	1.03			7.78	
24	Asp	4.47	2.61	2.70					8.05	
25	Leu								8.10	
26	Leu	3.82					0.63[b]	0.78[b]	8.22	
27	Lys								7.49	
28	Asn	4.42	3.08						7.71	7.28, 6.93 (CONH_2)[b]
29	Ala	3.35	1.13						8.89	
30	Tyr	4.69	2.64	2.67					8.40	6.39 (2,6), 6.28 (3,5)
31	Gln	4.08	2.05	2.15	2.20				7.40	
32	Asn	4.65	2.55						7.32	
33	Asp	4.78	2.57	3.08					7.88	
34	Leu	4.25					0.02	0.58	8.28	
35	His	3.58	1.67	1.82					9.02	8.01 (H2), 6.35 (H4)
36	Leu	4.24	1.70		1.30		0.98		7.42	
37	Pro									
38	Leu	4.33							7.97	
39	Leu	4.20	2.49		1.92		0.96	1.23	8.89	
40	Asn	4.64	2.80	2.87					8.18	
41	Leu						0.95	1.00	7.69	
42	Met	4.62	2.11	1.98	3.00	2.74			8.64	
43	Leu	4.80					0.90		7.64	
44	Thr	5.19	4.72		1.46				9.50	
45	Pro	4.75	2.35		1.63	1.37	3.30			
46	Asp	4.32	2.51	2.62					7.95	
47	Glu								7.80	
48	Arg								7.57	
49	Glu	4.03	2.05	2.18	2.37				7.70	
50	Ala	4.05	1.36						8.04	
51	Leu	3.93			1.60		0.50	0.65	8.10	
52	Gly	3.95							8.13	4.5 (α′)
53	Thr	3.97	4.32		1.25				8.03	
54	Arg								8.63	
55	Val	3.56	2.48		1.32	1.03			7.78	
56	Arg								7.08	
57	Ile	3.56	2.32		1.26	1.56	0.92		8.35	1.09 (γ-Me)
58	Val	3.57	2.35		1.36	0.98			8.46	
59	Glu	3.70	2.13	2.19	2.39				8.65	
60	Glu	4.51	1.97	2.17	2.39	2.72			8.25	
61	Leu	4.20	1.98		1.33		0.75	0.82	9.00	
62	Leu								8.28	
63	Arg								8.24	
64	Gly	3.94							7.63	
65	Glu	4.32	1.87		2.10	2.20			7.82	
66	Met								8.60	
67	Ser								8.13	
68	Gln	4.13	2.09	2.21	2.37				7.65	
69	Arg								8.37	
70	Glu	4.03	2.07	2.18	2.37				7.93	
71	Leu									
72	Lys									
73	Asn	4.51	2.90						7.83	6.86, 7.63 (CONH_2)
74	Glu	4.25	2.02	2.09	2.32				8.54	
75	Leu								7.99	
76	Gly	3.94							7.87	4.03 (α′)
77	Ala	4.59	1.25						7.01	

Table II (Continued)

no.	amino acid	α	β	β′	γ	γ′	δ	δ′	NH	other
78	Gly	4.01							8.29	4.21 (α′)
79	Ile	3.60	1.93		1.62	0.98	0.76		8.29	0.87 (γ-Me)
80	Ala	4.25	1.51						8.08	
81	Thr	4.24	4.43		1.55				7.58	
82	Ile	3.86	1.88		1.46	1.36	0.82		7.90	0.91 (γ-Me)
83	Thr	4.18	4.39		1.20				8.50	
84	Arg								7.68	
85	Gly								7.61	
86	Ser								8.24	
87	Asn	4.36	2.76	2.72					8.40	6.90, 7.54 ($CONH_2$)
88	Ser								8.06	
89	Leu	4.03			1.65	0.93			8.30	
90	Lys								7.74	
91	Ala	4.37	1.45						7.09	
92	Ala	4.55	1.54						7.31	
93	Pro	4.65	2.03	2.11	2.55		3.70			
94	Val	3.72	2.16		1.01	1.12			8.80	
95	Glu	4.17	2.01	2.17	2.38				9.62	
96	Leu						1.00	1.37	7.24	
97	Arg								7.89	
98	Gln	4.10	2.10	2.18	2.58				8.70	
99	Trp	4.36	3.34	3.43					7.48	7.23 (H2), 9.65 (NH), 7.41 (H4), 6.66 (H5), 7.13 (H6), 7.04 (H7)
100	Leu	3.52	2.05	2.12	1.12		0.76	0.86	8.83	
101	Glu	3.69	2.05	2.07					7.99	
102	Glu	3.94	2.05	2.07					7.41	
103	Val	3.75	1.31	−0.08	0.31				7.93	
104	Leu						0.78	0.85	8.35	
105	Leu						1.01	1.15	7.12	
106	Lys	4.37	1.88		1.40		1.63		7.19	2.95 (ε)
107	Ser									
108	Asp	4.42	2.58	2.68					7.92	
	L-Trp	4.02	3.28	3.43						7.32 (H2), 10.04 (NH), 7.80 (H4), 7.17 (H5), 7.26 (H6), 7.53 (H7)

[a] ±0.03 ppm. [b] Chemical shift at 35 °C.

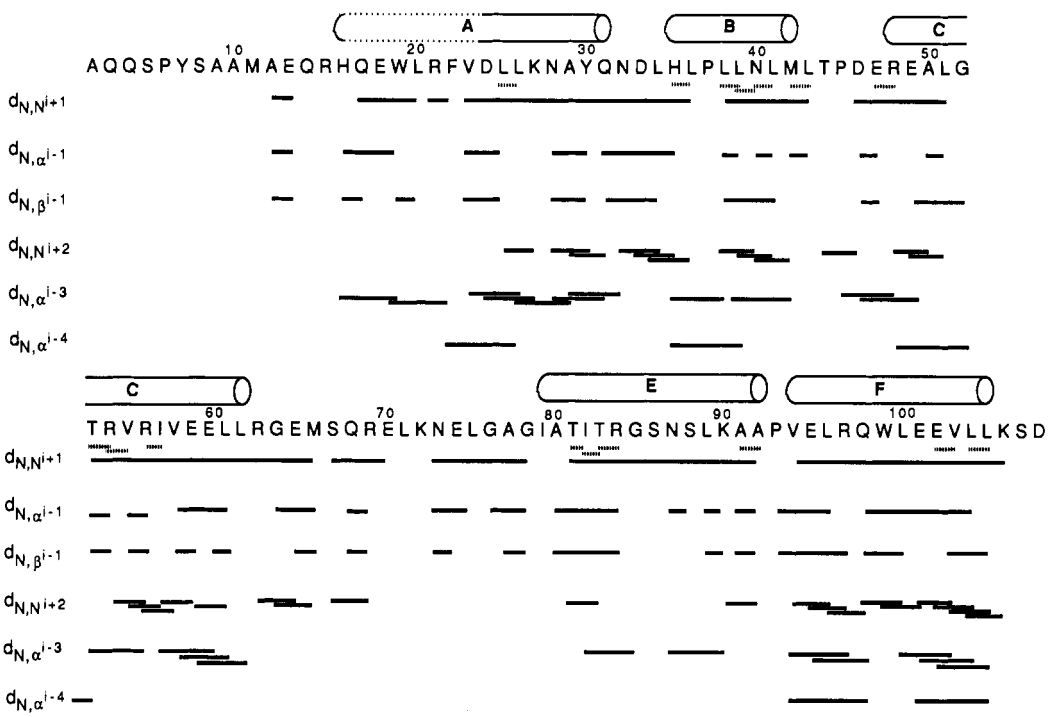

FIGURE 3: Amino acid sequence of *trp* repressor with the numbering reported by Gunsalus and Yanofsky (1980) for the nucleotide sequence. Short- and medium-range NOEs are represented by solid horizontal bars between residues. The broken bars indicate NH^i–NH^{i+1} NOEs that had greater intensity in spectra of analogues in which one or both of the residues were deuterated. Helical regions determined in this work are indicated by cylinders across the top and are labeled A–F, corresponding to the analogous regions in the crystal structure (Schevitz et al., 1985). The secondary structure of the first half of helix A was not specifically calculated for reasons of computational ease; however, the NOEs between αHis16 and NHTrp19 and from αTrp19 to NHPhe22 indicate that it is probably helical.

identified from the intraresidue α–β NOE in the NOESY of natural abundance wild-type *trp* repressor. At 5.19 ppm this proton is the only α proton with a chemical shift greater than 4.8 ppm and can therefore be considered unusual in this protein. The amide proton of 44 was assigned to the only NH that had NOEs to the α and γCH_3 protons of Thr44. This NH also has NOEs to side-chain protons consistent with a Pro residue—Pro45. These NH44–Pro45 NOEs were strongest

in spectra of GMAPS, but weak cross-peaks also appeared in spectra of SHAGED and VFT. The side-chain protons of Pro45 also have NOEs to two NHs which are part of a string of NH–NH connectivities beginning with an Asx residue and ending four residues later with an Ala. The Pro45 NOEs are to the first and third NHs of this string. Thus, this string of connectivities was assigned to Asp46, Glu47, Arg48, Glu49, and Ala50. Intraresidue NH–side-chain NOEs are observed for residues 46, 49, and 50, and several sequential NOEs tie this sequence together. The α and βCH_3 protons of Ala50 have strong NOEs to an NH that has intraresidue NOEs to two upfield-shifted methyl groups (0.65, 0.50 ppm) in spectra of HLW. This NH was assigned to Leu51. The methyl groups of Leu51 have weak NOEs to an NH with strong NOEs to two α protons that were assigned to Gly52. The NHs of residues 50, 51, and 52 all have chemical shifts with similar values so that any NH–NH cross-peaks between these residues would be lost in the diagonal. However, a weak NH^i–NH^{i+2} NOE between residues 49 and 51 links NH51 to the 46–50 string of connectivities. The side-chain resonances of residues 46, 49, 50, 51, and 52 all match unique COSY cross-peaks.

Residues 53–62. A string of 10 NH–NH connectivities that involves a sequence of Thr-X-Val-X-Ile-Val-X-X-Leu-X was identified from intraresidue NOEs to the NHs (X = initially unknown amino acid). This unique sequence must be residues 53–62. The chemical shift of NH53 is too close to that of 52 for a possible NOE to be resolved from the diagonal. Thr53, Val55, Ile57, Val58, and Leu61 all have strong, resolved intraresidue NOEs from NH to side-chain protons in spectra of analogues in which these residues are protonated. The assignment of the NHs of Arg54 and Arg56 were deduced from their positions in the sequence. Intraresidue NOEs from Glu59 and Glu60 were resolved in NOESY spectra of SHAGED. Five NH^i–NH^{i+2} NOEs and four NH^i–α^{i-3} NOEs within this sequence also help confirm the connectivities.

Residues 63–66. A short string of NH–NH connectivities was assigned to residues Arg63-Gly64-Glu65-Met66 on the basis of the Gly-Glx dipeptide. Gly64 is the only Gly in the *trp* repressor sequence that is adjacent to a Glx residue. In addition, the other four Gly residues can be uniquely placed in other strings of NH–NH connectivities. The side-chain–NH NOEs of the Gly and Glu residues are strong and well resolved in all analogues. At the frequency of NH64 a single very strong NOE is observed to an α-proton frequency, indicating the α protons of Gly64 have the same chemical shift.

Residues 67–70. A string of three NH–NH connectivities was identified involving an X-Glx-X-Glx peptide. The only two regions of *trp* repressor that fit this sequence are residues 46–49 and 67–70. The former can be assigned to another set of resonances (see above) on the basis of the Asp46 intraresidue cross-peaks. The resonances in question were therefore assigned to residues 67–70. Further support for this assignment is a very weak cross-peak between the NH of Met66 and the NH of Ser67 that appears in NOESY spectra of only a few analogues. The intraresidue NH–side-chain cross-peaks for Gln68 and Glu70 are well resolved, as shown in Figure 1.

Residues 73–79. A strong set of sequential NOEs [NH^i–$(NH,\alpha,\beta)^{i-1}$] from an Asx-Glx dipeptide are visible in NOESY spectra of all analogues. This dipeptide was assigned to residues 73 and 74 because the only other Asx-Glx dipeptide is 46-47 and the latter are assigned to other resonances. The NH–NH connectivities extend from Glu74 through the next three residues in spectra of all analogues. In the NOESY spectrum of SHAGED a weak NOE is also observed between the NH of Ala77 and the NH of Gly78. Strong intraresidue NOEs were observed for Asn73, Glu74, Gly76, Ala77, and Gly78, and the chemical shifts of the side-chain protons of each of these residues match a COSY spin system of the appropriate type. The α,α' protons of Gly78 have NOEs to an NH which has intraresidue NOEs to an Ile side chain, Ile79. No NH–NH NOE was observed between Gly78 and Ile79.

Residues 80–92. A long stretch of NH–NH connectivities was identified that started with a Thr-Ile-Thr tripeptide and ends with an Ala-Ala dipeptide. This must correspond to residues 81–92. Strong well-resolved intraresidue NOEs are observed for Thr81, Ile82, and Thr83 in spectra of VFT and GRIFSTV and for Asn87, Ala91, and Ala92 in spectra of all analogues. The side-chain protons of residues 81, 82, 83, 87, 91, and 92 also match an appropriate COSY spin system. Intraresidue NOEs to the α, γ, and δ protons but not to βs are observed from the NH of Leu89. The α and β protons of an Ala residue have NOEs to the NH of Thr81. The strongest NOEs from these Ala α and β protons are to an NH at 8.08 ppm, and therefore, this NH was assigned to Ala80, although there is no NH–NH connectivity to the NH of Thr81.

Residues 93–106 and 108. A string of NH^i–NH^{i+1} connectivities was identified involving 13 residues, starting with a well-resolved Val, connecting five residues later to a Trp and nine residues later to another Val. The chemical shifts of the Trp α and β protons and the side-chain protons of the second Val matched those of Trp99 and Val103, respectively, reported by Hyde et al. (1989). Thus, this string of NH–NH connectivities defines the amide protons of Val94 through Lys106. Strong, well-resolved intraresidue NOEs from NH to side-chain protons were observed for Val94, Glu95, Gln98, Trp99, Leu100, and Val103 in NOESY spectra of analogues in which these residues were protonated. Only NOEs from the methyl protons to amide protons were resolved for Leu96, Leu104, and Leu105 in spectra of HLW. For Glu101 and Glu102 only NOEs from amide to α and β protons were resolved. Of residues 94–105 only the two Vals and Trp99 have spin systems that could be completely identified in the COSY spectrum. The sequential NOEs, shown in Figure 3, are further evidence for the correct assignment of the residues in this sequence. Residue 93 is a Pro and therefore has no amide proton; however, in the NOESY spectrum of GMAPS weak NOEs from NH94 to side-chain protons with chemical shifts consistent with a Pro spin system were used to assign Pro93. The final NH in the connectivity chain was assigned to Lys106. The side chain of Lys106 was assigned to the only Lys spin system that could be identified in the COSY spectrum of natural abundance repressor. It was assumed that of the four Lys residues only the relatively mobile Lys106 near the carboxy terminal would have lines narrow enough to give rise to a complete COSY spin system. Similarly, the N-terminal Asp108 was assigned to the only Asx spin system that did not have intraresidue NOEs but did have HOHAHA cross-peaks from amide to side-chain protons.

L-Tryptophan Ligand. The resonances of the L-Trp ligand were easily identified by comparing NOESY spectra of analogues with natural abundance L-Trp to those with L-Trp-d_5 (indole ring deuterated). Intramolecular cross-peaks within the bound tryptophan are very strong and quite broad due to exchange broadening with the free state. NOEs from L-Trp to repressor protons are discussed elsewhere (Altman et al., unpublished results).

Assignment of Inter- and Intrasubunit NOEs. Since *trp* repressor is a symmetric dimer, one cannot a priori distinguish interresidue NOEs within the same monomer from those be-

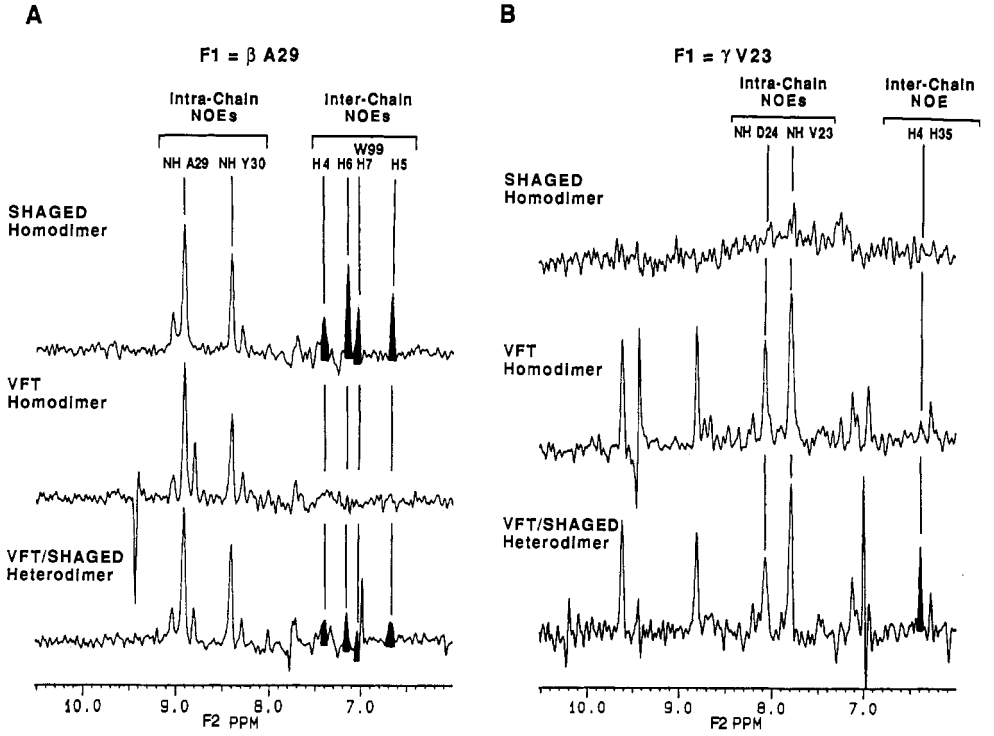

FIGURE 4: F_1 slices of the 500-MHz NOESY spectra (45 °C, mixing time of 200 ms) of SHAGED, VFT, and a heterodimer between these two analogues. (A) Slices at the frequency of the β-methyl protons of Ala29. (B) Slices at the frequency of the γ-methyl protons of Val23. Intersubunit NOEs are shaded. Each slice was scaled so that the indicated intrachain NOEs were of equal intensity. For a detailed explanation, see the text.

tween different monomers. It is reasonable to assume that NOEs between adjacent residues or residues separated by three or four amino acids in the sequence are intrasubunit NOEs. This does not hold for longer range NOEs. Intersubunit NOEs were identified by making three heterodimers between two differently deuterated repressors. For example, a heterodimer was made between analogues VFT in which Ala is protonated and Trp is deuterated and SHAGED in which both residues are protonated. By mixing equimolar amounts of each selectively deuterated homodimer, one generates a 1:2:1 mixture of VV/VS/SS dimers. Figure 4A shows slices taken at the frequency of the Ala29 methyl protons in the NOESY spectra of the two homodimers and the heterodimer mixture. All three slices are scaled so that the intrachain NOEs which are present in all three species (β29–NH29 and β29–NH30) are of equal intensity in each spectrum. In the NOESY of SHAGED homodimer the Ala29–Trp99 NOEs are readily observed, and two are of intensity comparable to the that of intrasubunit NOEs. In the NOESY of VFT homodimer no Ala29–Trp99 NOEs are observed because Trp is deuterated in both subunits. In the NOESY of the VFT/SHAGED mixture the Ala29–Trp99 NOEs have decreased to approximately one-fourth of their size in the SHAGED homodimer spectrum. This proves that this NOE is between two different subunits, the remaining NOE intensity arising from the SHAGED homodimer species in the mixture. Similarly, Figure 4B shows the appearance of an intersubunit NOE between Val23 and His35 in the VFT/SHAGED heterodimer. Since one of the residues involved in this NOE is deuterated in each homodimer, the presence of this NOE in the heterodimer mixture must be from Val23 in VFT to His35 in SHAGED. Table III summarizes the classification of eight other long-range NOEs from this and two other heterodimers.

Secondary Structure Calculations. Computations converged from initial average (maximal) errors of ca. 100 (500) to approximately 0.3 (5–7) standard deviation within 50–200

Table III: Summary of Long-Range NOEs Based on Hybridization Experiments

NOE	amino acid present in homodimer: SHAGED	VFT	rel intensity in heterodimer mixture[a]: wrt SHAGED	wrt VFT	inter-/intra-subunit NOE
βAla29	+	+			
H6 Trp99	+	–	–	+	inter
γ Val23	–	+			
H4 His35	+	–	+	+	inter
β Asn28	+	+			
H7 Trp99	+	–	–	+	inter
γ Val55	–	+			
H2,6 Phe22	–	+	+	–	inter
β Ala29	+	+			
α His35	+	–	0	+	*intra*

NOE	amino acid present in homodimer: nat abund[b]	VFT	rel intensity in heterodimer mixture: wrt nat abund	wrt VFT	inter-/intra-subunit NOE
δ Leu34	+	–			
H2 Trp99	+	–	–	+	inter
δ Leu51	+	–			
H4 Trp19	+	–	–	0	inter
δ Leu51	+	–			
H3,5 Phe22	+	+	–	+	inter

NOE	amino acid present in homodimer: GIT	VFT	rel intensity in heterodimer mixture: wrt GIT	wrt VFT	inter-/intra-subunit NOE
α Ile82	+	–			
γ Ile57	+	–	0	+	*intra*
β Met42[c]	+	+			
H2,6 Phe22	–	+	0	–	*inter*

[a] (+) = increased, (–) = decreased, and (0) = unchanged relative to intrasubunit NOE. [b] Nat abund = natural abundance *trp* repressor. [c] Met is partially protonated in the preparation of GIT used for the hybridization experiment, and weak NOEs were observed in VFT homodimer between F22 and residual protons on Met42.

successive cycles for regions 1–5 of the protein. The resulting helical regions are indicated in Figure 3. Comparison of the α-carbon backbone of each helical segment with that of an ideal helix of equal length gave average RMS deviations (Å) of 2.4, 2.3, 2.5, 3.2, 2.5, and 1.7 for residues 24–31, 35–42, 48–62, 67–76, 79–92, and 94–105, respectively. Using these boundaries for helical regions of the protein gives standard deviations ranging from 2.1 to 1.3 for each helix from the corresponding positions in ideal helices. Clearly, the five A, B, C, E, and F helical structures reported for the crystal structure are reproduced here with the available NOE data, while the lack of experimental data for the D segment results in a less well-defined region.

Discussion

The combination of high molecular weight and selective deuteration creates the special conditions necessary for the assignment procedure described above. The large molecular weight of the *trp* repressor results in significant spin diffusion in NOE experiments, especially at the mixing time of 200 ms used here. Although spin diffusion is usually undesirable, in this case spin diffusion within the side chains of protonated residues results in significant NOE intensity between the NH and its side-chain protons, allowing the potential assignment of all the side-chain protons to their respective NHs. At the same time, deuteration of surrounding residues prevents loss of magnetization due to spin diffusion to nearby side chains. The assignment procedure used here should be generally applicable to α-helical proteins or regions of proteins.

In the case of *trp* repressor, sequence-specific assignments were made for 90% of the backbone NH protons and 70% of the α and side-chain resonances. Except for leucines, most of the unassigned side chains remain so due to the lack of intraresidue NOEs. This is particularly true of the nine Arg and four Ser residues, which are likely to be on the surface of the protein and therefore may have greater conformational mobility. The unassigned Leu resonances are mainly due to spectral overlap; most of the β and γ resonances could not be assigned because of overlapping cross-peaks from β and γ protons of the 19 Glx residues, which could not be completely deuterated. Nevertheless, enough resonances have been assigned to determine the secondary structure of the *trp* repressor in solution. As expected from CD data (Lane & Jardetzky, 1987) and from the large number of NH–NH NOEs, the repressor is largely α-helical.

Helical boundaries are within one to two residues of those reported for the crystal structure for helices B, E, and F. It is interesting to note that helical boundaries based on the calculations do not always correspond to those predicted from the sequential NOEs alone. For example, from the NOEs in Figure 3 one would not necessarily expect helix E to start at residue 79 because there are no NOEs from Ile79 to the NHs of 80 or 82 (these NOEs would be resolved if they were present in the spectra). Similarly, the NOE data for residues 45–47 would indicate that these residues are part of helix C; however, the calculated structure shows an incomplete turn for this region. Although the calculations for A started with residue 24 for reasons of computational ease, the NOE data indicate that helix A may start at residue 16 in solution. There is no evidence of helical structure between residues 10 and 16, as in the crystal. Helix D in the crystal structure is the first helix of the helix–turn–helix DNA binding motif and is therefore very important functionally. The sequential NOE data alone for this region (residues 67–75) do not define a helix. However, there are two potential turns or loops (residues 68–70 and 73–75) that could be part of a distorted or conformationally mobile helix. It is significant that the N-terminal region and "helix" D correspond to regions of "flexibility" in different crystal space groups of the repressor (Lawson et al., 1988). We interpret the regions of sparse NOE data in the present study (residues 2–15 and, to a lesser extent, 66–75) as indicative of conformational mobility. This has been verified experimentally for the first seven residues (Arrowsmith et al., 1989). Further relaxation studies are planned to test this hypothesis with respect to the DNA binding region.

From the spectra of three heterodimers, it was possible to classify all long-range NOEs as being intra- or intermonomer NOEs, either directly or indirectly on the basis of residues' close proximity in the secondary structure to a pair involved in a directly determined NOE. On the basis of these NOEs and 11 others to the L-Trp ligand, the tertiary structure of the protein was calculated. The details of this calculation will be reported elsewhere (Altman et al., unpublished results); however, the general topology of the solution structure is quite similar to that of the crystal, the major differences being (as for the secondary structure) in the N-terminal region and "helix" D region.

Acknowledgments

Special thanks to Dr. L. Szilágyi for help with spectroscopy, Dr. L. Treat-Clemons for preparation of protein, Drs. C. Yanofsky and B. Hurlburt for providing the overproducing *E. coli* strains and the mutant TM44, and Drs. V. Ramesh and J. Carey for fruitful discussions. We thank Stardent Computer Inc. for their generous loan of a TITAN minisupercomputer.

References

Altman, R. B., & Jardetzky, O. (1989) *Methods Enzymol. 177*, 218.

Altman, R. B., Pachter, R., Carrara, E., & Jardetzky, O. (1990) Quantum Chemistry Program Exchange.

Arrowsmith, C. H., Carey, J., Treat-Clemons, L., & Jardetzky, O. (1989) *Biochemistry 28*, 3875.

Arrowsmith, C. H., Treat-Clemons, L., Szilágyi, L., Pachter, R., & Jardetzky, O. (1990) in *Die Makromolekulare Chemie, Macromolecular Symposia* (Höcker, H., & Sedláček, B., Eds.) Vol. 34, Hüthig & Wepf Verlag, Heidelberg.

Graddis, T. J., Klig, L. S., Yanofsky, C., & Oxender, D. L. (1988) *Proteins 4*, 173.

Gunsalus, R. P., & Yanofsky, C. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7117.

Hyde, E. I., Ramesh, V., Roberts, G. C. K., Arrowsmith, C. H., Treat-Clemons, L., Klaič, B., & Jardetzky, O. (1989) *Eur. J. Biochem. 183*, 545.

Klig, L. S., Carey, J., & Yanofsky, C. (1988) *J. Mol. Biol. 202*, 769.

Lane, A. N., & Jardetzky, O. (1985) *Eur. J. Biochem. 152*, 405.

Lane, A. N., & Jardetzky, O. (1987) *Eur. J. Biochem. 164*, 389.

Lawson, C. L., Zhang, R.-G., Schevitz, R. W., Otwinowski, Z., Joachimiak, A., & Sigler, P. B. (1988) *Proteins 3*, 18.

Lefèvre, J.-F., Lane, A. N., & Jardetzky, O. (1987) *Biochemistry 26*, 5076.

Madrid, M., & Jardetzky, O. (1988) *Biochim. Biophys. Acta 953*, 61.

Madrid, M., Mace, J. E., & Jardetzky, O. (1989) *J. Magn. Reson. 83*, 267.

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

Markley, J. L., Putter, I., & Jardetzky, O. (1968) *Science 161*, 1249.
Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321.
Paluh, J. L., & Yanofsky, C. (1986) *Nucleic Acids Res. 14*, 7851.
Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310.
Schevitz, R., Otwinowski, Z., Joachimiak, A., Lawson, C. L., & Sigler, P. B. (1985) *Nature 317*, 782.
Szilágyi, L., & Jardetzky, O. (1989) *J. Magn. Reson. 83*, 441.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *Mol. Biol. 169*, 949.
Zhang, R.-G., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., & Sigler, P. B. (1987) *Nature 327*, 591.

Articles

Assignment of the Backbone ^1H and ^{15}N NMR Resonances of Bacteriophage T4 Lysozyme†

Lawrence P. McIntosh,*,‡ A. Joshua Wand,§ David F. Lowry,∥ Alfred G. Redfield,∥ and Frederick W. Dahlquist*,‡

Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, Oregon 97403, Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, and Departments of Biochemistry and Physics, Brandeis University, Waltham, Massachusetts 02254

Received January 22, 1990; Revised Manuscript Received March 23, 1990

ABSTRACT: The proton and nitrogen (^{15}NH–H^{α}–H^{β}) resonances of bacteriophage T4 lysozyme were assigned by ^{15}N-aided ^1H NMR. The assignments were directed from the backbone amide ^1H–^{15}N nuclei, with the heteronuclear single-multiple-quantum coherence (HSMQC) spectrum of uniformly ^{15}N enriched protein serving as the master template for this work. The main-chain amide ^1H–^{15}N resonances and H^{α} resonances were resolved and classified into 18 amino acid types by using HMQC and ^{15}N-edited COSY measurements, respectively, of T4 lysozymes selectively enriched with one or more of α-^{15}N-labeled Ala, Arg, Asn, Asp, Gly, Gln, Glu, Ile, Leu, Lys, Met, Phe, Ser, Thr, Trp, Tyr, or Val. The heteronuclear spectra were complemented by proton DQF-COSY and TOCSY spectra of unlabeled protein in H_2O and D_2O buffers, from which the H^{β} resonances of many residues were identified. The NOE cross peaks to almost every amide proton were resolved in ^{15}N-edited NOESY spectra of the selectively ^{15}N enriched protein samples. Residue specific assignments were determined by using NOE connectivities between protons in the ^{15}NH–H^{α}–H^{β} spin systems of known amino acid type. Additional assignments of the aromatic proton resonances were obtained from ^1H NMR spectra of unlabeled and selectively deuterated protein samples. The secondary structure of T4 lysozyme indicated from a qualitative analysis of the NOESY data is consistent with the crystallographic model of the protein.

T4 lysozyme is an endomuramidase required for the lytic growth of bacteriophage T4 (Tsugita & Ikeya-Ocada, 1972). From the time of the pioneering work of Streisinger and colleagues, this enzyme has been the subject of extensive genetic, thermodynamic, spectroscopic, and crystallographic studies and stands today as one of the best characterized systems for investigating the basis of protein structure and stability (Streisinger et al., 1961; Alber et al., 1987; Becktel & Baase, 1987; Hudson et al., 1987; Matthews, 1987; Weaver & Matthews, 1987). In earlier work, we have used NMR[1] to study selected regions of wild-type and variant T4 lysozymes, focusing on questions including the kinetics of hydrogen exchange and the pH dependence of the electrostatic stability of these proteins (Griffey et al., 1985a; Weaver et al., 1989; Anderson et al., 1990). However, more extensive investigations

† This work was supported by NSF Grant DMB8905322 and a grant from the Lucille B. Markey Chartible Trust to F.W.D., by NIH Grant GM20168 to A.G.R., by NIH Grant GM35940 to A.J.W., and by NIH Grants CA06927 and RR05539, an award from the Fanny Ripple Foundation, a grant from the Pew Trust, and an appropriation from the Commonwealth of Pennsylvania awarded to the Institute for Cancer Research. L.P.M. is a recipient of a Natural Sciences and Engineering Council of Canada 1967 Scholarship and an Alberta Heritage Foundation Studentship.

* Authors to whom requests for reprints should be addressed.
‡ University of Oregon.
§ Fox Chance Cancer Center.
∥ Brandeis University.

[1] Abbreviations: The amino acids are denoted by standard one- and three-letter codes, Asx is both aspartate and asparagine, and Glx is both glutamate and glutamine. Atoms in amino acids are named according to the IUPAC–IUB convention [(1970) *J. Mol. Biol. 52*, 1–17]; COSY, two-dimensional *J*-correlated spectroscopy; DQF, double-quantum-filtered; DSS, sodium 2,2-dimethyl-2-silapentane-5-sulfonate; HMBC, heteronuclear multiple-bond coherence spectroscopy; HMQC, heteronuclear multiple-quantum coherence spectroscopy; HSMQC, heteronuclear single-multiple-quantum coherence spectroscopy; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE-correlated spectroscopy; ppm, parts per million; pH*, observed pH meter reading without corrections for the deuterium isotope effect; TPPI, time proportional phase incrementation; TOCSY, two-dimensional total correlation spectroscopy.

0006-2960/90/0429-6341$02.50/0