Journal of Biomolecular NMR, **11:** 307–318, 1998.
KLUWER/ESCOM

Subunit-specific backbone NMR assignments of a 64 kDa *trp* repressor/DNA complex: A role for N-terminal residues in tandem binding

Xi Shan[a], Kevin H. Gardner[b], D.R. Muhandiram[b], Lewis E. Kay[b] & Cheryl H. Arrowsmith[a,*]
[a]*Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, 610 University Avenue, Toronto, ON, Canada M5G 2M9* [b]*Protein Engineering Centers of Excellence and Departments of Medical Genetics, Biochemistry and Chemistry, University of Toronto, Toronto, ON, Canada M5S 1A8*

Received 5 August 1997; Accepted 25 October 1997

Key words: deuterium labelling, protein–DNA interactions, protein–protein interactions

Abstract

Deuterium decoupled, triple resonance NMR spectroscopy was used to analyze complexes of ^{2}H,^{15}N,^{13}C labelled intact and (des2–7) *trp* repressor (Δ2–7 trpR) from *E. coli* bound in tandem to an idealized 22 basepair *trp* operator DNA fragment and the corepressor 5-methyltryptophan. The DNA sequence used here binds two trpR dimers in tandem resulting in chemically nonequivalent environments for the two subunits of each dimer. Sequence- and subunit-specific NMR resonance assignments were made for backbone ^{1}HN, ^{15}N, ^{13}Cα positions in both forms of the protein and for ^{13}Cβ in the intact repressor. The differences in backbone chemical shifts between the two subunits within each dimer of Δ2–7 trpR reflect dimer-dimer contacts involving the helix-turn-helix domains and N-terminal residues consistent with a previously determined crystal structure [Lawson and Carey (1993) *Nature*, **366**, 178–182]. Comparison of the backbone chemical shifts of DNA-bound Δ2–7 trpR with those of DNA-bound intact trpR reveals significant changes for those residues involved in N-terminal-mediated interactions observed in the crystal structure. In addition, our solution NMR data contain three sets of resonances for residues 2–12 in intact trpR suggesting that the N-terminus has multiple conformations in the tandem complex. Analysis of Cα chemical shifts using a chemical shift index (CSI) modified for deuterium isotope effects has allowed a comparison of the secondary structure of intact and Δ2–7 tprR. Overall these data demonstrate that NMR backbone chemical shift data can be readily used to study specific structural details of large protein complexes.

Introduction

NMR spectroscopy has been extremely successful in recent years in determining the three-dimensional structures of DNA-binding domains of proteins and small protein/DNA complexes. This is in large part due to the fact that most sequence-specific DNA-binding domains of proteins are small (<100 residues), independently folded domains. These studies have provided a wealth of information on structural requirements for protein-DNA interactions and binding specificity. However, another important aspect of protein-DNA interactions involves the control of DNA-binding affinity by factors such as protein-protein interactions, phosphorylation or cofactor binding. These types of regulatory interactions typically require additional domains of the protein, if not the intact molecule. Due to limits on the size of proteins or complexes that can be studied routinely by solution NMR (<20 kDa) few NMR studies have been reported of intact DNA-binding proteins or larger multi-component complexes to date.

E. coli trp repressor (trpR) controls the expression of genes involved in tryptophan biosynthesis in response to cellular levels of tryptophan (Klig et al., 1988). TrpR contains 107 amino acids (numbered 2–108) and is therefore one of the smallest intact DNA-binding proteins with multiple regulatory features such as protein-protein interactions and corepressor (L-trp) binding. Genetic (Kumamoto et al., 1987)

* To whom correspondence should be addressed.

Figure 1. The DNA sequences of operators which bind tandem copies of trpR. The palindromic CTAG repeats which can bind two helix-turn-helix motifs (shaded ovals) are underlined. *TrpEDCBA* and *aroH* are natural operators. The 'crystal' DNA is the sequence used to form the cocrystal of tandemly bound trp repressors (Lawson and Carey, 1993). In the cocrystal successive duplexes abut one another (shown by a •) to form palindromic CTAG junctions in phase for repressor binding. This mode of binding resulted in crystals with a left-handed superhelix of trpR dimers covering end-to-end stacked DNA duplexes. The sequence used here for NMR is shown with a schematic of two tandem dimers with the two subunits indicated by A and B. Subunit A, but not B, shares a CTAG site with subunit A of an adjacent dimer leading to different chemical shifts for the same residues in subunits A and B of a given dimer.

and biochemical (Liu and Matthews, 1993; Czernik et al., 1994; Yang et al., 1996) evidence indicates that two trpR dimers bind in tandem to natural operators that contain at least one or two palindromic 5′-CTAG-3′ repeats at 8 bp intervals (Figure 1). A structural model of such a tandem complex was proposed based on a crystal structure in which tandemly bound trpR dimers bridged successive 16-basepair DNA duplexes by joining together the …CT-3′ sequence of one duplex with the 5′-AG... of the adjacent duplex (shown as -**CT•AG**- in Figure 1). This structure also indicated a role for the N-terminal 'arms' of trpR (residues 2–7) in tandem protein-protein interactions consistent with the fact that deletion of residues 2–7 lowers the overall binding affinity of trpR for natural operator sites by approximately 18% (Carey, 1989).

Previously we have used NMR spectroscopy to study a 64 kDa complex of two tandem trpR dimers bound to a single 22 basepair (bp) symmetric *trp* operator and the corepressor analog, 5-methyltryptophan (Shan et al., 1996). The complex itself has twofold symmetry; however, the two subunits of each trpR homodimer are not in chemically equivalent environments resulting in unique chemical shifts for most residues in subunit A and B of each dimer (Figure 1). The protein was uniformly labelled with ^{15}N and ^{13}C

Table 1. ^{1}HN, ^{15}N and ^{13}Cα chemical shifts for Δ2–7 trpR bound in tandem to DNA[a]

Residue	Subunit A			Subunit B		
	^{1}HN	^{15}N	^{13}Cα	^{1}HN	^{15}N	^{13}Cα
S8						
A9						
A10	8.35	123.1	52.54	8.27	123.0	52.57
M11			55.09	8.13	118.8	55.41
A12	7.91	122.5	52.93	8.16	124.2	53.82
E13	8.41		57.57	8.44	119.1	57.76
Q14	8.83	122.5	57.17	8.28	120.9	57.65
R15	8.53	119.2	57.40			
- - - -						
H16	8.24	119.5	59.35			59.00
Q17	8.47	117.5	58.31	8.35	117.9	58.22
E18	8.76	121.1	59.24	8.55	120.7	58.87
W19	8.38	122.6	59.16	8.24	122.4	59.13
L20	7.75	117.6	57.08	7.75	117.6	57.08
R21	7.88	120.7	58.99	7.88	120.1	58.99
F22	8.10	121.4	61.10	8.16	121.5	61.14
V23	7.79	119.6	66.06	7.84	119.6	66.05
D24	7.79	121.6	56.77	7.79	121.6	56.77
L25	8.22	124.3	57.35	8.22	124.3	57.35
L26	8.34	119.9	56.65	7.96	117.7	56.55
K27	7.28	120.7	59.48	7.28	120.7	59.48
N28	7.55	116.7	55.38	7.55	116.7	55.38
A29	8.97	126.8	55.24	8.97	126.8	55.24
Y30	8.38	119.2	61.26	8.38	119.2	61.26
Q31	7.29	117.1	57.46	7.29	117.1	57.46
N32	7.30	115.8	52.68	7.30	115.8	52.68
D33	7.89	117.5	55.90	7.89	117.5	55.90
L34	8.22	118.1	53.10	8.22	118.1	53.10
H35	9.05	121.1	58.80	9.05	121.1	58.80
L36	7.45	119.7	59.21	7.45	119.7	59.21
P37			64.89			64.89
L38	7.92	120.1	58.02	7.92	120.1	58.02
L39	8.90	121.3	57.91	8.90	121.3	57.91
N40				8.20	118.7	54.37
L41			57.89	7.67	120.2	57.65
M42	8.52	114.9	55.36	8.54	115.2	55.36
L43	7.65	115.7	53.29	7.65	115.7	53.29
T44	9.68	118.0	58.79	9.68	118.0	58.79
- - - -						
P45			65.28			65.53
D46	8.24	116.7	56.58	8.53	117.2	56.70
E47	8.15	123.1	58.38	8.15	123.1	58.38
R48	7.67	117.4	59.99	7.43	116.9	59.81
E49	7.67	117.4	58.81	7.57	117.4	58.81
A50	8.09	122.4	54.61			54.40
L51			57.91	7.91	119.4	57.92
G52	8.04	106.3	46.69	8.11	106.3	46.63
T53	8.04	120.1	66.36	7.84	119.9	66.63
R54	8.13	119.9	60.23	8.19	119.7	60.15

Table 1 continued.

Residue	Subunit A			Subunit B			Residue	Subunit A			Subunit B		
	^{1}HN	^{15}N	^{13}Cα	^{1}HN	^{15}N	^{13}Cα		^{1}HN	^{15}N	^{13}Cα	^{1}HN	^{15}N	^{13}Cα
V55	8.12	118.8	65.67			65.65	T83	9.30	118.7	65.92	8.65	117.0	66.15
R56	7.46	121.9	57.32	7.40	121.9	57.18	R84	7.86	121.4	59.85	8.02	123.7	59.21
I57	8.30	119.0	66.13				G85	7.27	110.4	48.40	7.91	110.4	48.18
V58	8.17	119.0	67.12	8.17	119.0	67.12	S86	8.76	117.6	60.10	8.47	117.6	60.89
E59	8.33	117.5	59.53	8.34	118.0	59.74	N87	9.01	116.2	55.35	8.55	119.3	55.32
E60	8.32	113.9	56.78	8.21	114.5	57.18	S88	8.02	118.6	62.45	7.97	118.1	62.07
L61	9.03	123.6	57.40	8.93	123.2	57.16	L89	8.34	124.6	57.60	8.46	124.3	57.44
L62	7.96	117.7	56.55	8.11	119.1	56.52	K90	7.71	118.4	58.89	7.93	118.4	58.58
R63	8.27	119.3	58.72	8.10	119.2	58.58	A91	7.01	118.1	51.12	7.06	118.4	51.28
G64	7.54	102.3	46.42	7.52	102.8	46.15	A92	7.37	123.6	49.89	7.37	123.6	49.89
E65	7.96	118.9	56.58	7.87	119.3	56.79	- - - -						
M66	8.58	119.3	54.60	8.33	118.8	55.16	P93			62.01			62.01
S67	8.66	119.0	57.15	8.69	118.2	56.95	V94	8.81	125.1	65.56	8.81	125.1	65.56
Q68	10.1	122.4	59.96	8.73	120.4	59.72	E95	9.76	119.2	59.45	9.76	119.2	59.45
R69	8.77	120.2	58.89	8.09	117.5	58.27	L96	7.25	117.6	56.70	7.25	117.6	56.70
E70	7.59	119.7	58.28	7.59	119.7	58.28	R97	7.83	119.7	60.65	7.83	119.7	60.65
L71	8.30	122.8	57.43	8.40	121.7	57.23	Q98	8.70	116.3	58.38	8.70	116.3	58.38
K72	8.83	119.8	59.45	8.17	119.3	59.00	W99	7.44	121.4	61.32	7.42	121.4	61.28
N73	7.78	119.2	55.07	7.60	117.0	54.90	L100	8.81	117.7	57.08	8.81	117.7	57.08
E74	8.25	121.3	58.44	8.15	119.3	57.95	E101	7.92	116.9	59.29	7.92	116.9	59.29
L75	7.87	114.3	54.37	8.15	114.8	54.48	E102	7.41	118.2	58.30	7.41	118.2	58.30
G76	8.06	111.4	46.44	7.86	109.9	46.54	V103	7.87	113.7	53.65	7.91	113.8	63.67
A77	7.40	118.5	49.60	7.14	119.2	49.17	L104	8.24	117.9	55.55	8.24	117.9	55.55
G78	9.44	105.9	43.58	9.79	106.7	43.72	L105	7.05	115.9	54.47	7.05	115.9	54.47
I79	9.35	127.5	61.77	8.82	123.7	62.49	K106	7.18	119.7	55.82	7.18	119.7	55.82
A80	9.07	120.9	53.89	9.19	123.9	54.63	S107	8.15	117.8	57.66	8.15	117.8	57.66
T81	7.52	118.6	65.02	7.13	117.1	64.45	D108	7.93	127.7	55.08	7.93	127.7	55.08
I82	7.13	122.3	64.67	7.60	121.7	64.72							

[a] A dashed line indicates a point where the sequential connectivities are broken for both subunits in triple resonance spectra.

in order to perform triple resonance 3D NMR experiments as well as ^{2}H-labelled in order to overcome the short ^{13}C T_2 relaxation times associated with such a large complex (Shan et al., 1996). Although we were able to assign the backbone resonances of the protein in this complex, we were not able to assign them to specific subunits within each dimer (A or B). We report here the analysis of chemical shift data for both intact and Δ2–7 trpR/DNA tandem complexes and their subunit-specific assignments. The data indicate that sites of protein-protein contact between tandem dimers mirror those observed in the crystal structure. However, the solution data indicate that the N-terminal arms may exist in additional conformations beyond what is observed in the crystal.

Materials and methods

Triply (^{15}N, ^{13}C, > 90% ^{2}H) labelled trpR was prepared as described by Shan et al. (1996) and (^{15}N, ^{13}C, ∼70% ^{2}H) trpR was prepared as outlined in Yamazaki et al. (1994a,b). Both samples were used to prepare the tandem trpR/operator complexes as described in Shan et al. (1996). Final NMR conditions were 1.2 mM trpR (dimer), 0.6 mM double-stranded operator DNA, 3.0 mM 5-methyl-L-tryptophan, and 10 mM sodium phosphate adjusted to pH 6.0, 7% D_2O/93% H_2O, 45 °C. After several weeks of NMR experiments at 45 °C, the 70% ^{2}H sample began to show signs of limited proteolytic degradation. The progress of the reaction was monitored over a period of several more weeks (sample stored at 4 °C) by recording HSQC spectra (at 45 °C) and SDS-PAGE. The latter indicated

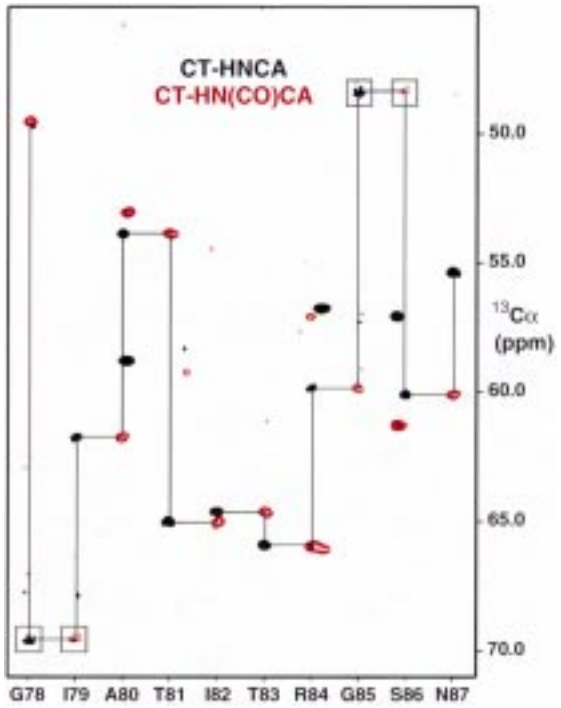

Figure 2. Strip plots of the CT-HNCA (black) and CT-HN(CO)CA (red) for subunit A of Δ2–7 trpR corresponding to the DNA binding region of the protein. Each strip shows the correlation between the ^{15}N(*i*), HN(*i*) and Cα(*i*) or Cα(*i–1*) of the indicated residue. Boxed residues have a negative phase. The Cα of Gly78 is folded in the carbon dimension.

only a small decrease in the size of the protein, although many changes in chemical shift were observed in HSQC spectra. N-terminal sequencing (Biotechnology Service Center, University of Toronto) indicated a single reaction product corresponding to cleavage of the protein between Tyr7 and Ser8, resulting in an 'armless' repressor (residues 8–108, hereafter referred to as Δ2–7trpR; the N-terminal Met1 is removed in vivo (Joachimiak et al., 1983)). Once the reaction appeared to be complete by SDS-PAGE, 3D constant-time (CT) HNCA, CT-HN(CO)CA, HN(CA)CB and HN(COCA)CB spectra were acquired, processed and analyzed as described in Yamazaki et al. (1994a,b). No further changes in the spectra were detected over a period of 12 months (sample stored at 4 °C). Experimental NMR details for the intact tandem trpR complex have been previously reported (Shan et al., 1996).

Chemical shift index Secondary structure analyses of each form of trpR were performed using the ^{13}Cα chemical shift index (CSI) method (Wishart and

Table 2. ^{1}HN, ^{15}N and ^{13}Cα chemical shifts for intact trpR bound in tandem to DNA[a]

Residue	Subunit A				Subunit B			
	^{1}HN	^{15}N	^{13}Cα	^{13}Cβ	^{1}HN	^{15}N	^{13}Cα	^{13}Cβ
A2								
E3					8.68	120.1	55.79	28.65
E4	8.44	122.4	55.36	28.82	8.32	120.9	54.99	29.03
S5	8.26	118.7	55.75	62.77	8.43	117.5	55.65	63.00
- - - -								
P6			63.15	30.87			62.32	33.36
Y7	7.92	119.5	57.49	37.51	8.33	122.4	57.69	37.73
S8	7.86	117.5	58.05	63.19	8.02	118.3		63.32
A9	8.45	126.4	53.92	17.34				
A10	8.07	121.1	53.67	17.42	8.26	123.0	52.57	17.99
M11	7.86	118.6	56.39	31.43	8.13	118.8	55.41	31.71
A12	8.06	123.1	54.49	17.40	8.15	124.2	53.78	17.87
E13	8.15	119.2	58.34	28.27	8.43	119.1	57.79	28.38
Q14	7.94	121.1	58.75	27.49				
R15	8.55	117.2	58.03	28.83				
- - - -								
H16			58.90	28.42	8.32	120.6	59.23	28.41
Q17	8.31	118.2	58.30	27.22	8.18	117.8	58.45	27.16
E18	8.39	120.5	58.42	27.45	8.13	119.9	58.58	27.83
W19			59.09	27.56	8.29	122.8	58.65	28.16
L20	7.75	117.8	57.10	39.23	8.03	118.1	56.98	39.29
R21	7.89	120.6	59.02	28.64	7.71	119.8	58.99	28.62
F22	8.11	121.4	61.09	36.38	7.74	120.6	61.42	36.35
V23	7.77	119.5	66.10	30.08	7.89	119.5	65.93	30.24
D24	7.74	121.5	56.75	40.56	8.02	121.6	56.76	40.59
L25	8.24	124.4	57.38	41.47	7.96	124.1	57.47	41.12
L26	8.34	119.9	56.65	38.88			56.60	38.85
K27	7.42	120.9	59.44	31.15	7.29	120.9	59.51	31.14
N28	7.57	116.7	55.42	38.06	7.57	116.7	55.42	38.06
A29	8.88	126.7	55.27	16.70	8.97	126.7	55.18	16.66
Y30	8.38	119.3	61.28	37.27	8.38	119.3	61.28	37.27
Q31	7.32	117.3	57.57	27.44	7.28	117.2	57.51	27.43
N32	7.30	115.8	52.71	39.62	7.30	115.8	52.71	39.62
D33	7.88	117.6	55.91	39.08	7.88	117.6	55.91	39.08
L34	8.24	118.1	53.11	42.09	8.24	118.1	53.11	42.09
H35	9.03	121.1	58.80	27.75	9.03	121.1	58.80	27.75
L36	7.46	119.6	59.24	36.42	7.46	119.6	59.24	36.42
- - - -								
P37			64.79	29.8			64.89	29.85
L38	7.93	120.3	58.03	40.13	7.93	120.3	58.03	40.13
L39	8.86	121.0	57.92	40.43	8.91	121.2	57.95	40.30
N40	8.06	118.6	54.38	37.29	8.16	118.5	54.39	37.24
L41	7.66	120.3	57.90				57.93	41.20
M42	8.63	115.2	55.47	31.06	8.44	114.7	55.36	30.82
L43	7.68	115.9	53.49	15.79				
T44	9.45	117.5	58.82	66.99	9.66	118.1	58.81	38.81
- - - -								
P45			65.27	30.79			62.91	30.88
D46	8.26	116.5	56.48	38.06	7.83	119.2	56.91	37.65

Table 2 continued.

Residue	Subunit A				Subunit B				Residue	Subunit A				Subunit B			
	^{1}HN	^{15}N	^{13}Cα	^{13}Cβ	^{1}HN	^{15}N	^{13}Cα	^{13}Cβ		^{1}HN	^{15}N	^{13}Cα	^{13}Cβ	^{1}HN	^{15}N	^{13}Cα	^{13}Cβ
- - - -									G78	9.40	105.9	69.65		9.73	106.8	69.79	
E47									I79	9.34	127.2	61.56	37.10	8.74	123.3	62.46	36.43
R48	7.62	117.1	59.92		7.51	116.8	59.82	29.37	A80	9.04	120.9	53.92	18.65	9.14	124.2	54.77	17.45
E49	7.61	117.0	58.80	28.08	7.74	118.0	58.62	28.05	T81	7.54	118.6	65.09	65.09	7.20	117.1	64.59	67.68
A50	8.06	122.1	54.64	16.27	8.35	122.8	54.52	16.31	I82	7.14	122.1	64.47	35.15	7.68	121.8	64.73	36.83
L51	8.02	118.8	58.02	40.32	7.88	119.1	58.02	39.94	T83	9.29	118.6	66.00	67.54	8.64	117.1	66.19	67.93
G52	8.01	105.9	46.82			8.00	105.1	46.78	R84	7.82	121.2	59.67		7.97	123.6	59.24	28.67
T53	7.94	120.3	66.13	40.13					G85	7.19	110.2	48.44		7.96	110.4	48.19	
- - - -									S86	8.65	117.0	59.97	62.99	8.48	117.6	60.90	61.87
R54									N87	9.03	116.2	55.41	36.56	8.51	119.3	55.34	36.76
V55	7.86	118.2	66.17	30.34			66.00	30.30	- - - -								
R56	7.36	120.8	58.11	28.24	7.36	121.4	57.35	27.35	S88								
I57	8.32	118.7	65.92	36.79	8.13	118.5	66.13	36.61	L89	8.30	124.8	57.61	41.35	8.46	124.3	57.46	41.08
V58	8.14	119.0	67.14	30.29	8.14	119.0	67.14	30.29	K90	7.82	117.1	59.02	31.87	7.92	118.4	58.60	31.28
E59	8.32	117.5	59.51	29.01	8.42	118.3	59.81	28.91	A91	7.10	118.2	51.21	18.45	7.04	118.4	51.29	18.30
E60	8.31	114.0	56.83	28.11	8.11	114.9	57.32	28.40	A92	7.36	124.1	49.90	16.92	7.34	123.8	49.68	16.92
L61	8.98	123.6	57.47	39.41	8.85	122.9	57.19		- - - -								
L62					8.16	119.1	56.47	40.97	P93			62.04	31.25			61.97	31.22
R63	8.27	119.1	58.77	29.75	8.08	119.2	58.58	29.51	V94	8.80	125.1	65.54	30.56	8.89	125.0	65.96	30.60
G64	7.55	102.3	46.59		7.52	103.3	46.14		E95	9.75	119.3	59.52	27.27	9.85	119.2	59.50	27.35
E65	7.99	118.9	56.66	30.23	7.90	119.4	56.79	29.88	L96	7.24	117.7	56.73	40.77	7.24	117.7	56.73	40.77
M66	8.53	119.1	54.51	34.79	8.25	118.9	55.25	34.01	R97	7.75	119.0	60.85	29.22	7.84	119.8	60.65	29.00
S67	8.66	118.7	57.12	63.55	8.61	117.7	57.07	63.91	Q98	8.68	116.5	58.43	27.65	8.68	116.5	58.43	27.65
Q68	10.0	122.2	60.01	25.71	8.68	120.5	59.41	27.31	W99	7.45	121.4	61.26	27.77	7.40	121.4	61.25	27.77
R69	8.65	119.9	58.94	28.65	8.09	117.9	58.01	29.10	L100	8.80	117.7	57.09	41.40	8.80	117.7	57.09	41.40
E70	7.63	120.2	58.42	28.56	7.63	119.8	58.17	28.87	E101	7.88	117.0	59.34	28.44	7.88	117.0	59.34	28.44
L71	8.29	122.9	57.46	41.01	8.36	121.7	57.17	41.02	E102	7.38	118.2	58.33	28.49	7.38	118.2	58.33	28.49
K72	8.82	119.7	59.44	29.27	7.91	120.3	58.69	31.33	V103	7.92	113.7	63.64	30.94	7.87	113.7	63.64	30.95
N73	7.78	119.0	55.10	37.37	7.63	117.2	54.80	37.93	L104	8.22	117.9	55.56	41.26	8.22	117.9	55.56	41.26
E74	8.22	121.3	58.46	28.12	8.35	119.8	57.70	28.71	L105	7.03	115.9	54.54	39.90	7.03	115.9	54.54	39.90
L75	7.85	114.4	54.39	42.22	8.15	115.5	54.62	41.85	K106	7.18	119.8	55.85	31.75	7.18	119.8	55.85	31.75
G76	8.02	111.1	46.41		7.89	109.7	46.37		S107	8.15	117.7	57.62	63.42	8.15	117.7	57.62	63.42
A77	7.39	118.5	49.73	20.81	7.06	119.4	49.43	20.69	D108	7.92	127.7	55.11	41.13	7.92	127.7	55.11	41.13

[a] A dashed line indicates a point where the sequential connectivities are broken for both subunits in triple resonance spectra.

Sykes, 1994). The random coil ^{13}Cα values used in this method were modified to account for deuterium isotope effects on ^{13}C nuclei (Hansen, 1988). For Cα carbons, upfield isotope shifts of 0.50 ppm are expected, based on several comparisons of ^{13}Cα shifts in protonated and deuterated forms of the same protein (Venters et al., 1996; Gardner et al., 1997; Garrett et al., 1997). Alpha-helical residues were identified by the criteria described by Wishart and Sykes (1994) using a MATLAB script written by Catherine Zwahlen (University of Toronto).

Results and discussion

Sequence-specific resonance assignments

We have previously reported the sequence-specific, subunit-nonspecific backbone (^{1}HN, ^{15}N, ^{13}Cα, ^{13}Cβ) assignments for the tandem complex of intact trpR (Shan et al., 1996). In order to resolve the many overlapping correlations in this alpha-helical protein, it was necessary to use CT spectra correlating both the ^{13}Cα and ^{13}Cβ resonances with the backbone NH. Furthermore, in order for the CT spectra to be of sufficient sensitivity high levels of deuteration were necessary.

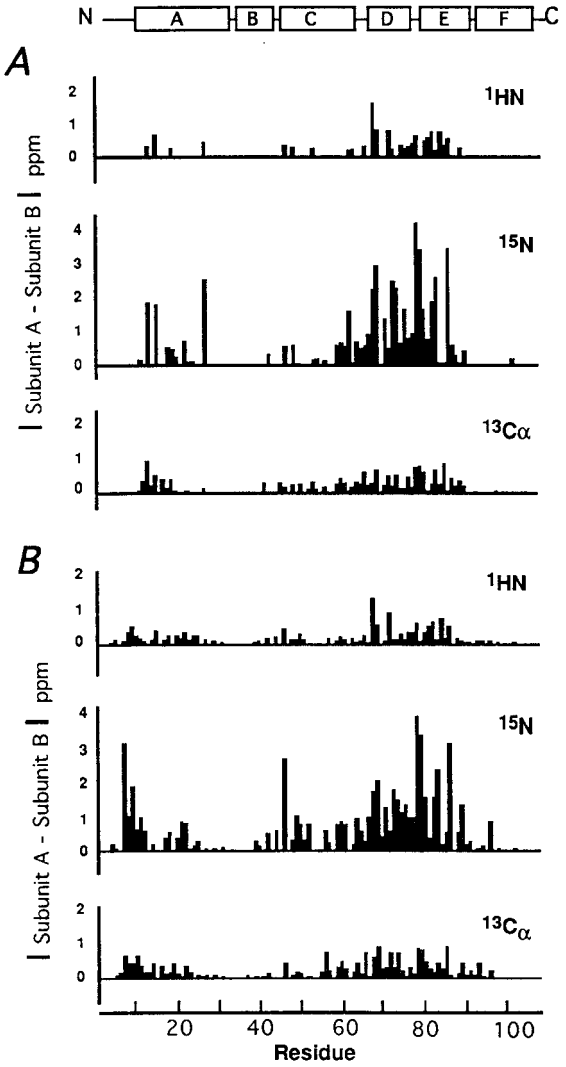

Figure 3. Plots of the absolute value of the difference in chemical shifts between subunits A and B of (A) Δ2–7 trpR and (B) intact trpR (Shan et al., 1996).

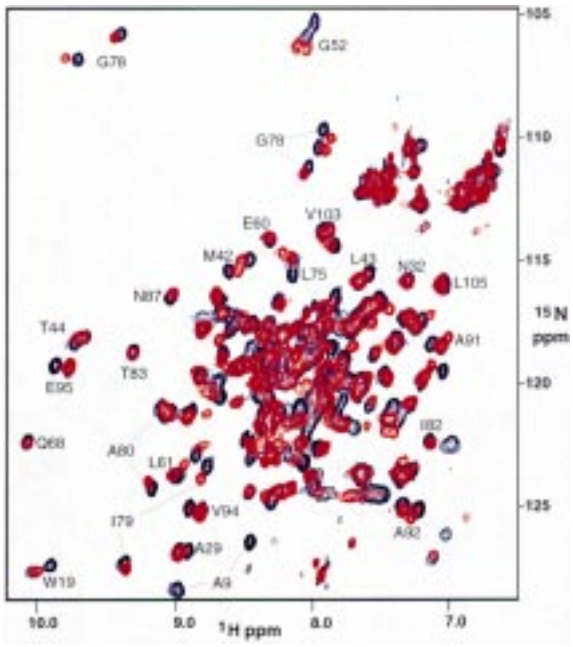

Figure 4. The HSQC spectrum of intact trpR (blue) overlaid with that of Δ2–7 trpR (red). Selected resolved resonances are labelled. This spectrum of Δ2–7 trpR was recorded when the proteolytic cleavage of residues 2–7 was approximately 90% complete. Therefore some of the residues with more intense peaks still show a small peak due to the ∼10% intact protein in the sample (for example residues Val94 and Glu95).

We found that deuterium incorporation of ∼70% was sufficient for recording good quality CT {^{1}HN, ^{15}N, ^{13}Cα} correlations, but that due to the significant times during which carbon ^{13}Cα and ^{13}Cβ magnetization resides in the transverse plane, higher deuterium levels were required to obtain good sensitivity in CT spectra recording Cβ chemical shifts (Yamazaki et al., 1994a,b; Shan et al., 1996). Our sample of the Δ2–7 protein is only 70% deuterated, precluding the use of CT experiments for recording Cβ correlations. Nevertheless, for the following reasons we were still able to make sequence-specific assignments for most residues in the Δ2–7 complex. First, there was a greater degree of chemical shift degeneracy between the two subunits in the Δ2–7 protein resulting in fewer unique resonances to assign in the triple resonance spectra. Second, many of the ambiguities in Cα connectivities could be resolved based on comparisons with the spectra of the intact protein which had been assigned using both Cα and Cβ correlations. Figure 2 shows a strip plot of the CT-HNCA and CT-HN(CO)CA spectra of the Δ2–7 trpR complex.

Figure 3A illustrates the differences in ^{1}HN, ^{15}N and ^{13}Cα chemical shifts between the two subunits of Δ2–7 trpR. As we observed for the intact complex (Figure 3B, and Shan et al., 1996), the two regions of the protein that differ the most between subunits are the DNA-binding domain (residues 68–90) and the N-terminus. Since these differences map to the regions of contact between dimers observed in the crystal structure, the chemical shift differences are most likely due to differences in dimer-dimer contacts in the two subunits (see Figure 1). This data alone, however, is not sufficient to identify which subunit of the dimer gives rise to each resonance. Subunit-specific assignments are needed in order to better understand tandem

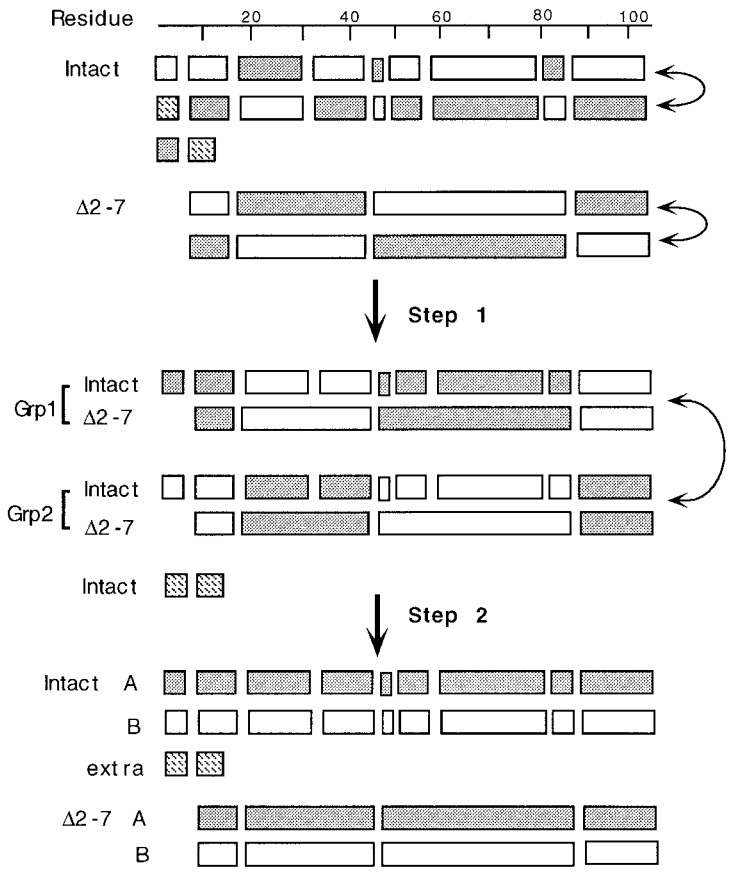

Figure 5. Schematic of the process used to assign sequential resonance connectivities to subunits A and B of intact and Δ2–7 trpR. The boxes represent stretches of sequential connectivities identified in triple resonance experiments. The shaded and white boxes represent resonances that arise from subunits A and B, respectively. The hatched boxes represent the extra resonances of residues 3–12. In step 1 the stretches of sequence were paired so as to minimize the chemical shift difference between intact and Δ2–7 trpR. In step 2 the chemical shift differences for Grp1 and Grp2 were compared with the crystal structure of a tandem complex (Lawson and Carey, 1993) to generate subunit-specific assignments. See the main text for a more detailed discussion.

binding and the role of the N-terminus in tandem interactions. The close correspondence between the largest chemical shift changes in Figure 3 and the sites of protein-protein contact in the crystal structure allowed us to assign resonances to subunits A or B of each dimer by comparing the chemical shifts of intact trpR with Δ2–7 trpR as described in detail below.

Subunit specific assignments

Figure 4 compares HSQC spectra of the tandem complexes of intact trpR and Δ2–7 trpR. Clearly Δ2–7 is still bound to the DNA (chemical shifts are very different from those of free trpR) and likely in a very similar conformation to that in the intact complex. However, there are small, but significant changes in chemical shift throughout the protein associated with deletion of the six N-terminal residues. For certain residues on the 'back' side of the protein away from the DNA, only one of two peaks remains in the Δ2–7 protein (for example Trp19, Ala29, Val94 and Glu95). Other peaks, for example Gly78 and Gly52, still give rise to two separate peaks for each subunit, both different from those of the intact trpR complex. The unequal differences in chemical shift between dimers of intact vs Δ2–7 trpR, such as those described above, can be used to clarify which subunits give rise to each set of resonances (see below).

The triple resonance assignment process is predicated on linking together correlations between the ^{15}N and HN chemical shifts of residues *i* with the Cα and/or Cβ shifts of residues *i* and *i–1*. Proline residues and unassigned residues naturally lead to breaks in the series of sequential connectivities. In the case of an asymmetric (with respect to chemical shift) homo-

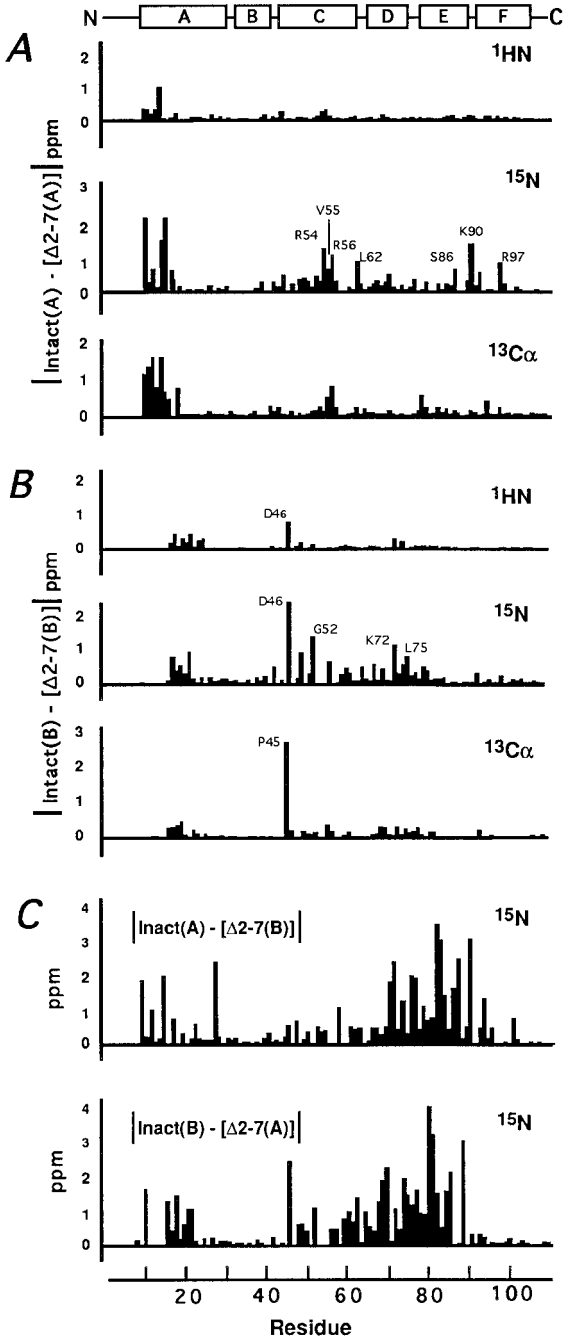

Figure 6. Plots of the absolute value of the difference in chemical shifts for (A) subunit A of intact trpR and subunit A of Δ2–7 trpR, (B) subunit B of intact trpR and subunit B of Δ2–7 trpR, and (C) subunit A of intact trpR and subunit B of Δ2–7 trpR (upper panel), and subunit B of intact trpR with subunit A of Δ2–7 trpR (lower panel). The positions of the six helices are indicated along the top. Residues outside the N-terminal region that have chemical shift changes greater than 0.5 ppm are labelled and discussed in the text.

Table 3. ^{1}HN, ^{15}N and ^{13}Cα chemical shifts for the third form of N-terminal residues in the tandem complex of intact trpR

Residue	Subunit C			
	^{1}HN	^{15}N	^{13}Cα	^{13}Cβ
A2				
E3	8.68	120.14	55.79	28.65
E4	8.49	122.21	55.35	28.69
S5	8.37	118.69	55.84	62.82
P6			63.32	31.12
Y7	7.84	116.39	56.82	37.03
S8	7.51	116.43	57.60	63.68
A9	9.00	128.22	54.35	17.53
A10	8.34	120.48	54.35	17.40
M11	7.68	117.60	56.06	30.78
A12	7.97	122.54	54.66	17.14

dimer such as tandemly bound trpR, these break points result in a series of isolated stretches of sequence which cannot be assigned specifically to a given subunit of the dimer. Spectra of both intact and Δ2–7 trpR have several such break points in the sequential connectivities due to either proline residues (no NH), unassigned residues, or regions where the chemical shifts of both subunits are identical. For the intact protein there were 9 stretches of sequence that could be interchanged between subunits in the absence of any additional information (Figure 5). In the Δ2–7 complex there were only 4 ambiguous stretches. The smaller number of ambiguities in Δ2–7 trpR is due to the absence of Pro6, fewer unassigned residues, and longer stretches in which the chemical shifts of both subunits are degenerate. In addition, we observed an 'extra' set of resonances (total of 3 sets) for residues 3–12 in the intact protein (see below).

We pieced together these isolated stretches of sequential assignments into groupings (termed group 1 (Grp1) and group 2 (Grp2)) based on the assumption that the absolute values of the differences in chemical shift, |Grp1(Intact) – Grp1(Δ2–7)| and |Grp2(Intact) – Grp2(Δ2–7)| for each backbone atom position are a minimum (step 1 in Figure 5). Thus, we assumed that the structure and chemical environment of subunit A of intact trpR would be more like that of subunit A in Δ2–7 trpR than subunit B of either intact or Δ2–7 trpR. This analysis led to the pairing of stretches of sequence between the two forms of the protein (middle panel of Figure 5). Except for the N-terminal region

(up to ∼ residue 22) the largest differences in chemical shift between intact and Δ2–7 trpR were limited to several specific residues which mapped remarkably well to sites of contact with the N-terminal 'arm' of the protein in the tandem cocrystal (Lawson and Carey, 1993). It is noteworthy that the chemical shift differences between intact and Δ2–7 trpR were not the same for the Grp1 and Grp2 pairs of the above analysis. Since the protein-protein contacts involving the N-terminal residues are not the same for subunits A and B in the crystal structure it is likely that the differences in Grp1 and Grp2 pairs reflected the different subunit contacts with residues 2–7. Thus, by comparing the chemical shift differences of the Grp1 and Grp2 pairs with differences in contacts in the crystal structure we were able to generate a self-consistent set of subunit-specific assignments (step 2 in Figure 5). Specifically, for the complex studied here in solution (see Figure 1) we assigned subunit A as the subunit containing tandem interactions via its N-terminus, and subunit B as the 'outer' subunit with an unbound or free form of the N-terminus. A number of examples will serve to illustrate how subunit-specific assignments were accomplished.

In the crystal structure the terminal amino group of the A subunit interacts with Asp46 of subunit B of the same dimer. In the intact tandem repressor this interaction results in reorientation of the side chain of Asp46 and a small movement of the Pro45 ring compared to the intact, 1:1 trpR/DNA crystal structure (Otwinowski et al., 1988; Lawson and Carey, 1993). On this basis the sequential assignment stretches which had a large chemical shift difference between intact and Δ2–7 trpR for Pro45 and Asp46 were assigned to subunit B. Thus, our solution data suggests a readjustment of Pro45 and Asp46 in the absence of residues 2–7 in the tandem complex.

In the tandem crystal structure Leu62, Ser86, Lys90 and Arg97 of subunit A all line a hydrophobic pocket which interacts with the aromatic ring of Tyr7 of subunit A of the adjacent dimer. Thus, the sequence stretches containing residues Leu62, Ser86, Lys90 and Arg97 were assigned to subunits A or B on the basis of having either large (A) or small (B) chemical shift differences between intact and Δ2–7 trpR.

Because the Δ2–7 protein lacks residues 2–7, the corresponding residues of intact trpR could not be assigned to a specific subunit by the procedure described above. Instead, subunit-specific assignments for residues 2–5 were based on the similarity of their resonance intensities in relation to those of the adjacent residues 6–12. In our original data (Shan et al., 1996) three sets of resonances were observed for residues Glu3-Glu12 suggesting an 'extra' form of the N-terminal region of one of the subunits. In contrast, the Δ2–7 repressor has only two sets of resonances for all residues (subunits A and B). In the triple resonance spectra two of the three sets of correlations for residues 2–12 were of medium intensity while the third was of higher intensity. We reasoned that residues 2–12 of subunit B, which do not participate in intersubunit contacts, would have longer T_2 relaxation times and more intense resonances. This is the case for these residues in the 1:1 trpR:DNA complex (Yamazaki et al., 1994b) where tandem interactions are not possible. These signal intensities also suggest that the 'extra' set of resonances for residues 2–12 arises from a form of the protein with similar mobility (and concentration) to that of subunit A – perhaps an alternative conformation of these residues involved in protein-protein or protein-DNA contacts. It is worth noting that we have eliminated the possibility that the third set of resonances are early products of the proteolytic cleavage that generated the Δ2–7 repressor for two reasons. First, the third set of resonances was present in the very first 3D spectrum before the HSQC showed evidence of proteolysis. Second, there are clear sequential connectivities in the triple resonance data between Tyr7 and Ser8 in all three sets of resonances, indicating that they are from the intact form of the protein.

Figure 6 shows the final subunit-specific chemical shift comparisons for intact vs. Δ2–7 trpR. For comparison we include the ^{15}N chemical shift differences between opposite subunits of intact and Δ2–7 trpR (Figure 6C) showing how the 'wrong' pairing of assignments results in much larger chemical shift differences for all but the extreme N-terminal region. The differences between the intact and Δ2–7 forms of the same subunit are much fewer and smaller than those between the two subunits within each dimer (Figures 3A and 3B versus 6A and 6B) or the differences between opposite subunits of intact and Δ2–7 trpR (Figures 6A and 6B versus 6C). Because of the strong dependence of the chemical shift on local environment, the small chemical shift differences seen in Figures 6A and 6B are a sensitive probe for small differences in molecular contacts or conformation. The close correspondence of the chemical shift difference data with the regions of protein-protein contact in the crystal structure is illustrated in Figure 7.

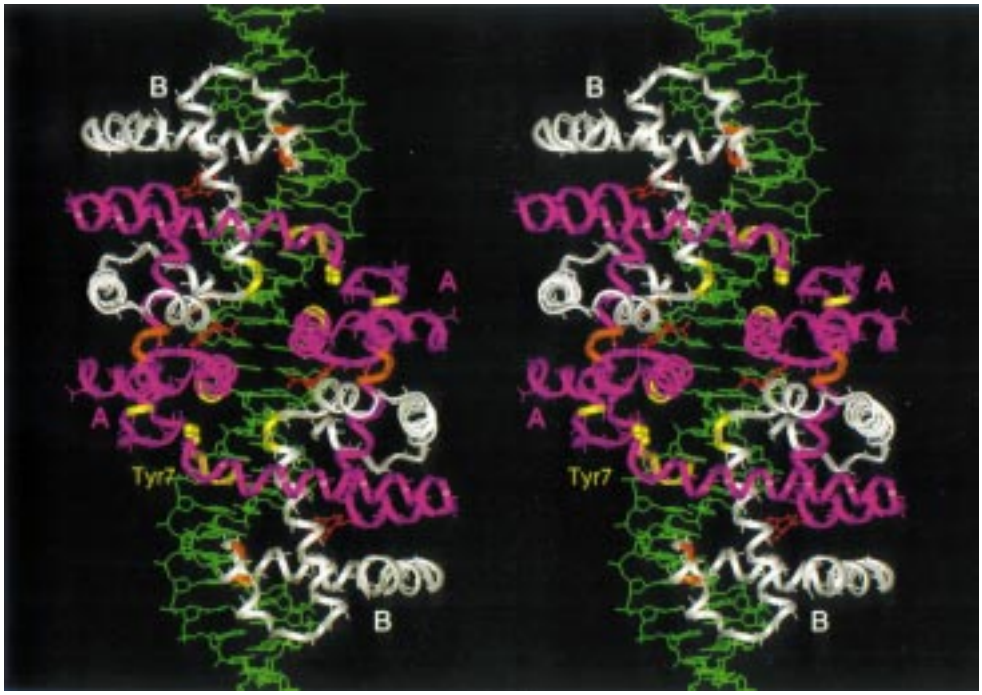

Figure 7. Stereoview of a backbone ribbon display of two tandem dimers from the tandem cocrystal showing dimer-dimer contacts (Lawson and Carey, 1993). The purple and white subunits correspond to subunits A and B of Figure 1, respectively. Residues 2–7 of subunit A are shown in yellow, as are those residues with chemical shift changes greater than 0.5 ppm in Δ2–7 trpR due to direct interactions with Ala2, or Tyr7. The side chain of Tyr7 is shown in yellow space-filling mode. Other residues that have significant backbone changes are shown in orange. The corepressor, 5-methyltryptophan, is shown as a stick model in red.

Once the subunit-specific assignments were complete a number of additional small, but significant changes in chemical shift that occur with deletion of residues 2–7 were noted. The solution data shows small changes at Glu49 and Gly52 of subunit B which may reflect the close proximity of the carboxylate oxygen of Glu3 to both residues in the crystal structure of the intact tandem complex (3.8 Å to the Cβ of Glu49 and 6.6 Å to the backbone nitrogen of Gly53). Residues Arg54, Val55 and Arg56 of subunit A and Lys72 and Leu75 of subunit B all show changes greater than 0.5 ppm between intact and Δ2–7 trpR. However, none of these residues has direct contacts with residues 2–7 in the crystal. These small changes remain unexplained in terms of the crystal structure and may therefore reflect small differences in the solution and crystalline states. Residues 54–56 are in the ligand binding pocket and their difference in chemical shift in Δ2–7 trpR may reflect slightly different concentrations of the corepressor, 5-methyltryptophan, used in the samples of intact and Δ2–7 trpR. Since deletion of residues 2–7 caused perturbations in the N-terminus extending as far as residue 22 in the A subunit, these effects may be transmitted to Lys72 and Leu75 of the opposite subunit which lie adjacent to the N-terminal region of helix A.

Secondary structure analysis

A modified version of the CSI (Wishart and Sykes, 1994) was used to compare the secondary structure of each subunit in intact and Δ2–7 trpR (Figure 8). All subunits contain the same 6 helices (A–F) previously identified by structural studies of holorepressor (Zhao et al., 1993; Schevitz et al., 1995) and the 1:1 protein/DNA complex (Otwinowski et al., 1988). Except as noted below for helix A, the helix boundaries of all subunits are approximately (± one residue) as follows: helix A (residues 12–33), helix B (residues 35–41), helix C (residues 45–65), helix D (residues 67–74), helix E (residues 80–90) and helix F (residues 93–103). These helix boundaries are within 2 residues of those determined by crystallographic methods (Otwinowski et al., 1988; Schevitz et al., 1995) and somewhat longer than those derived by solution NMR for the free holorepressor (Zhao et al., 1993). It is interesting to note that in the present study,

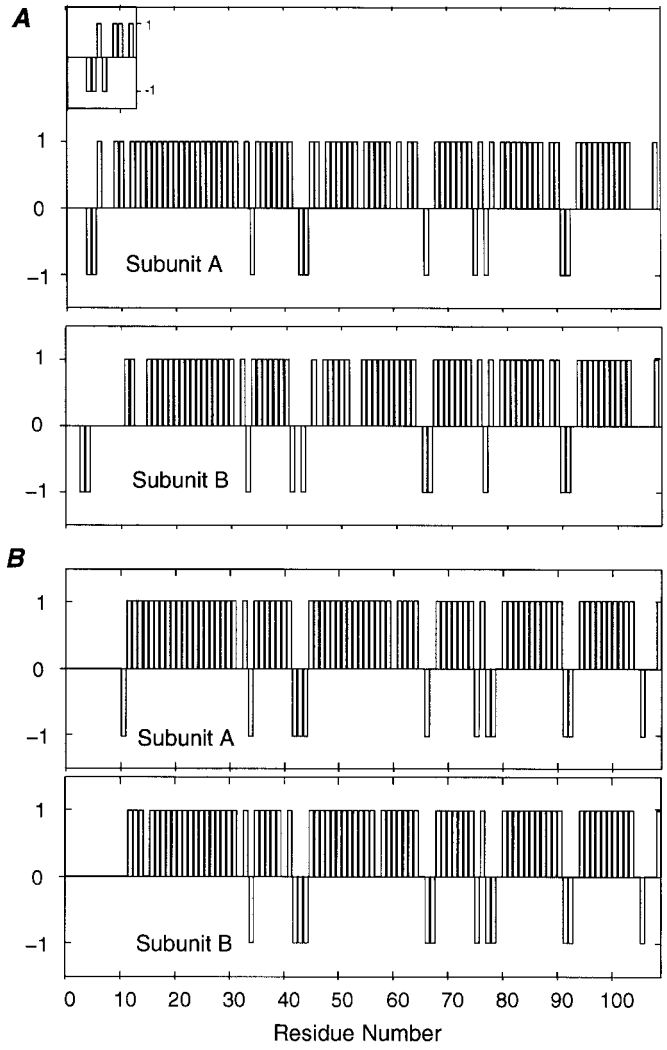

Figure 8. CSI analysis of Cα values for trpR subunits A and B in the tandem complex. (A) Intact trpR. The inset is for the alternative (extra) conformation of residues 2–12. (B) Values for Δ2–7 trpR.

helix A is 3 residues longer at the N-terminus (starting at residue 9) for subunit A and the alternative (extra) conformation of intact trpR relative to the helix length for subunit B. Subunit A has tandem interactions involving the N-terminus which may stabilize helix A. Thus, the CSI data also suggests that the alternative conformation of residues 2–13 is more like that of the bound subunit.

Conclusions

We have presented chemical shift data characterizing the 64 kDa complex of trp repressor with and without N-terminal residues 2–7. A careful analysis of these data in the context of a previously reported crystal structure establishes the sensitivity of backbone chemical shifts to small perturbations in the protein. The trpR dimers in the crystal structure of the tandem complex (Lawson and Carey, 1993) were identical to those in the 1:1 dimer:operator complex (Otwinowski et al., 1988) with the exception of the N-terminal residues and Asp46 which were involved in tandem binding. Our solution data of intact and Δ2–7 trpR reflect these same differences. Although the current analysis was based on the crystallographic data it would not have been successful if the solution structure did not closely correspond to the crystal structure. Furthermore, the solution data provide evidence for an alternative con-

formation of the N-terminus in solution. The similarity of the extra set of N-terminal resonances both in intensity and in predicted secondary structure suggests that the extra form of residues 2–12 is from an alternative state of the 'bound' N-terminus, similar to subunit A. These data demonstrate that significant structural conclusions can be drawn from backbone chemical shift data of large protein complexes and establish that triple resonance NMR spectroscopy of triply labelled proteins is a powerful method for addressing specific structural questions in large systems.

Acknowledgements

This work is supported by the National Cancer Institute of Canada with funds from the Canadian Cancer Society (C.H.A. and L.E.K.), by the Medical Research Council (MRC) of Canada (C.H.A. and L.E.K.) and the Natural Sciences and Engineering Research Council of Canada (L.E.K.). K.H.G. was supported by a postdoctoral fellowship from the Helen Hay Whitney Foundation. C.H.A. and L.E.K. are MRC Research Scholars.

References

Carey, J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 975.
Carey, J. (1989) *J. Biol. Chem.*, **264**, 1941–1945.
Czernik, P.J., Shin, D.S. and Hurlburt, B.K. (1994) *J. Biol. Chem.*, **269**, 27869–27875.
Gardner, K.H., Rosen, M.K. and Kay, L.E. (1997) *Biochemistry*, **36**, 1389–1401.
Garrett, D.S., Seok, Y.-J., Liao, D.-I., Peterkofsky, A., Gronenborn, A.M. and Clore, G.M. (1997) *Biochemistry*, **36**, 2517–2530.
Hansen, P.E. (1988) *Prog. NMR Spectrosc.*, **20**, 207–255.
Joachimiak, A., Kelley, R.L., Gunsalus, R.P., Yanofsky, C. and Sigler, P.B. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 668–672.
Klig, L.S., Carey, J. and Yanofsky, C. (1988) *J. Mol. Biol.*, **202**, 769–777.
Kumamoto, A.A., Miller, W.G. and Gunsalus, R.P. (1987) *Gene Dev.*, **1**, 556–564.
Lawson, C.L. and Carey, J. (1993) *Nature*, **366**, 178–182.
Liu, Y.-C. and Matthews, K.S. (1993) *J. Biol. Chem.*, **268**, 23239–23249.
Otwinowski, Z., Schevitz, R.W., Zhang, R.-G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F. and Sigler, P.B. (1988) *Nature*, **335**, 321–329.
Schevitz, R., Otwinowski, Z., Joachimiak, A., Lawson, C.L. and Sigler, P.B. (1985) *Nature*, **317**, 782–786.
Shan, X., Gardner, K.H., Muhandiram, D.R., Rao, N.S., Arrowsmith, C.H. and Kay, L.E. (1996) *J. Am. Chem. Soc.*, **118**, 6570–6579.
Venters, R.A., Farmer II, B.T., Fierke, C.A. and Spicer, L.D. (1996) *J. Mol. Biol.*, **264**, 1101–1116.
Wishart, D.S. and Sykes, B.D. (1994) *J. Biomol. NMR*, **4**, 171–180.
Yamazaki, T., Lee, W., Revington, M., Mattiello, D.L., Dahlquist, F.W., Arrowsmith, C.H. and Kay, L.E. (1994a) *J. Am. Chem. Soc.*, **116**, 6464–6465.
Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R. and Kay, L.E. (1994b) *J. Am. Chem. Soc.*, **116**, 11655–11666.
Yang, J., Gunasekera, A., Lavoie, T.A., Jin, L. and Lewis, D.E.A. (1996) *J. Mol. Biol.*, **258**, 37–52.
Zhao, D., Arrowsmith, C.H., Jia, X. and Jardetzky, O. (1993) *J. Mol. Biol.*, **229**, 735–746.