César Fernández^a, Thomas Szyperski^a, Akira Ono^b, Hideo Iwai^a, Shin-ichi Tate^b, Masatsune Kainosho^{b,*} and Kurt Wüthrich^{a,*}

^aInstitut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Hönggerberg, CH-8093 Zürich, Switzerland

^bDepartment of Chemistry, Faculty of Science, Tokyo Metropolitan University, 1–1 Minamiohsawa, Hachioji 192-03, Japan

Received 19 December 1997; Accepted 23 January 1998

Key words: isotope labeling, nucleic acids, protein–DNA complex, sequential resonance assignment

Abstract

Nearly complete ¹H, ¹³C and ¹⁵N NMR assignments have been obtained for a doubly labeled 14-base pair DNA duplex in solution both in the free state and complexed with the uniformly ¹⁵N-labeled *Antennapedia* homeodomain. The DNA was either fully ¹³C, ¹⁵N-labeled or contained uniformly ¹³C, ¹⁵N-labeled nucleotides only at those positions which form the protein–DNA interface in the previously determined NMR solution structure of the *Antennapedia* homeodomain–DNA complex. The resonance assignments were obtained in three steps: (i) identification of the deoxyribose spin systems via scalar couplings using 2D and 3D HCCH-COSY and soft-relayed HCCH-COSY; (ii) sequential assignment of the nucleotides via ¹H–¹H NOEs observed in 3D ¹³C-resolved NOESY; and (iii) assignment of the imino and amino groups via ¹H–¹H NOEs and ¹⁵N–¹H correlation spectroscopy. The assignment of the duplex in the 17 kDa protein–DNA complex was greatly facilitated by the fact that ¹H signals of the protein were filtered out in ¹³C-resolved spectroscopy and by the excellent carbon chemical shift dispersion of the DNA duplex. Comparison of corresponding ¹³C chemical shifts of the free and the protein-bound DNA indicates conformational changes in the DNA upon complex formation.

Abbreviations: Antp(C39S), mutant *Antennapedia* homeodomain with Cys39 replaced by Ser; COSY, correlation spectroscopy; CPMG, Carr–Purcell–Meiboom–Gill; ct, constant time; DNA, deoxyribonucleic acid; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; ppm, parts per million; RNA, ribonucleic acid; 2D, two-dimensional; 3D, three-dimensional.

Introduction

Isotope labeling combined with heteronuclear NMR techniques enhances the spectral separation of the proton resonances, facilitates sequence-specific resonance assignments and enables measurements of otherwise inaccessible scalar coupling constants. The resulting increased number of conformational constraints is of special interest for nucleic acids, which yield comparatively few NOE constraints, and trans-

lates into an improvement of the quality of the NMR structures. Efficient preparation of 13 C, 15 N-doubly-labeled RNA thus enabled studies of larger RNA molecules (Puglisi et al., 1995; Ye et al., 1995; Allain et al., 1996; for reviews, see Varani et al., 1996, and Nikonowicz and Pardi, 1993). In contrast, efficient labeling protocols for DNA have been developed only more recently (Ono et al., 1994a,b; Zimmer and Crothers, 1995; Kainosho, 1997). Here, we prepared a 17 kDa protein–DNA complex comprising a 13 C, 15 N-doubly-labeled 14-base pair DNA duplex (Figure 1) bound to the uniformly 15 N-labeled

^{*}To whom correspondence should be addressed.

Experiment	Samples ^a	Max. evo	lution time (us) ^b	Numbe	r of com	olex points	s Reference
		t1,max	t2,max	t3,max	u lu	12 n ₃		I
2D [¹ H, ¹ H]-NOESY ($\tau_m = 50 \text{ ms}, \text{H}_2\text{O}, 750 \text{ MHz}$)	5	39	69		1150 2	2048		Anil Kumar et al. (1980)
2D [¹ H, ¹ H]-NOESY ($\tau_{m} = 300 \text{ ms}, D_{2}O, 750 \text{ MHz}$)	5	50	128		800	2048		Anil Kumar et al. (1980)
2D ct-[¹³ C, ¹ H]-COSY for deoxyriboses (750 MHz)	1,2,3,4	49 (C)*	114		564	1024		Vuister and Bax (1992)
2D ct-[¹³ C, ¹ H]-COSY for bases (750 MHz)	1,2,3,4	47 (C)	112		112	1024		Vuister and Bax (1992)
3D ct-HCCH-COSY (750 MHz)	3	25*	6.6 (C)*	85*	300	40 512	2	Ikura et al. (1991)
2D ct-HCCH-COSY (750 MHz)	1,2	6.6 (C)	57		60	1024		Ikura et al. (1991)
2D soft-relayed ct-HCCH-COSY ^c (750 MHz)	1,2	30 (C)	57		320	1024		Figure 2
3D triple-relayed soft ct-HCCH-COSY (750 MHz)	1,2	26	10 (C)	102	128	14 512	0	Figure 2
3D ¹³ C-resolved in-phase [¹ H, ¹ H]-COSY (500, 600 MHz)	1,2,3,4	15 (C)*	20^*	5 4*	60	85 250	10	Szyperski et al. (1998)
3D $^{13}C\text{-resolved}~[^1\text{H},^1\text{H}]\text{-}\text{NOESY}$ for deoxyriboses ($\tau_{\rm m}=50~\text{ms},\text{H}_2\text{O},750~\text{MHz})$	1,2,3,4	34*	8 (C)*	85*	180	36 512	6	Ikura et al. (1990)
3D $^{13}\text{C}\text{-resolved}$ [^1H, ^1H]-NOESY for deoxyriboses (t_m = 50 ms, H_2O, 750 MHz)	1,2,3,4	31	8 (C)	73	220	32 512	0	Ikura et al. (1990)
2D [¹⁵ N, ¹ H]-COSY (500 MHz)	1,2,3,4	58 (N)	93		260	1024		Mueller et al. (1995)
2D ¹ H-NOESY-relayed [¹⁵ N, ¹ H]-COSY with CPMG (500 MHz)	1,2,3,4	58 (N)	93		260	1024		Mueller et al. (1995)
See text for more information on the NMR experiments. ^a 1_{2}^{1} fully ¹³ C, ¹⁵ N-labeled DNA duplex; 2, partially ¹³ C, ¹⁵ N-labeled DNA duple	x; 3, fully	¹³ C, ¹⁵ N-	labeled duple	ex in the	Antp(C	39S) hon	reodomair	1 complex; 4, partially ¹³ C,
^{1.5} N-labeled duplex in the $Antp(C39S)$ homeodomain complex; 5, unlabeled DN/ ^b The type of nucleus is indicated as follows: (C). ¹³ C; (N). ¹⁵ N. No label is given	A duplex in for ¹ H che	the <i>Antp</i> (C mical shifi	39S) homeo t evolution p	domain eriods. ∤	comple; vn asteri	 sk indica 	tes that 31	P was decoupled during the

Table 1. NMR experiments recorded with the ^{13}C , ^{15}N -doubly-labeled DNA duplexes

evolution period. ^c Includes the 2D relayed, soft double-relayed and soft triple-relayed ct-HCCH-COSY experiment.



Figure 1. Survey of the NOE connectivities observed for the presently used ${}^{13}C, {}^{15}N$ -doubly-labeled DNA duplex in solution, in the free state (A) and in the *Antp(C39S)* homeodomain–DNA complex (B). Sequential NOEs (d_s) are indicated by horizontal bars, and interstrand NOEs (d_{pi}) by vertical arrows. Solid arrows are used to indicate imino proton-aromatic proton NOEs, d_{pi} (A2; T3), and broken arrows denote imino proton-amino proton NOEs, d_{pi} (G1; C4) and d_{pi} (T3; A6) (for the notation used see Wüthrich, 1986). The sequence numbering of the nucleotides is indicated at the 5' and 3' termini of each strand. Nucleotides enriched in the partially ${}^{13}C, {}^{15}N$ -doubly-labeled DNA duplex are underlined.

Antennapedia homeodomain. In addition to uniform double–labeling, we used a DNA duplex with only those nucleotides labeled that form the protein–DNA interface (Billeter et al., 1993) (Figure 1). ¹H, ¹³C and ¹⁵N resonance assignments were obtained for the labeled DNA duplex in solution both in the free state and in the homeodomain complex.

Understanding protein–DNA interactions is one of the major challenges in structural biology. The NMR solution structure of the *Antennapedia* homeodomain-DNA complex has previously been determined using uniformly ¹³C- or ¹⁵N-labeled protein and unlabeled DNA (Billeter et al., 1993; Qian et al., 1993). Additional NMR studies yielded also novel insights into the role of hydration water in protein–DNA recognition (Billeter et al., 1996; Qian et al., 1993). This well-characterized system is used here to develop techniques for NMR assignments of isotope-labeled DNA in complexes with proteins.

Materials and Methods

Preparation of the NMR samples

Fully and partially ¹³C,¹⁵N-doubly-labeled DNA oligomers were synthesized on a DNA synthesizer (Applied Biosystems Model 392–28) by the solid-phase phosphoroamidite method, using isotopically





Figure 2. Newly implemented pulse sequences: (A) 2D soft ct-(H4')C4'(C3'C2'C1')H1'-COSY (the number of relay steps is n = 2); (B) 2D soft ct-(H5',5"(C5'(C4'C3'C2'C1')H1'-COSY, n = 3, which differs from (A) in the selection scheme used for the initial ¹H-¹³C INEPT; and (C) 3D soft ct-H5', 5''(C5'C4'C3'C2')C1'H1'-COSY and 3D soft ct-H5', 5''C5'(C4'C3'C2'C1')H1'-COSY, n = 3. Rectangular 90° and 180° pulses are indicated by thin and thick black vertical bars, respectively, and phases are indicated above the pulses. Where no radio-frequency phase is marked, the pulse is applied along x. The $[90^{\circ}_{x} 180^{\circ}_{y} 90^{\circ}_{x}]$ -composite ¹³C-pulses are represented by triplets of bars. For the ¹H pulses, the carrier is placed at the position of the water line at 4.64 ppm, and the ¹³C carrier for rectangular pulses was set to 65 ppm for schemes (A) and (B), and to 77.5 ppm for scheme (C). The selective E-BURP1 excitation pulse (Geen and Freeman, 1991) on 1 H in (A) (pulse length 1.3 ms) is placed near the center of the 1 H4' region at 4.1 ppm, and the selective I-BURP inversion pulses (Geen and Freeman, 1991) on 13 C in (B) and (C) (pulse length 1.9 ms) were placed near the center of the ${}^{13}C5'$ signals at 65 ppm. The durations and amplitudes of the sine-bell-shaped PFGs are 400 µs and 10 G/cm for G1, 2 ms and 32 G/cm for G2, 1 ms and 32 G/cm for G3, and 400 µs and 15 G/cm for G4. The recovery delays after G_2 and G_3 are set to 1 ms. The delays have the following values: $\tau_1 = 1.7$ ms, $\tau_2 = 2.1$ ms, $\tau_3 = 3.3$ ms, $\tau_4 = 5.7$ ms, $\delta = 23 \text{ ms. In (C)}, \tau_a = \tau'_a = \tau_4, \tau_b = \tau_4 + t_2/2 \text{ and } \tau'_b = \tau_4 - t_2/2 \text{ are used for 3D soft ct-H5'}, 5''(C5'C4'C3'C2')C1'H1'-COSY, while \\ \tau_a = \tau_4 + t_2/2, \tau_{\prime a} = \tau_4 - t_2/2 \text{ and } \tau_b = \tau'_b = \tau_4 \text{ are used for 3D soft ct-H5'}, 5''C5'(C4'C3'C2')H1'-COSY. A DIPSI-2 sequence$ with rf = 2.5 KHz (Shaka et al., 1988) is used to decouple ¹H during the heteronuclear magnetization transfers and a GARP sequence with rf = 3.1 kHz (Shaka et al., 1985) is used to decouple ¹³C during proton detection. Phase cycling: $\phi_1 = x$; $\phi_2 = 4 \{x\}, 4\{y\}, 4\{-x\}, 4\{-y\};$ is accomplished by alternating the phases ϕ_1 and ϕ_9 according to States-TPPI (Marion et al., 1989). The same technique is used in 3D soft ct-H5',5"C5'(C4'C3'C2'C1')H1'-COSY, except that ϕ_1 and ϕ_5 are alternated instead of ϕ_1 and ϕ_9 .

labeled monomer units that had been synthesized according to a previously described strategy (Ono et al., 1994b). Approximately 1 μ mol of oligomer was obtained from 5 μ mol of nucleoside bound to the resin, and the purity of the labeled oligomers was higher than 99% as estimated by HPLC analysis on a C₁₈ column (Inertsil ODS-2, GL Science). Uniformly ¹⁵N-labeled *Antp(C39S)* homeodomain was produced as described by Müller et al. (1988). Preparation of the 1:1 complex between the Antp(C39S) homeodomain and the isotopically labeled DNA was performed as described by Otting et al. (1990).

NMR samples of the free DNA duplex at a concentration of about 2 mM were prepared in 90% H₂O/10% D₂O or in 99.98% D₂O containing 50 mM potassium phosphate and 20 mM KCl at pH = 6.0. Samples of



Figure 3. Contour plots from 2D [^{13}C , ^{1}H]-COSY spectra of the fully ^{13}C , ^{15}N -labeled DNA duplex. (A) spectra recorded for studies of the deoxyriboses; (B) spectra acquired for studies of the bases. The inset in (A) shows an expansion of the C1'–H1' region, where the cross peaks are labeled according to their sequence positions (see Figure 1). The dotted rectangles identify spectral regions corresponding to the specified CH_n moieties. The spectra were recorded on a Varian Unity+ spectrometer operating at a proton frequency of 750 MHz.

the *Antp*(*C39S*) homeodomain–DNA complex at concentrations of about 1.5 mM were prepared in 90% $H_2O/10\%$ D₂O or 99.98% D₂O at pH = 6.0.

NMR spectroscopy

NMR measurements were performed at $T = 36 \degree C$ on Bruker DRX500 and DRX600 spectrometers and a Varian U750+ spectrometer equipped with ¹H- $\{^{13}C, ^{15}N\}$ and $^{1}H-\{^{13}C, ^{31}P\}$ triple-resonance probeheads. Table 1 affords a survey of the NMR experiments performed for the present studies. Quadrature detection in the indirect dimensions was achieved using States-TPPI (Marion et al., 1989). The ¹H carrier position was set to the water line at 4.64 ppm, and the ¹⁵N carrier was placed at 91 ppm. The ¹³C carrier was set to 65 ppm in 2D experiments recorded for deoxyriboses, to 77.5 ppm in 3D experiments recorded for deoxyriboses in order to fold the ¹³C2' signals, and to 145 ppm in experiments performed for the bases. ¹H chemical shifts are relative to internal DSS, and ¹⁵N and ¹³C chemical shifts are relative to DSS calculated with the Ξ values reported by Wishart et al. (1995). For data processing and spectral analysis, we used the programs PROSA (Güntert et al., 1992) and XEASY (Bartels et al., 1995), respectively.

Newly implemented multiple-relayed soft HCCH-COSY experiments (Figure 2) were recorded to identify the deoxyribose spin systems of the free ¹³C,¹⁵Nlabeled DNA. In the 2D experiments, high resolution was attained in the indirect ¹³C dimension by extending the ct carbon evolution period by a complete ¹J(¹³C, ¹³C) dephasing/rephasing cycle (Szyperski et al., 1997a). These experiments exhibited good sensitivity for the free DNA duplex. The application of soft pulses for the selective excitation of magnetization in the initial INEPT transfer and proper tuning of the delays for the carbon-carbon magnetization transfer ensured selective observation of relayed correlations. In particular, double-relayed 2D soft ct-(H4')C4'(C3'C2'C1')H1'-COSY was recorded to assign the ¹³C4' resonances (Figure 2A). Application of an E-BURP1 pulse (Geen and Freeman, 1991) leads to selective excitation of the H4' protons in order to obtain selectively only the C4'-H1' correlations and suppress the generation of direct C1'-H1' correlations, which would overlap with the C4'-H1' cross peaks. The ¹³C5' and ¹H5', ¹H5' chemical shifts were measured in triple-relayed HCCH-COSY, i.e., 2D soft ct-(H5',5")C5'(C4'C3'C2'C1')H1'-COSY (Figure 2B), 3D soft ct-H5',5'(C5'C4'C3'C2')C1'H1'-



COSY and 3D soft ct-H5',5'C5'(C4'C3'C2'C1')H1'-COSY (Figure 2C). The pulse sequence of Figure 2C is a 3D extension of the experiment shown in Figure 2B, where the ct ¹³C chemical shift evolution takes place either for ¹³C1' or for ¹³C5'. Here, the I-BURP1 pulse (Geen and Freeman, 1991) employed during the initial INEPT step ensures selective observation of the H5',H5"-H1' correlations.

Assignment strategy

The ¹H, ¹³C and ¹⁵N chemical shift assignments of the ¹³C, ¹⁵N-doubly-labeled DNA duplexes were performed in three steps. First, the deoxyribose spin systems were identified in HCCH-COSY spectra. The H2' and H2'' protons were stereospecifically assigned based on the intensity of the $d_i(1'; 2', 2'')$ NOEs observed in a 3D ¹³C-resolved [¹H,¹H]-NOESY spectrum recorded with a short mixing time (see Wijmenga et al., 1993), and the values for the ³J_{1'2'} and ³J_{1'2'} scalar coupling constants measured in 3D

¹³C-resolved in-phase [¹H,¹H]-COSY spectra (Grzesiek et al., 1995; Szyperski et al., 1998). Second, the base C2H, C5H, C6H and C8H moieties were distinguished according to their characteristic and largely conformation-independent carbon chemical shifts that were obtained using [13C,1H]-COSY. C5H and C6H could independently also be identified from observation of the ${}^{1}J_{C5C6}$ scalar coupling. Third, we employed the conventional sequential assignment protocol (Feigon et al., 1983; Frechet et al., 1983; Haasnoot et al., 1983; Hare et al., 1983; Chazin et al., 1986) based on the observation of intranucleotide $d_i(2',2'';$ 6.8) and $d_i(1'; 6.8)$ NOEs and sequential $d_s(2',2'';$ 6.8) and $d_s(1'; 6.8)$ NOEs in ¹³C-resolved [¹H,¹H]-NOESY (for the notation, see Wüthrich, 1986). The sequential walk was started at the 5' and 3' terminal nucleotides, which could be readily identified from their unique C5'H and C3'H chemical shifts, and because the 5' terminal C6H or C8H groups exhibit only intranucleotide NOEs to deoxyribose protons



Figure 4. Contour plots of different spectral regions taken from the following experiments: (A) 2D ct- $[^{13}C, ^{1}H]$ -COSY. (B) 2D ct-(H)C(C)H-COSY. (C) 2D ct-(H)C(CC)H-COSY. (D) 2D soft ct-(H4')C4'(C3'C2'C1')H1'-COSY; and (E) 2D soft ct-(H5',5'')C5'(C4'C3'C2'C1')H1'-COSY. All spectra were recorded with the fully ^{13}C , ^{15}N -labeled DNA duplex. The five spectral regions show the correlations between H1' and C1', C2', C3', C4' and C5', respectively. Note that the lower resolution in the ^{13}C dimension in (B) is due to the use of a shorter ^{13}C evolution time (Table 1).

(Wüthrich, 1986). The thymidine methyl groups were assigned through observation of intranucleotide $d_i(T6; TCH_3)$ NOEs. The sequential assignment of X–T and X–C dinucleotides was then confirmed by observation of $d_s(6,8; TCH_3)$ and $d_s(6,8; C5)$ NOEs, respectively. Finally, amino protons, imino protons and C2H of A were assigned in 3D ¹³C-resolved [¹H,¹H]-NOESY and 2D [¹H,¹H]-NOESY spectra recorded in H₂O. Sequential imino proton–imino proton NOEs were de-

tected in a 2D [1 H, 1 H]-NOESY spectrum recorded in H₂O, and the imino 15 N chemical shifts were obtained from 2D [15 N, 1 H]-COSY. Since the amino proton resonance lines of G and A were broadened by exchange processes, we used 2D 1 H-NOESY-relayed [15 N, 1 H]-COSY implemented with a CPMG sequence for heteronuclear magnetization transfer (Mueller et al., 1995) to detect additional NOEs to the NH₂ groups.

Results

Identification of the deoxyribose spin systems

Identification of the deoxyribose spin systems for the free DNA duplex was based on 2D [¹³C,¹H]-COSY (Figure 3), 2D HCCH-COSY and multiple-relayed 2D and 3D HCCH-COSY (Figures 4 and 5; Table 1). The 2D experiments were used to achieve high resolution in the indirect ¹³C dimension, which would not be feasible with the 3D experiments within a reasonable measurement time. Figure 4 shows spectral regions taken from the spectra that were used to assign the ${}^{13}C2'$, ${}^{13}C3'$, ${}^{13}C4'$ and ${}^{13}C5'$ resonances via through-bond correlation to the ¹H1' resonances, which exhibit excellent dispersion. Complete resonance assignments were obtained for the deoxyriboses of the free DNA duplex, with the sole exception of six H5', H5" proton pairs. The spectral overlap in the C5'-H5', H5'' region could be resolved using the 3D soft ct-H5',5"(C5'C4'C3'C2')C1'H1' experiment. Figure 5A shows $[\omega_1({}^1\text{H5}', 5''), \omega_3({}^1\text{H1}')]$ strips from this experiment taken in the ${}^{13}C1'$ plane, which provide correlations between H1', C1' and H5',H5". To obtain the C5' assignments, we recorded a 3D soft ct-H5',5"C5'(C4'C3'C2,C1')H1'-COSY experiment with frequency labeling during t₂ on C5' instead of C1' (Figure 5B). The two 3D spectra yielded complete assignments of the C5'H2 groups for the free DNA duplex (chemical shifts have been deposited in the Bio-MagResBank (http://www.bmrb.wisc.edu; accession number: 4104)).

For the 17 kDa *Antp(C39S)*–DNA complex, the multiple-relay experiments turned out to have low sensitivity because of the short $T_2(^{13}C)$ relaxation times. We therefore used 3D ct-HCCH-COSY (Ikura et al., 1991) and 3D ^{13}C -resolved in-phase [^{1}H , ^{1}H]-COSY (Grzesiek et al., 1995) for the deoxyribose assignments, where the latter experiment was primarily recorded to determine the $^{3}J_{HH}$ scalar couplings of the sugar rings (Szyperski et al., 1998). We also



Figure 5. $[\omega_1(^1H), \omega_3(^1H)]$ strips taken from 3D soft triple-relayed spectra recorded for the assignment of the C5'H₂ moieties. The strips (A) are taken at the ¹³C1' chemical shifts (indicated at the top of the strips) from the 3D soft ct-H5',5''(C5'C4'C3'C2')C1'H1'-COSY spectrum. The strips (B) are taken at the ¹³C5' chemical shifts in the 3D soft ct-H5',5''C5'(C4'C3'C2,C1')H1'-COSY spectrum. Both spectra were recorded with the fully ¹³C,¹⁵N-labeled free DNA duplex. The strips are centered about the H1' chemical shifts along ω_3 , and the H5' and H5'' chemical shifts were measured along ω_1 . The nucleotide positions are indicated at the top of the strips.

made reference to the chemical shift assignment of the free DNA duplex, and eventually complete proton and carbon assignments were obtained for all deoxyriboses of the protein-bound DNA, except for the C5'– H5',H5" resonances of T₁₈, A₁₉ and A₂₀ (Tables 2 and 3). In these nucleotides, chemical shift degeneracy of H5', H5' and H4' prevented resonance assignments based on 3D HCCH-COSY and ¹³C-resolved [¹H,¹H]-NOESY.

Sequential resonance assignment

In both the free DNA duplex (Figure 1A) and the *Antp*(*C39S*)–DNA complex (Figure 1B) a complete set of sequential $d_s(2',2'; 6,8)$ and $d_s(1'; 6,8)$ NOEs was identified in the 3D ¹³C-resolved [¹H,¹H]-NOESY spectra. The $d_s(6,8; TCH_3)$ and $d_s(6,8; C5)$ sequential NOEs for X–T and X–C dinucleotides, respectively, were also observed, with the sole exception that the $d_s(6,8; C5)$ NOE for the dinucleotide T₂₇–C₂₈ could not be detected in the complex. To illustrate the high quality of the data used for sequential resonance assignment via ¹H–¹H NOEs, Figure 6 shows [$\omega_1(^1H)$,

 $\omega_3(^1\text{H})$] strips taken at the C1' and C6/C8 carbon chemical shifts of the first six nucleotides from the 3D ^{13}C -resolved [$^1\text{H}, ^1\text{H}$]-NOESY spectra acquired with the *Antp*(*C39S*)–DNA complex.

Sequential assignment of imino and amino groups

Assignments of the imino and amino protons were obtained from 2D [¹⁵N,¹H]-COSY, 2D [¹H,¹H]-NOESY and 2D 1H-NOESY-relayed [15N,1H]-COSY recorded with a CPMG sequence for the heteronuclear magnetization transfer. The G H1 proton chemical shifts were inferred from observation of interstrand NOEs to the amino protons H41 and H42 of the base-paired C, which in turn are connected by an intranucleotide NOE to C H5. The interstrand connectivities d_{pi}(G1; C4) could be observed for all nonterminal C \equiv G base pairs. Based on these connectivities, G H1, C H41 and C H42 for the free DNA and the Antp(C39S)-DNA complex could be assigned for all except the terminal base pairs, where fraying of the helical ends leads to rapid exchange of the labile protons with the solvent (Figure 1). Assignment of the T H3 and A H2 base



Figure 6. Contour plots of $[\omega_1({}^1\text{H}), \omega_3({}^1\text{H})]$ -strips taken from 3D ${}^{13}\text{C}$ -resolved $[{}^1\text{H}, {}^1\text{H}]$ -NOESY spectra ($\tau_m = 50$ ms) recorded for the *Antp(C39S)* homeodomain complex with the fully ${}^{13}\text{C}, {}^{15}\text{N}$ -labeled DNA duplex. The strips labeled at the bottom with H1' are taken from the spectrum recorded for studies of the deoxyriboses (Table 1) and are centered along ω_3 about the H1' chemical shifts of the first six nucleotides of the upper strand in Figure 1. The strips labeled with H6 or H8 are taken from the spectrum recorded for studies of the bases and are centered about the H6 or H8 chemical shifts, respectively. The corresponding ${}^{13}\text{C1'}$ and ${}^{13}\text{C6}$ or ${}^{13}\text{C8}$ chemical shifts are indicated at the top of each strip. As an illustration, cross peaks representing intranucleotide NOEs with ${}^1\text{H1'}$ of G1 are identified on the left of the figure. Cross peaks corresponding to the intranucleotide H1'-H2' NOEs are enclosed in solid rectangles, and the sequential H2'-H6/H8 and H2''-H6/H8 NOE connectivities are indicated with arrows.

protons of A=T base pairs was achieved via interstrand d_{pi}(T3; A2) and intranucleotide d_i(T3; TCH₃) NOEs. The d_{pi}(T3; A2) NOE connectivities were observed for all A=T base pairs in the free DNA duplex and the Antp(C39S)-DNA complex (Figure 1), yielding complete resonance assignments for T H3 and A H2. For the centrally located base pairs $A_8=T_{21}$, T₉=A₂₀, T₁₀=A₁₉ and A₁₁=T₁₈, interstrand NOEs between the imino protons H3 of T and the amino protons H61 and H62 of A, d_{pi}(T3; A6), could be detected in the 2D ¹H-NOESY-relayed [¹⁵N,¹H]-COSY spectrum. The A H61 and A H62 chemical shifts could not be obtained from these data because of the small ¹⁵N and ¹H chemical shift dispersion of the adenine amino groups in the free DNA duplex, but the H61 and H62 protons of A8, A11, A19 and A_{20} could be assigned in the Antp(C39S)–DNA complex. Two intranucleotide imino proton-amino proton

NOEs $d_i(G1; G2)$ could be detected for G_{22} and G_{23} . The ¹⁵N-¹H and ¹³C2-¹H2 correlations were then established in 2D [¹⁵N,¹H]-COSY and 2D [¹³C,¹H]-COSY spectra, respectively. As an illustration of the high quality data obtained, Figure 7A shows the region of the 2D ¹H-NOESY-relayed [¹⁵N,¹H]-COSY spectrum recorded with CPMG that comprises the interstrand amino proton-imino proton contacts and shows how the assignment of the amino proton resonances of C and A was performed. This spectrum was essential for identifying the A H61 and A H62 chemical shifts, since most of these resonance lines could not be observed by 2D [¹H,¹H]-NOESY (Figure 7B) due to the exchange broadening arising from the rotation of the NH₂ group about the C-N bond (Wijmenga et al., 1993).



Figure 7. (A) 2D ¹H-NOESY-relayed [15 N,¹H]-COSY spectrum of the *Antp(C39S)* homeodomain–DNA complex containing fully 13 C,¹⁵N-labeled DNA recorded using a CPMG pulse train for heteronuclear magnetization transfer (Mueller et al., 1995). This experiment was used for the detection of NOEs with exchange-broadened amino protons. Cross peaks belonging to direct 15 N–¹H correlations are enclosed in solid rectangles that are labeled with the type of the NH_n moiety. The pattern of cross peaks representing interstrand amino proton–imino proton NOEs is illustrated with two examples: dashed lines identify the connectivity between the H6 amino protons of A₁₉ and the H3 imino proton of T₁₀, and dot-dashed lines indicate the connectivity between the H4 amino protons of C₂₄ and the H1 imino proton of G₅. Negative contours, indicating cross peaks folded once in the ¹⁵N dimension, are drawn with broken contour lines, and a correspondingly adapted chemical shift scale is shown on the left side of the figure. (B) Expanded region of the 2D [1 H, ¹H]-NOESY spectrum recorded for the *Antp(C39S)* homeodomain–DNA complex, showing NOEs to the imino protons of T. The assignments of the interstrand aromatic proton–imino proton NOEs, d_{pi}(A2; T3), are indicated by the sequence numbers at the top and on the left side of the spectrum. Cross-peak positions corresponding to amino proton–imino proton NOEs, d_{pi}(A6; T3), observed in (A) are indicated by circles in this spectrum.

Table 2. ¹H chemical shifts in ppm^a of the DNA duplex bound to Antp(C39S) homeodomain in H₂, pH 6.0, 36°.

Residue	H1'	H2′, H2″	H3′	H4′	H5′,H5″	Base protons
G ₁	5.46	2.27, 2.47	4.72	4.08	3.61, 3.61	H8 7.74
A ₂	5.85	2.69, 2.82	4.99	4.33	4.08, 3.98	H8 8.13; H2 7.40
A ₃	5.84	2.58, 2.81	5.00	4.38	4.32, 4.15	H8 7.97; H2 7.16
<u>A</u> 4	6.04	2.58, 2.86	4.80 +	4.44	4.33, 4.24	H8 7.89; H2 7.36
<u>G</u> 5	5.84-	2.38, 2.70	5.07 -	4.31	4.19, 4.17	H8 7.38; H1 12.89 -
<u>C</u> 6	6.06-	1.75 +, 2.54	4.83	4.25	4.06, 4.00	H6 7.14; H5 4.66 +; H41 8.32 -; H42 5.72 +
<u>C</u> 7	5.28+	1.80 +, 2.26	4.70	3.92	4.06, 4.01	H6 7.24+; H5 5.39; H41 8.35; H42 5.77+
A ₈	6.29	2.60, 2.95	4.95	4.41	4.07, 4.01	H8 8.28; H2 7.56; H61 7.90; H62 6.17
Т9	5.54+	1.68 +, 2.35	4.66	4.00	4.32, 4.07	H6 6.86 +; C7H ₃ 1.41; H3 13.44
T ₁₀	5.68	1.94, 2.23	4.81	4.01	4.01, 3.85	H6 7.13+; C7H ₃ 1.48; H3 13.41
A ₁₁	5.49 +	2.65, 2.70	4.95	4.10 +	3.86 +, 3.90	H8 8.28; H2 6.86+; H61 7.18; H62 6.61
<u>G</u> ₁₂	5.41	2.63, 2.68	4.97	4.11 +	4.24, 3.88 +	H8 7.78; H1 12.39
<u>A</u> 13	5.99	2.59, 2.81	4.99	4.29	4.20, 3.95 +	H8 8.00; H2 7.73
G14	5.95	2.34, 2.20	4.56	4.09	4.20, 4.09	H8 7.58
C ₁₅	5.82	2.13, 2.50	4.59	4.04	3.81, 3.73	H6 7.81; H5 5.89
T ₁₆	6.06	2.13, 2.48	4.81	4.17	4.00, 4.03	H6 7.52; C7H ₃ 1.58; H3 13.66
<u>C</u> 17	5.91	1.90 +, 2.26 +	4.72	4.09	3.97 +, 3.88 +	H6 7.47; H5 5.56; H41 8.22 ; H42 6.98
T ₁₈	4.69 +	2.10, 2.22 +	4.70	3.94 +		H6 7.47; C7H ₃ 1.66; H3 13.92 -
<u>A</u> 19	6.14-	2.93-, 3.03-	4.89	4.33		H8 8.15; H2 7.07-; H61 6.82; H62 6.19
<u>A</u> 20	5.81 +	2.31 +, 2.75	5.01	4.25		H8 7.44+; H2 7.45; H61 6.95; H62 6.83
<u>T</u> ₂₁	5.72	1.65, 2.28	4.70	3.99	3.95 +, 3.89 +	H6 6.68+; C7H3 0.94+; H3 13.21
G22	5.56	2.55, 2.61	4.91	4.25	4.13, 4.09	H8 7.72; H1 12.46
G ₂₃	5.84	2.50, 2.65	4.88	4.31	4.14, 4.08	H8 7.61; H1 12.65; H21 6.72
C ₂₄	5.87	2.06, 2.48	4.63	4.18	4.10, 4.03	H6 7.30; H5 5.21; H41 7.96; H42 6.34
T ₂₅	6.00	2.10, 2.55	4.80	4.16	4.07, 4.07	H6 7.36; C7H ₃ 1.48; H3 13.86
T ₂₆	6.13	2.17, 2.57	4.84	4.15	4.17, 4.17	H6 7.43; C7H ₃ 1.60; H3 13.73 +
T ₂₇	6.14	2.16, 2.47	4.85	4.12	4.14, 4.03	H6 7.40; C7H ₃ 1.66; H3 13.77 +
C ₂₈	6.20	2.20, 2.20	4.50	3.98	4.13, 4.01	H6 7.63; H5 5.85

^a The chemical shifts are relative to internal DSS. Where stereospecific assignments of diastereotopic pairs of 2', 2'' protons were obtained, these are printed in italics, and the first number is the chemical shift of the 2' proton. Labeled nucleotides in the partially labeled samples are underlined. Numbers are in bold if $|\Delta\delta$ (free DNA-complex)| > 2.0 ppm, where shifts to higher field relative to the free DNA are indicated with '+' and shifts to lower field with '-'. The chemical shifts have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu; accession number: 4104).

Discussion

Double labeling of the DNA with ¹³C and ¹⁵N combined with the presently applied assignment strategy yielded nearly complete ¹H, ¹³C and ¹⁵N resonance assignments for the free and the protein-bound DNA (Tables 2 and 3). The resonance assignment of the DNA duplex in the 17 kDa homeodomain–DNA complex was greatly facilitated by the large ¹³C chemical shift dispersion and the fact that the protein proton signals were filtered out in ¹³C-resolved spectroscopy. Since for the imino groups the ¹⁵N chemical shift dispersion is very limited (Figure 7A), the major benefit from ¹⁵N labeling arises from observation of imino proton–amino proton NOE connectivities in 2D ¹H-NOESY–relayed [¹⁵N,¹H]-COSY (Mueller et al., 1995). With regard to the structure determination, it is of interest that the ¹³C labeling of the DNA allowed efficient measurement of ${}^{3}J_{CP}$ and ${}^{3}J_{HH}$ coupling constants for the DNA backbone and the deoxyribose rings, respectively (Szyperski et al., 1997b, 1998).

The standard NOE-based sequential assignment method is robust for B-type DNA duplexes (Wüthrich, 1986) and allows sequential resonance assignments with a small set of spectra (Table 1). For the present project, there was thus no need to record triple-resonance NMR experiments for the sequential assignment of the DNA duplex via through-bond correlations, which is of interest since the short $T_2(^{31}P)$ relaxation times at higher field strengths limit through-bond

Residue	C1′	C2′	C3′	C4′	C5′	base ¹³ C and ¹⁵ N
G ₁	85.7	39.9	79.4	89.4	64.4	C8 139.0
A_2	85.0	40.0	80.2	88.1	68.3	C8 141.2; C2 154.1
A ₃	84.1	40.7	78.7	87.2	68.1	C8 140.6; C2 153.6
\underline{A}_4	85.4	42.2	76.6 +	86.1	68.1	C8 140.1; C2 153.7
\underline{G}_5	85.5 -	42.8 -	76.2 +	85.5 +	67.9	C8 136.0+; N1 147.0; N2 75.6
<u>C</u> 6	87.2	40.5	80.0 -	86.8-	67.5	C6 142.7; N4 98.6
<u>C</u> 7	85.4 +	39.8	74.9 +	84.6 +	66.2	C6 142.8; C5 98.1; N4 98.1
A ₈	85.1	40.7	80.2	88.1 -	68.3	C8 141.8; C2 154.4; N6 81.7-
Т9	85.0	38.9	75.4 +	84.6	65.7	C6 137.5; C7 14.3; N3 159.1
T ₁₀	85.3	38.6	78.9 -	86.3	67.9	C6 139.7; C7 14.7; N3 158.5
A ₁₁	84.6	40.0	78.4 +	87.6	68.3	C8 141.9; C2 153.2; N6 79.9
<u>G</u> ₁₂	84.0	40.0	79.4 -	87.6	69.2	C8 137.5; N1 146.7
<u>A</u> 13	85.0	40.6	79.6	87.5	68.7	C8 140.9; C2 154.8
G14	84.5	42.7	72.3	87.6	67.5	C8 138.2
C ₁₅	88.5	40.8	75.9	88.1	63.1	C6 143.3; C5 98.9
T ₁₆	86.4	39.1	78.7 +	86.8	67.9	C6 139.7; C7 14.5; N3 158.7
<u>C</u> 17	85.8	38.6	78.4-	85.7	68.3-	C6 143.2; C5 99.0; N4 98.0
T ₁₈	86.2	37.7 +	76.8 +	86.9		C6 140.0; C7 14.4; N3 160.4-
<u>A</u> 19	85.4	39.9	79.2	87.3		C8 142.7; C2 153.0; N6 77.4+
<u>A</u> 20	82.8 +	42.3	76.1 +	86.3		C8 139.0+; C2 154.4; N6 81.5
<u>T</u> ₂₁	85.8	40.5	76.5	85.9	68.3-	C6 137.2; C7 14.8; N3 158.4
G ₂₂	84.1	40.3	79.3	87.2	68.8	C8 138.2; N1 147.0; N2 75.5
G ₂₃	84.6	41.0	79.3	87.2	68.8	C8 137.2; N1 147.0; N2 75.3
C ₂₄	86.8	40.3	76.8	85.9	67.9	C6 142.2; C5 97.7; N4 97.0
T ₂₅	86.3	39.4	79.8 -	86.3	67.5	C6 139.5; C7 14.6; N3 160.4-
T ₂₆	85.4	39.2	78.0	85.9	66.6	C6 140.0; C7 14.5; N3 160.4-
T ₂₇	85.4	39.4	77.0	85.9	66.6	C6 140.0; C7 14.4; N3 160.5-
C ₂₈	87.2	41.2	72.1	87.2	67.0	C6 144.7; C5 98.9

Table 3. ¹³C and ¹⁵N chemical shifts in ppm^a of the DNA duplex bound to the *Antp(C39S)* homeodomain in H₂O, pH 6.0, 36 °C.

^a The chemical shifts are relative to internal DSS with the following Ξ ratios: ${}^{13}C/{}^{1}H = 0.251449530$, ${}^{15}N/{}^{1}H = 0.101329118$ (Wishart et al., 1995). Labeled nucleotides in the partially labeled samples are underlined. Numbers are in bold if $|\Delta\delta|$ (free DNA–complex)| > 1.0 ppm, where shifts to higher field relative to the free DNA are indicated with '+' and shifts to lower field with '-'. The chemical shifts have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu; accession number: 4101).

sequential resonance assignment protocols (Pardi et al., 1983) to smaller systems (Varani et al., 1996). However, in the absence of a regular helical conformation, e.g., in protein–DNA complexes in which the DNA duplex is highly distorted, or for single-strand DNA and its protein complexes (Omichinski et al., 1997), additional efforts would be needed to establish intranucleotide links between deoxyribose and base via scalar couplings (for a review, see Dieckmann and Feigon, 1994).

Since the deoxyribose ¹³C chemical shifts are sensitive to local conformational changes (e.g., Ghose et al., 1994; Oldfield, 1995), they enable an initial characterization of variations in the DNA conformation upon binding to a protein. In the Antp(C39S)–DNA complex, the major chemical shift changes are observed at the protein—DNA interface (Tables 2 and 3). Additional small chemical shift differences suggest that parts of the DNA duplex that do not directly contact the protein are also subject to subtle conformational changes upon complex formation. A more precise description of these conformational features will result from the ongoing refinement of the NMR structure of the *Antp* homeodomain–DNA complex, which makes use of additional data resulting from the use of the ¹³C, ¹⁵N-labeled DNA, i.e., an extensive collection of vicinal ${}^{1}H{-}^{1}H$ and heteronuclear coupling constants (Szyperski et al., 1997b, 1998), and a significantly increased number of NOE upper distance constraints when compared to the previous data collection with unlabeled DNA (Billeter et al., 1993).

Acknowledgements

Financial support was obtained from the Schweizerischer Nationalfonds (Project 31.49047.96), from the ETH Zürich for a special grant within the framework of the Swiss-Japanese R&D Roundtable Collaboration, and by Crest (Core Research for Evolutional Science and Technology) of the Japan Science and Technology Corporation (JST). C.F. is indebted to the 'Schweizerische Bundesstipendienkommission' for a fellowship. We thank Mrs. E. Ulrich for careful processing of the manuscript.

References

- Allain, F.H.T., Gubser, C.C., Howe, P.W.A., Nagai, K., Neuhaus, D. and Varani, G. (1996) *Nature*, **380**, 646–650.
- Anil Kumar, Ernst, R.R. and Wüthrich, K. (1980) Biochem. Biophys. Res. Commun, 95, 1–6.
- Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, 6, 1–10.
- Billeter, M., Qian, Y.Q., Otting, G., Müller, M., Gehring, W. and Wüthrich, K. (1993) J. Mol. Biol., 234, 1084–1097.
- Billeter, M., Güntert, P., Luginbühl, P. and Wüthrich, K. (1996) *Cell*, **85**, 1057–1065.
- Bodenhausen, G. and Ruben, D. (1980) Chem. Phys. Lett., 69, 185– 188.
- Chazin, W.J., Wüthrich, K., Hyberts, S., Rance, M., Denny, W.A. and Leupin, W. (1986) *J. Mol. Biol.*, **190**, 439–453.
- Dieckmann, T. and Feigon, J. (1994) Curr. Opin Struct. Biol., 4, 745–749.
- Feigon, J., Leupin, W., Denny, W.A. and Kearns, D.R. (1983) *Biochemistry*, 22, 5943–5951.
- Frechet, D., Cheng, D.M., Kan, L.S. and Ts'o, P.O.P. (1983) *Biochemistry*, 22, 5194–5200.
- Geen, H. and Freeman, R. (1991) J. Magn. Reson., 93, 93-96.
- Ghose, R., Marino, J.P., Wiberg, K.B. and Prestegard, J.H. (1994)J. Am. Chem. Soc., 116, 8827–8828.
- Grzesiek, S., Kuboniwa, H., Hinck, A.P. and Bax, A. (1995) J. Am. Chem. Soc., 117, 5312–5315.
- Güntert, P., Dötsch, V., Wider, G. and Wüthrich, K. (1992) J. Biomol. NMR, 2, 619–629.
- Haasnoot, C.A.G., Westering, H.P., van der Marel, G.A. and van Boom, J.H. (1983) *J. Biomol. Struct. Dyn.*, **1**, 131–149.
- Hare, D.R., Wemmer, D.E., Chou, S.H., Drobny, G. and Reid, B.R. (1983) J. Mol. Biol., 171, 319–336.

- Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) J. Magn. Reson., 86, 204–209.
- Ikura, M., Kay, L.E. and Bax, A. (1991) J. Biomol. NMR, 1, 299– 304.
- Kainosho, M. (1997) Nat. Struct. Biol., 4, 858-861.
- Legault, P., Jucker, F.M. and Pardi, A. (1995) *FEBS Lett.*, **362**, 156–160.
- Marion, D., Ikura, K., Tschudin, R. and Bax, A. (1989) J. Magn. Reson., 85, 393–399.
- Mueller, L., Legault, P. and Pardi, A. (1995) J. Am. Chem. Soc., 117, 11043–11048.
- Müller, M., Affolter, M., Leupin, W., Otting, G., Wüthrich, K. and Gehring, W.J. (1988) *EMBO J.*, **7**, 4299–4304.
- Nikonowicz, E.P. and Pardi, A. (1993) J. Mol. Biol., 232, 1141– 1156.
- Oldfield, E. (1995) J. Biomol. NMR, 5, 217-225.
- Omichinski, J., Pedone, P.V., Felsenfeld, G., Gronenborn, A. and Clore, G.M. (1997) Nat. Struct. Biol., 122–132.
- Ono, A., Tate, S. and Kainosho, M. (1994b) In Stable Isotope Applications in Biomolecular Structure and Mechanisms (Eds., Trewhella, J., Cross, T.A. and Unkefer, C.J.), Los Alamos National Laboratory New Mexico, pp. 127–144.
- Ono, A., Tate, S., Ishido, Y. and Kainosho, M. (1994a) J. Biomol. NMR., 4, 581–586.
- Otting, G., Qian, Y.Q., Billeter, M., Müller, M., Affolter, M., Gehring, W.J. and Wüthrich, K. (1990) *EMBO J.*, 9, 3085–3092.
- Pardi, A., Walker, R., Rapaport, H., Wider, G. and Wüthrich, K. (1983) J. Am. Chem. Soc., 105, 1652–1653.
- Puglisi, J.D., Chen, L., Blanchard, S. and Frankel, A.D. (1995) Science, 270, 1200–1203.
- Qian, Y.Q., Otting, G., Billeter, M., Müller, M., Gehring, W. and Wüthrich, K. (1993) J. Mol. Biol., 234, 1070–1083.
- Santoro, J. and King, G.C. (1992) J. Magn. Reson., 97, 202-207.
- Shaka, A.J., Barker, P.B. and Freeman, R. (1985) J. Magn. Reson., 64, 547–552.
- Shaka, A.J., Lee, C.J. and Pines, A.J. (1988) J. Magn. Reson., 77, 274–293.
- Szyperski, T., Fernández, C. and Wüthrich, K. (1997a) J. Magn. Reson., 128, 228–232.
- Szyperski, T., Ono, A., Fernández, C., Iwai, H., Tate, S., Wüthrich, K. and Kainosho, M. (1997b) J. Am. Chem. Soc., 119, 9901– 9902.
- Szyperski, T., Fernández, C., Ono, A., Kainosho, M. and Wüthrich, K. (1998) J. Am. Chem. Soc., 120, 821–822.
- Varani, G., Aboul-ela, F. and Allain, F.H.T. (1996) Prog. NMR Spectrosc., 29, 51–127.
- Vuister, G.W. and Bax, A. (1992) J. Magn. Reson., 98, 428-435.
- Wijmenga, S.S., Mooren, M.M.W. and Hilbers, C.W. (1993) In NMR of Macromolecules. A Practical Approach (Ed., Roberts, G.C.K.), Oxford University Press, pp. 217–288.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, H.J.D., Markley, J.L. and Sykes, B.D. (1995) *J. Biomol. NMR*, 6, 135–140.
- Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, NY.
- Ye, X., Kumar, R.A. and Patel, D.J. (1995) Chem. Biol., 2, 827-840.

37

Zimmer, D.P. and Crothers, D.M. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 3091–3095.