Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 4091–4095, April 1994
Biophysics

The des(1–6)Antennapedia homeodomain: Comparison of the NMR solution structure and the DNA-binding affinity with the intact Antennapedia homeodomain

(protein–DNA interactions/transcriptional regulation)

YAN QIU QIAN*, DIANA RESENDEZ-PEREZ†, WALTER J. GEHRING†, AND KURT WÜTHRICH*

*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland; and †Biozentrum der Universität Basel, Abteilung Zellbiologie, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Contributed by Kurt Wüthrich, December 20, 1993

ABSTRACT **The nuclear magnetic resonance (NMR) solution structure of an N-terminally truncated mutant Antennapedia homeodomain, des(1–6)Antp(C39S), has been determined from 935 nuclear Overhauser effect upper distance constraints and 148 dihedral angle constraints by using the programs DIANA and OPAL. Twenty conformers representing the solution structure of des(1–6)Antp(C39S) have an average root-mean-square distance relative to the mean coordinates of 0.56 Å for the backbone atoms of residues 8–59. Comparison with the intact Antp(C39S) homeodomain shows that the two proteins have identical molecular architectures. The removal of the N-terminal residues 1–6, which are flexibly disordered in the intact homeodomain, causes only strictly localized structure variations and does not noticeably affect the adjoining helix I from residues 10–21. The DNA-binding constant of des(1–6)Antp(C39S) is ≈10-fold reduced relative to the intact Antp(C39S) homeodomain, which can now be attributed to the absence of the previously reported contacts of the N-terminal polypeptide segment of the intact Antp(C39S) homeodomain with the minor groove of the DNA duplex.**

Homeotic genes are master control genes that regulate the development of eukaryotic organisms (for reviews, see refs. 1 and 2). They share a conserved DNA segment, the homeobox, and encode gene regulatory proteins that control the transcription of large sets of target genes. The homeobox encodes the homeodomain, which is the DNA-binding domain of the usually much larger homeotic proteins. Structure determinations of the DNA complexes of *Antp* (3–5), engrailed (6), and *Matα2* (7) revealed that three regions of the homeodomains are primarily involved in direct DNA contacts—i.e., the N-terminal hexapeptide, the loop at the start of the helix preceding the turn of the helix–turn–helix motif, and the recognition helix. This paper describes a systematic attempt at evaluating the relative importance of the three contacting regions of the homeodomain for the stability and specificity of DNA binding and recognition.

In the NMR solution structure of a DNA complex of the *Antp* homeodomain (3), the N-terminal polypeptide segment 0–6, which is flexibly disordered in the free homeodomain (8, 9), makes contacts with the DNA in the minor groove, and this structural feature was subsequently also found in the crystal structures of two DNA complexes with the homologous homeodomains engrailed (6) and *Matα2* (7). Deletion of the N-terminal hexapeptide in a mutant fushi tarazu (*ftz*) homeodomain, *des(1–6)ftz*, resulted in an increase of the equilibrium dissociation constant for DNA binding (10). In the present paper, we prepared a mutant Antp homeodomain with residues 1–6 deleted and Cys-39 replaced by Ser, des(1–6)Antp(C39S), measured its DNA-binding affinity, determined the NMR solution structure,‡ and compared these results with the corresponding data on the intact mutant Antp(C39S) homeodomain. [Antp(C39S) is a 68-residue polypeptide from the Antp protein, with Met at position 0, the Antp homeodomain sequence at positions 1–60 except that Cys-39 is replaced by Ser, and 7 additional residues of the Antp protein at positions 61–68.]

MATERIALS AND METHODS

Construction of the Antp Homeodomain Expression Plasmid pAop2CS-des(1–6). The N-terminal deletions in the Antp homeodomain were achieved by insertion of mutant oligonucleotides into the previously constructed plasmid pAop2CS (11, 12). The oligonucleotides 5′-TATGACGTACACCCGGTAC-3′ and 5′-CGGCTGTACGTAC-3′ were annealed and cloned into the corresponding *Nde* I and *Kpn* I sites of pAop2CS, from which the natural *Nde* I/*Kpn* I insert had previously been removed. The sequence of the new plasmid, pAop2CS-des(1–6), was confirmed by DNA sequencing using the dideoxynucleotide chain-terminator method (13).

Expression and Purification of des(1–6)Antp(C39S). The des(1–6)Antp(C39S) polypeptide was expressed in *Escherichia coli* BL21(DE3) Lys E using a T7 expression vector (14). Best induction was obtained after growing freshly transformed cells in super broth at 18°C overnight in a high-density Fermentor (Lab-Line Instruments). Cells were harvested by centrifugation and lysed in a buffer containing 0.2 mg of lysozyme per ml, followed by high-salt extraction with gentle shaking during 20 min on a rotating arm in the 4°C cold room. The lysate was centrifuged at 35,000 rpm in a 50 Ti Beckman rotor during 1 h at 4°C, and the clear supernatant was dialyzed twice for 3–4 h and once overnight at 4°C against a buffer solution containing 50 mM potassium phosphate (pH 7.0), 1 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride, and 0.4 M KCl. The des(1–6)Antp(C39S) homeodomain polypeptide was retained, washed, and eluted from a Bio-Rex 70 column with a salt gradient of 0.2–1.0 M KCl in the same buffer. The protein was further purified through a Mono S 10/10 FPLC column with a KCl elution gradient from 0.4 to 1.0 M. All fractions were monitored by SDS/PAGE on 15% acrylamide gels. About 30 mg of des(1–6)Antp(C39S) homeodomain was purified per 4 liters of induced culture.

Mobility-Shift Assays. Binding reactions for saturation experiments of the des(1–6)Antp(C39S) were carried out as

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; RMSD, root mean square distance.

‡The atomic coordinates and structure factors have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton NY 11973 (reference 1SAN).

Table 1. ^{1}H chemical shifts in des(1–6)Antp(C39S)

Residue*	Chemical shift δ, ppm†			
	NH	αH	βH	Others
Met-6		4.12	2.14,*2.03(−0.07)*	γCH$_2$ *2.48(−0.07)*, εCH$_3$ *1.54(−0.59)*
Thr-7	*8.49(+0.34)*	*4.37(+0.20)*	3.98	γCH$_3$ *1.09(+0.14)*
Tyr-8	*8.42(+0.28)*	*4.76(−0.13)*	2.70, 3.03	δH 7.00, 7.00, εH 6.69, 6.69
Thr-9	*9.09(−0.05)*	4.41	4.80	γOH 5.91,‡ γCH$_3$ 1.33
Leu-40	8.14	4.91	*1.63(+0.09)*, 1.30	γH 1.55, δCH$_3$ 0.88, 0.66
Thr-41	*8.67(−0.09)*	4.56	4.72	γOH 5.60,‡ γCH$_3$ 1.27
Gln-44	7.92	4.02	2.58, 2.02	γCH$_2$ 2.53, *2.51(+0.10)*; εNH$_2$ *7.32(−0.08)*, 6.84

*All residues are listed for which the chemical shift of at least one proton resonance in des(1–6)Antp(C39S) differs from the corresponding shift in Antp(C39S) by >0.05 ppm.

†Data for H_2O solution at pH 4.3 and 20°C; chemical shifts that differ from the corresponding values in the intact homeodomain by more than |0.05| ppm are italicized, and the shift difference, $\Delta\delta$ = δ[des(1–6)Antp(C39S)] − δ[Antp(C39S)], is given in parentheses.

‡Hydroxyl protons were observed at pH 4.3 and 4°C.

described by Affolter *et al.* (15) by using the oligodeoxynucleotide BS2-24 with the sequence 5′-CTGAGAAAAAGCCATTAGAGATCG-3′ as a ligand. The single-stranded oligodeoxynucleotides were purified by preparative FPLC (Pharmacia Mono Q HR10/10) with a step gradient from 0.3 to 0.8 M NaCl (pH 12.0) and desalted by ultrafiltration (YM-1 diaflow filter), which corresponds nearly identically to the procedures used previously (3).

NMR Measurements. NMR spectra were recorded on Bruker AMX 600 and AMX 500 spectrometers, using a 7.6 mM solution of des(1–6)Antp(C39S) in water at pH 4.3 and 20°C. The input for the structure calculation was obtained from two 600-MHz two-dimensional nuclear Overhauser effect spectroscopy (NOESY) spectra recorded with a mixing time of 60 ms: in H_2O solution the data size was 900 points in t_1 and 4096 points in t_2, t_{1max} = 56 ms, t_{2max} = 254 ms, water suppression by presaturation (16), total measuring time 64 h, quadrature detection in the phase-sensitive mode with time-proportional phase incrementation (TPPI) (17); in ^2H_2O solution the recorded data size was 1024 points in t_1 and 2048 points in t_2, t_{1max} = 79.9 ms, t_{2max} = 160 ms, total measuring time 47 h, phase-sensitive mode by States TPPI (18). Amide proton exchange rates were determined from the time course of the signal intensity measured with the program EASY (19) in a series of NOESY spectra recorded after dissolving the lyophilized protein in ^2H_2O with a mixing time of 100 ms and a recording time of 3 h. Values for $^3J_{HN\alpha}$ were extracted by inverse Fourier transformation of in-phase multiplets from the same NOESY spectrum in H_2O that was used to collect the NOE upper distance constraints (20), and $^3J_{\alpha\beta}$ values were measured using two-dimensional exclusive correlated spectroscopy (21) in ^2H_2O.

Determination of the Three-Dimensional Structure. The same strategy as for the structure determination of the intact Antp(C39S) homeodomain (12) was used, except that in the structure calculations with the program DIANA (22) one redundant dihedral angle constraints (REDAC) cycle was used for improved convergence (23). Energy minimization was performed with the program OPAL (P. Luginbühl, P. Güntert, M. Billeter, and K.W., unpublished data), which uses the AMBER force field (24) supplemented with previously described pseudoenergy terms for distance constraints and dihedral angle constraints (25).

RESULTS

^{1}H NMR Assignments and Structure Determination. Sequence-specific ^{1}H NMR assignments for des(1–6)Antp(C39S) were obtained with homonuclear two-dimensional NMR, following the usual strategy for small proteins (26–28). The total correlated spectroscopy and NOESY cross peak patterns of des(1–6)Antp(C39S) and the intact Antp(C39S) homeodomain

Table 2. Quantitative characterization of the 20 DIANA conformers representing the solution structure of des(1–6)Antp(C39S) before and after energy minimization with the program OPAL

Parameter	DIANA*†	DIANA + OPAL*‡
DIANA target function, Å^2	1.02 ± 0.23 (0.59–1.44)	
NOE violations		
Number >0.2 Å	1.8 ± 1.5 (0–5)	0.0
Sum of violations, Å	5.4 ± 0.74 (4.2–6.8)	8.1 ± 0.5 (6.8–8.8)
Maximum violation, Å	0.23 ± 0.04 (0.14–0.32)	0.10 ± 0.00 (0.10–0.11)
Dihedral angle constraint violations		
Number >5.0°	0.4 ± 0.5 (0–1)	0.0
Sum of violations, °	13.1 ± 4.5 (4.3–21.2)	23.6 ± 4.3 (16.5–33.0)
Maximum violation, °	4.0 ± 1.4 (1.7–6.7)	2.1 ± 0.2 (1.7–2.6)
AMBER energy, kcal/mol	174 ± 44 (104–277)	−1053 ± 51 (−1124 to −960)
RMSD values, Å		
Backbone N, C$^\alpha$, C′ (8–56)	0.49 ± 0.07 (0.33–0.59)	0.50 ± 0.08 (0.38–0.66)
Same +20 best side chains§	0.56 ± 0.06 (0.42–0.68)	0.56 ± 0.06 (0.45–0.70)
All heavy atoms (8–56)	1.16 ± 0.08 (0.99–1.34)	1.17 ± 0.09 (0.97–1.37)

*Numbers given are the average values ± SD, with the minimum and maximum values among the 20 conformers given in parentheses. Root-mean-square distance (RMSD) values are relative to the mean of the 20 conformers.

†DIANA conformers before energy minimization.

‡DIANA conformers after energy refinement with the program OPAL.

§Includes the backbone atoms N, C$^\alpha$, and C′ of residues 8–56 and the side-chain heavy atoms of the following 20 residues, which have global side-chain displacements smaller than 1.24 Å: 11–14, 16, 20, 26, 32, 34–40, 45, 47–49, and 51.

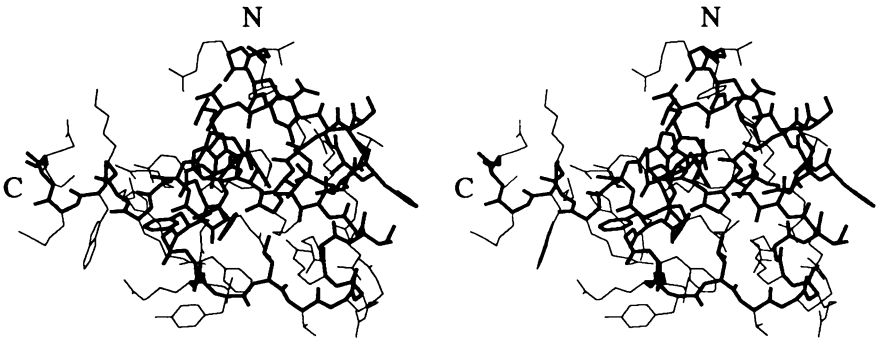

FIG. 1. Stereoview of one of the 20 energy-refined DIANA conformers used to characterize the solution structure of des(1–6)Antp(C39S) (Table 2). The backbone atoms N, C^{α}, and C′ of residues 8–59 and the side-chain heavy atoms of the 20 best-defined residues 11–14, 16, 20, 26, 32, 34–40, 45, 47–49, and 51 are drawn with thick lines; the other side chains are drawn with thin lines. N and C identify positions of residues 8 and 59, respectively.

are quite similar, and the residues with large chemical shift differences between the two proteins are all located near the N terminus, with some additional shifts exceeding |0.05| ppm in the turn of the helix–turn–helix motif and at the start of the recognition helix (Table 1).

For structure determination of des(1–6)Antp(C39S) a total of 1303 NOEs were assigned, and 57 $^3J_{HN\alpha}$ and 42 $^3J_{\alpha\beta}$ coupling constants representing unique rotamer states about χ^1 (29) were measured. After initial processing with the programs HABAS (30) and DIANA (22), an input of 935 NOE upper distance limits on internuclear distances and 148 constraints on the dihedral angles ϕ, ψ, and χ^1 was obtained. Stereospecific assignments for 36 nondegenerate proton pairs of βCH_2, γCH_2, and δCH_2 groups, and for 4 pairs of isopropyl methyl groups were determined with the programs HABAS and GLOMSA (12, 22). For all 6 side-chain amide groups of Gln and Asn, and for 2 ηNH_2 groups of Arg, individual assignments were determined from the relative intensity of NOEs to γCH_2, βCH_2, or εNH, respectively.

In the final DIANA calculations with 50 randomized starting structures, the best 20 conformers had residual values of the DIANA target function in the range 0.59–1.44 Å^2 (Table 2). Restrained energy minimization with the program OPAL yielded a low-energy conformation with further reduced maximum constraint violations and only a small increase of the sum of constraint violations (Table 2). The pairwise RMSD values of the 20 conformers relative to their means before and after energy minimization were virtually unchanged (Table 2). The value of 0.50 ± 0.08 Å for the average RMSD calculated for the backbone atoms N, C^{α}, and C′ of residues 8–56 is representative of a high-quality structure determination.

The NMR Structure of des(1–6)Antp(C39S) and Comparison with the Intact Antp(C39S) Homeodomain. The structure of des(1–6)Antp(C39S) (Fig. 1) includes helix I with residues 10–21, helix II with residues 28–38, and helix III with residues 42–52 and an extension by a more flexible helix IV containing residues 53–59. There is also a characteristic loop of residues 22–27 between helix I and helix II and a turn of residues 39–41 linking helices II and III. The visual impression of close coincidence between des(1–6)Antp(C39S) and Antp(C39S) (Fig. 2) is confirmed by the RMSD value of 0.86 Å calculated for the backbone atoms N, C^{α}, and C′ of residues 8–59 of the two mean structures. In both proteins, the C-terminal peptide segment 60–67 is flexibly disordered. In des(1–6)Antp(C39S) (Fig. 1), the helical secondary structures plus the 20 best-defined side chains (see Table 2) form the same globular core as in the intact Antp(C39S) homeodomain (12).

Close similarity between des(1–6)Antp(C39S) and the intact Antp(C39S) homeodomain is also supported by the fact that the amide proton exchange data are identical within the accuracy of the measurements (Fig. 3), but the chemical shifts (Table 1) and the hydrogen bonding networks in the two proteins indicate that there are nonetheless some localized differences near the turn between helices II and III and near the C-terminal end of helix IV. Although the two proteins contain otherwise identical hydrogen bonds (Table 3), the truncated homeodomain contains a hydrogen bond Arg-43 (NH)–Thr-41 (C′=O) in place of the hydrogen bond Ile-45 (NH)–Thr-41 (C′=O) observed in Antp(C39S), but Ile-45 NH rather than Arg-43 NH shows slow exchange with the solvent (Fig. 3). An additional hydrogen bond Ser-39 (γOH)–Leu-38 (C′=O) in the truncated protein has no counterpart in Antp(C39S) (Table 3). The different hydrogen bonds implicated for the turn can be accommodated in very similar structures, since the mean atom coordinates for the intact homeodomain are well within the group of 20 conformers for des(1–6)Antp(C39S) (Fig. 2). The different hydrogen bonds with residues 58–60 appear to be associated with a slightly

FIG. 2. Stereoview of a superposition of the polypeptide backbone of residues 8–59 of the 20 energy-refined des(1–6)Antp(C39S) conformers (thin lines) and the mean structure of the energy-refined Antp(C39S) homeodomain structure (thick line). The backbone atoms N, C^{α}, and C′ of residues 8–56 were superimposed for minimum RMSD. N and C identify positions of residues 8 and 59, respectively.

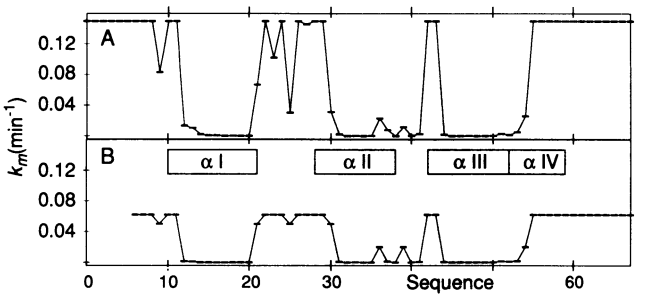

FIG. 3. Plots versus amino acid sequence of the backbone amide proton exchange rates, k_m. (*A*) Intact Antp(C39S) homeodomain at pH = 4.3 and 20°C [3.3 mM sample of uniformly ^{15}N-enriched Antp(C39S) homeodomain, ^{1}H frequency = 500 MHz, measurements with ^{15}N-^{1}H correlated spectroscopy, where only a lower limit was obtained for exchange rates faster than 0.15 min^{-1} (31)]. (*B*) Des(1–6)Antp(C39S) (protein concentration, 7.6 mM; ^{1}H frequency = 500 MHz, measurements with ^{1}H-^{1}H NOESY, where only a lower limit was obtained for exchange rates faster than 0.06 min^{-1}). In the center, locations of the four helices in the homeodomain sequence are indicated.

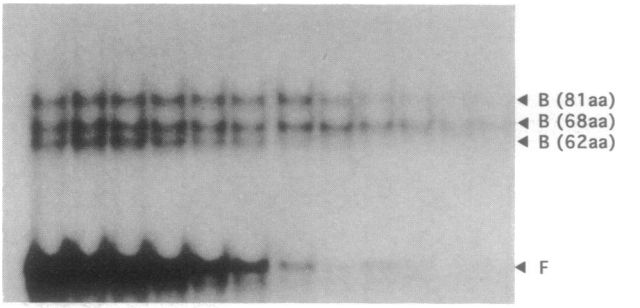

FIG. 4. Gel-shift assay for determination of the equilibrium dissociation constant (K_d) of homeodomain–DNA complexes. Constant amounts of the truncated des(1–6)Antp(C39S) (62 aa), the intact Antp(C39S) (68 aa), and the elongated Antp(YPWM) (81 aa) homeodomain polypeptides at a final concentration of 4 nM were incubated for 1 h in binding buffer containing the labeled BS-24 oligonucleotide at decreasing concentrations starting from 400 nM. Bound (B) and free (F) DNA was quantitated by liquid scintillation counting, and the results were used to calculate the K_d values.

different orientation of the last turn of the helix (Fig. 2) adjacent to the flexible C-terminal peptide segment.

Comparison of the DNA-Binding Affinities of des(1–6)Antp(C39S), Intact Antp(C39S), and Antp(YPWM). The equilibrium dissociation constants (K_d) were measured for des(1–6)Antp(C39S) with residues 6–67, intact Antp(C39S) with residues 0–67, and Antp(YPWM) with residues −14 to 67 (31). [Antp(YPWM) is an 81-residue polypeptide with the Antp(C39S) homeodomain at positions 1–60, the residues preceding and following the homeodomain in the Antp protein at positions −13 to −1 and 61–67, respectively, and Met at position −14.] Saturation experiments were used in which a fixed amount of polypeptide was incubated with increasing amounts of the oligodeoxynucleotide BS2-24 (15). Since in our experience the absolute values of such measurements show relatively large variations between different experiments, we added all three polypeptides to the same reaction mixture, so that the complex formation and complex stabilities could be

Table 3. Hydrogen bonds identified in the NMR solution structures of des(1–6)Antp(C39S) and the intact Antp(C39S) homeodomain

Donor*	Acceptor*	des(1–6) Antp(C39S)†	Antp (C39S)†	Donor*	Acceptor*	des(1–6) Antp(C39S)†	Antp (C39S)†
Thr-9 HN	Gln-12 δC=O	14‡	(§)	Leu-38 HN	Ile-34O′	20‡	20§
Gln-12 ε^2NH	Leu-38 O′	20‡	(§)	Ser-39 γOH	Ala-35O′	20‡	16(§)
Thr-13 HN	Thr-9 O′	20‡	20§	Ser-39 γOH	Leu-38O′	13	
Thr-13 γOH	Thr-9 O′	19		Leu-40 HN	Ala-35O′	13‡	15(§)
Leu-14 HN	Arg-10 O′	20‡	20§	Arg-43 HN	Thr-41O′	18	
Glu-15 HN	Tyr-11 O′	20‡	20§	Ile-45 HN	Thr-41O′	‡	20(§)
Leu-16 HN	Gln-12 O′	20‡	20§	Lys-46 HN	Glu-42O′	20‡	20§
Glu-17 HN	Thr-13 O′	20‡	20§	Ile-47 HN	Arg-43O′	20‡	20§
Lys-18 HN	Leu-14 O′	20‡	20§	Trp-48 HN	Gln-44O′	20‡	20§
Glu-19 HN	Glu-15 O′	20‡	20§	Phe-49 HN	Ile-45O′	20‡	20§
Phe-20 HN	Leu-16 O′	20‡	20§	Gln-50 HN	Lys-46O′	19‡	20§
His-21 HN	Glu-17 O′	20‡	20§	Asn-51 HN	Ile-47O′	20‡	20§
Phe-22 HN	Lys-18 O′	15	19	Arg-52 HN	Trp-48O′	20‡	19§
Asn-23 HN	Glu-19 O′	18	20§	Arg-53 HN	Phe-49O′	11‡	19(§)
Asn-23 δ^2NH	Tyr-25 O′	11		Met-54 HN	Gln-50O′	13‡	19(§)
Tyr-25 HN	Asn-23 δC=O	11‡	(§)	Lys-55 HN	Asn-51O′	15	20
Arg-31 HN	Thr-27 O′	14‡	20§	Trp-56 HN	Arg-52O′	16	20
Ile-32 HN	Arg-28 O′	13‡	20§	Lys-57 HN	Arg-53O′	16	20
Glu-33 HN	Arg-29 O′	19‡	20§	Lys-58 HN	Lys-55O′		11
Ile-34 HN	Arg-30 O′	20‡	20§	Glu-59 HN	Trp-56O′	18	15
Ala-35 HN	Arg-31 O′	18‡	20§	Asn-60 HN	Trp-56O′	18	
His-36 HN	Ile-32 O′	20‡	20§	Asn-60 δ^2NH	Trp-56O′	19	
Ala-37 HN	Asp-33 O′	20‡	20§				

*Hydrogen bonds are listed if they are present in at least 11 of the 20 energy-refined DIANA conformers representing the solution structure of either of the two proteins. These two columns specify the sequence positions of the residues in the intact Antp(C39S) homeodomain and the atom groups involved in the hydrogen bond, where HN and O′ stand for backbone amide proton and carbonyl oxygen, respectively.

†Columns list how many of the 20 energy-refined DIANA conformers representing the solution structure of des(1–6)Antp(C39S) or Antp(C39S), respectively, contain this hydrogen bond according to the criteria that $d \leq 2.4$ Å and $\theta \leq 35°$, where d is the proton–acceptor distance and θ is the angle between the donor–proton bond and the line connecting the acceptor and donor heavy atoms.

‡This amide proton of des(1–6)Antp(C39S) exchanges sufficiently slowly to be observable in a NOESY spectrum recorded during 3 h immediately after dissolving the protein in ^2H_2O at pH 4.3 and 20°C.

§This amide proton of the Antp(C39S) homeodomain exchanges slowly after dissolving the protein in ^2H_2O at pH 4.3 and 20°C, and the corresponding hydrogen bond was included by appropriate upper and lower distance limits in the input for the DIANA calculations (12). (§), Slowly exchanging protons for which no hydrogen bond constraints were included in the input for the DIANA calculations. Note that no supplementary hydrogen bond constraints were used in the input for the calculation of the des(1–6)Antp(C39S) structure.

directly compared. Fig. 4 shows that Antp(C39S) has the lowest K_d, followed by Antp(YPWM) and des(1–6)Antp(C39S). The absolute K_d values at 20°C in three independent experiments were calculated to be $2.3 \pm 0.5 \times 10^{-9}$ M for the intact Antp(C39S) homeodomain, $2.1 \pm 0.5 \times 10^{-8}$ M for des(1–6)Antp(C39S), and $1.0 \pm 0.1 \times 10^{-8}$ M for Antp(YPWM).

DISCUSSION

Prokaryotic repressors with a helix–turn–helix motif make most of their DNA contacts with the recognition helix in the major groove (32). In addition, the λ repressor has an extended N-terminal arm, which also makes critical DNA contacts, but these are again confined to the major groove. In contrast, the N-terminal arm of the homeodomain makes critical contacts in the minor groove of the DNA, which is more similar to the binding mode proposed for the hin recombinase (33). The importance of the N-terminal arm is underlined by recent genetic experiments that indicate that the functional specificity of two very similar homeodomains, those of Antp and Sex combs reduced (Scr), resides in the N-terminal segment (34, 35). By exchanging residues 1, 4, 6, and 7, which distinguish Scr from Antp, the functional specificity of the protein can be changed from that of Scr to that of Antp. Whether this effect is due to differences in DNA binding between the two proteins and/or selective interactions with other transcription factors is not known.

Although one could imagine that the N-terminal arm might be involved in the folding of the homeodomain core, thus assuming an intramolecular chaperone function, the present structure determination of des(1–6)Antp(C39S) indicates that the deletion of the flexible N-terminal hexapeptide does not noticeably affect the static and dynamic properties of the molecular structure of the homeodomain. In particular, there is no detectable effect on the immediately adjacent helix I from residues 10–21, indicating that the core homeodomain folds autonomously and that the arbitrarily generated ends of the polypeptide chain do not necessarily become disordered upon truncation of the chain.

Since the present work has shown that des(1–6)Antp(C39S) and Antp(C39S) have virtually identical three-dimensional structures, the observed 10-fold reduction of the DNA binding constant for des(1–6)Antp(C39S) can now be attributed to the absence of the flexible N-terminal arm of residues 1–6. On the level of the homeodomain–DNA interaction there appears to be no straightforward explanation for the reduced affinity of the elongated Antp(YPWM) homeodomain, which leads one to speculate that *in vivo* the highly conserved YPWM motif might be essential for selective interactions with other transcription factors.

We thank M. Grilc for performing part of the assignment work, Drs. M. Billeter and P. Güntert for helpful discussions, and R. Marani for careful processing of the manuscript. Financial support from the Schweizerischer Nationalfonds (Projects 31.32033.91 and 31.28707.90) is gratefully acknowledged.

1. McGinnis, W. & Krumlauf, R. (1992) *Cell* **68**, 283–302.
2. Wüthrich, K. & Gehring, W. J. (1992) in *Transcriptional Regulation*, eds. McKnight, S. L. & Yamamoto, K. R. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 535–577.
3. Otting, G., Qian, Y. Q., Billeter, M., Müller, M., Affolter, M., Gehring, W. J. & Wüthrich, K. (1990) *EMBO J.* **9**, 3085–3092.
4. Qian, Y. Q., Otting, G., Billeter, M., Müller, M., Gehring, W. J. & Wüthrich, K. (1993) *J. Mol. Biol.* **234**, 1070–1083.
5. Billeter, M., Qian, Y. Q., Otting, G., Müller, M., Gehring, W. J. & Wüthrich, K. (1993) *J. Mol. Biol.* **234**, 1084–1097.
6. Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B. & Pabo, C. O. (1990) *Cell* **63**, 579–590.
7. Wolberger, C., Vershon, A. K., Liu, B. S., Johnson, A. D. & Pabo, C. D. (1991) *Cell* **67**, 517–528.
8. Billeter, M., Qian, Y. Q., Otting, G., Müller, M., Gehring, W. J. & Wüthrich, K. (1990) *J. Mol. Biol.* **214**, 183–197.
9. Qian, Y. Q., Billeter, M., Otting, G., Müller, M., Gehring, W. J. & Wüthrich, K. (1989) *Cell* **59**, 573–580.
10. Percival-Smith, A., Müller, M., Affolter, M. & Gehring, W. J. (1990) *EMBO J.* **9**, 3967–3974.
11. Müller, M., Affolter, M., Leupin, W., Otting, G., Wüthrich, K. & Gehring, W. J. (1988) *EMBO J.* **7**, 4299–4304.
12. Güntert, P., Qian, Y. Q., Otting, G., Müller, M., Gehring, W. J. & Wüthrich, K. (1991) *J. Mol. Biol.* **217**, 531–540.
13. Sanger, F., Nicklen, S. & Coulsen, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
14. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–89.
15. Affolter, M., Percival-Smith, A., Müller, M., Leupin, W. & Gehring, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4093–4097.
16. Wider, G., Hosur, R. V. & Wüthrich, K. (1983) *J. Magn. Reson.* **52**, 130–135.
17. Marion, D. & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113**, 967–974.
18. Marion, D., Ikura, M., Tschudin, R. & Bax, A. (1989) *J. Magn. Reson.* **85**, 393–399.
19. Eccles, C., Güntert, P., Billeter, M. & Wüthrich, K. (1991) *J. Biomol. NMR* **1**, 111–130.
20. Szyperski, T., Güntert, P., Otting, G. & Wüthrich, K. (1992) *J. Magn. Reson.* **99**, 552–560.
21. Griesinger, C., Sørensen, O. W. & Ernst, R. R. (1985) *J. Am. Chem. Soc.* **107**, 6394–6396.
22. Güntert, P., Braun, W. & Wüthrich, K. (1991) *J. Mol. Biol.* **317**, 517–530.
23. Güntert, P. & Wüthrich, K. (1991) *J. Biomol. NMR* **1**, 447–456.
24. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. (1986) *J. Computat. Chem.* **7**, 230–252.
25. Widmer, H., Billeter, M. & Wüthrich, K. (1989) *Proteins* **6**, 357–371.
26. Billeter, M., Braun, W. & Wüthrich, K. (1982) *J. Mol. Biol.* **155**, 321–346.
27. Wagner, G. & Wüthrich, K. (1982) *J. Mol. Biol.* **155**, 347–366.
28. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
29. Nagayama, K. & Wüthrich, K. (1981) *Eur. J. Biochem.* **115**, 653–657.
30. Güntert, P., Braun, W., Billeter, M. & Wüthrich, K. (1989) *J. Am. Chem. Soc.* **111**, 3997–4004.
31. Qian, Y. Q., Otting, G., Furukubo-Tokunaga, K., Affolter, M., Gehring, W. J. & Wüthrich, K. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10738–10742.
32. Pabo, C. O. & Sauer, R. T. (1992) *Annu. Rev. Biochem.* **61**, 1053–1095.
33. Affolter, M., Percival-Smith, A., Müller, M., Billeter, M., Qian, Y. Q., Otting, G., Wüthrich, K. & Gehring, W. J. (1991) *Cell* **64**, 879–880.
34. Furukubo-Tokunaga, K., Flister, S. & Gehring, W. J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6360–6364.
35. Zeng, W., Andrew, D. J., Mathies, L. D., Horner, M. A. & Scott, M. P. (1993) *Development* **118**, 339–352.