The EMBO Journal vol.7 no.13 pp.4299-4304, 1988

Isolation and sequence-specific DNA binding of the *Antennapedia* homeodomain

Martin Müller, Markus Affolter, Werner Leupin[1], Gottfried Otting[1], Kurt Wüthrich[1] and Walter J.Gehring

Biozentrum der Universität Basel, CH-4056 Basel and [1]Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland

Communicated by W.J.Gehring

The homeodomain encoded by the *Antennapedia* (*Antp*) gene of *Drosophila* was overproduced in a T7 expression vector in *Escherichia coli*. The corresponding polypeptide of 68 amino acids was purified to homogeneity. The homeodomain was analysed by ultracentrifugation and assayed for DNA binding. The secondary structure of the isolated homeodomain was determined by nuclear magnetic resonance spectroscopy. DNA-binding studies indicate that the isolated homeodomain binds to DNA *in vitro*. It selectively binds to the same sites as a longer *Antp* polypeptide and a full-length *fushi tarazu* (*ftz*) protein. Therefore, the homeodomain represents the DNA-binding domain of the homeotic proteins.

Key words: Antennapedia/homeodomain/homeobox/ sequence-specific DNA-binding protein

Introduction

Homeotic genes and several other genes controlling development share a characteristic DNA segment, the homeobox. This DNA sequence of ~180 bp was first discovered in *Drosophila* (McGinnis *et al.*, 1984a,b; Scott and Weiner, 1984), but was subsequently found in many other metazoa including vertebrates and also in man (McGinnis *et al.*, 1984a,b; Carrasco *et al.*, 1984; Levine *et al.*, 1984). It encodes a section of homeotic proteins which is called the homeodomain, and comprises a sequence of ~60 amino acids that has been highly conserved throughout evolution (Gehring, 1987a). Computer searches through protein sequence data banks have revealed a small but significant degree of homology (22-26% identity) to the yeast mating type proteins MAT a1 and α2 of *Saccharomyces cerevisiae* and P1 of *Schizosaccharomyces pombe* (Shepherd *et al.*, 1984; Gehring, 1987a). Since the MAT genes of yeast encode gene regulatory proteins and determine cell type specificity, this observation, along with genetic evidence, suggested a gene regulatory function for homeotic genes. On the basis of structural considerations and some sequence homologies, it had been proposed earlier that the MAT a1 protein may contain a helix-turn-helix motif (Ohlendorf *et al.*, 1983) as is found in various prokaryotic gene regulatory proteins that bind to specific DNA sequences. Limited sequence homology between the homeodomain and the prokaryotic DNA-binding proteins has also been found (Laughon and Scott, 1984; Shepherd *et al.*, 1984). This suggested that the homeodomain represents the DNA-binding domain of the homeotic proteins and that it contains a helix-turn-helix motif. In addition, DNA-sequence comparison between prokaryotic DNA-binding proteins and homeodomain sequences of *Drosophila* indicated that the amino acid sequences in the putative helix-turn-helix motif are much more conserved among the *Drosophila* homeodomains than among the prokaryotic DNA-binding proteins (Gehring, 1987a,b). In particular, the homeodomains of the *Antennapedia* (*Antp*) family, which includes *fushi tarazu* (*ftz*), have largely invariant putative recognition helices, which suggests that they may bind to the same or similar DNA sequences.

In order to test these hypotheses we have expressed an artificial gene encoding the *Antp* homeodomain in *Escherichia coli*, and purified the polypeptide to homogeneity. We describe the isolation of the homeodomain of the *Antp* protein and some of its DNA-binding properties.

Results

Expression of the Antennapedia homeobox and purification of the homeodomain polypeptide

Large amounts of homeodomain were prepared by introducing the *Antp* homeobox into an *E.coli* expression system developed by Rosenberg *et al.* (1987), which makes use of bacteriophage T7 polymerase to direct selective, high level expression of cloned genes. As there are no suitable restriction sites at the 5′ end of the *Antp* homeobox, two complementary oligodeoxynucleotides were synthesized which encode the first 10 amino acids of the homeodomain. Insertion of the oligodeoxynucleotides into the unique *Nde*I site in the expression vector pAR3038 provided the ATG start codon and allowed for cloning of the remaining homeobox sequences using the *Kpn*I site within the *Antp* homeodomain. At the C terminus the homeodomain is extended by seven amino acids due to a *Bam*HI site at that position of the *Antp* DNA sequence. An artificial translation stop codon was introduced by ligating a 6-bp *Hpa*I linker to the mung bean nuclease-treated *Bam*HI site. The last amino acids in front of the stop codon are not changed by this procedure (see Figure 1).

In the new recombinant plasmid pAop2 the 68 amino acid *Antp* homeodomain peptide is under the control of the T7 promoter. All amino acids except the N-terminal methionine are *Antp* derived. For overexpression, the bacterial strain BL21(DE3) (Studier and Moffatt, 1986) was transformed with pAop2. The *Antp* homeodomain polypeptide was purified to homogeneity from IPTG-induced cultures (see Figure 2, and Materials and methods). In fully induced cells the homeodomain comprises at most 5% of the total soluble protein.

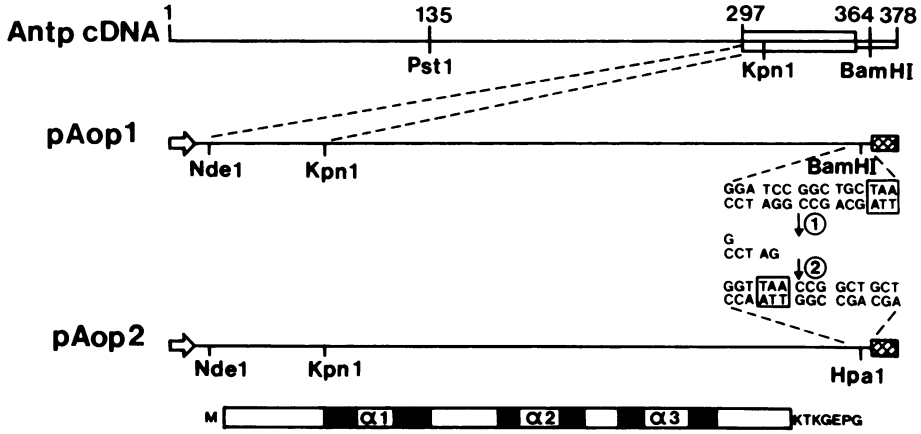

Fig. 1. Cloning strategy for construction of the *Antp* homeodomain expression plasmid pAop2. Only restriction sites important for cloning are indicated. Numbers above the *Antp* cDNA give amino acid residue positions. The partial *Antp* protein used for DNase I protection assays starts at amino acid residue 135 and ends at amino acid residue 364. The *Antp* homeodomain contains the amino acid residues 297–364. The homeobox and homeodomain (defining the homeobox as 180 bases and the homeodomain as 60 amino acids, respectively) are boxed in. In pAop1 the translation stop lies within the T7 transcription terminator. Compared to pAop2 it encodes three additional amino acids. (**1**) Digestion with restriction endonuclease *Bam*HI. (**2**) 5′ Overhangs produced by *Bam*HI restriction were digested with mung bean nuclease and subsequently 6-bp *Hpa*I linkers were ligated. pAop2 encodes the 68-amino acid *Antp* homeodomain peptide which was purified. Amino acids outside the homeodomain are given in single letter code. Black boxes within the homeodomain show helical structures as determined by NMR (Otting *et al.*, 1988). They are named α1, α2 and α3 according to Pabo and Sauer (1984). ⇨: T7 promoter; ⊠ T7 transcription terminator.

Sedimentation properties of Antennapedia homeodomain

The sedimentation properties of purified homeodomain were investigated using analytical ultracentrifugation. Under reducing conditions [25 mM Tris–HCl, pH 7.0, 100 mM NaCl, 2 mM dithiothreitol (DTT)] a monophasic broad sedimentation profile was observed (see Figure 3A). This type of broad profile is expected for small peptides which have high diffusion coefficients. The sedimentation coefficient was calculated to be $s_{20,w} = 1.0 \pm 0.1$S. From sedimentation equilibrium data a linear ln$A$ versus r^2 plot was obtained again indicating homogeneity of the protein preparation. Runs performed at different peptide concentrations and temperatures gave similar values with an average of $M_r = 9040 \pm 340$, as compared with $M_r = 8545$ derived from the sequence. The relationship between the s value and the mol. wt is indicative of an ellipsoid with an axial ratio between 1:4 and 1:8.

In the absence of DTT the material was no longer homogenous as indicated by a curved ln$A$ versus r^2 plot. The minimum mol. wt as estimated from the slope of the plot near the meniscus was ~10 000, indicating the presence of monomers in the mixture. A quantitative evaluation of a large fraction of material with higher mol. wt was not achieved but the curvature of the ln$A$ versus r^2 is consistent with the presence of large amounts of *Antp* homeodomain dimer. The average sedimentation coefficient of the mixture was $s_{20,w} = 1.3$S.

These data suggest that the homeodomain can dimerize in solution via a disulphide bond between two *Antp* homeodomains due to the single cysteine at position 39. The *Antp* homeodomain was analysed further by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) in the presence or absence of β-mercaptoethanol in the loading buffer.

As shown in Figure 3B, lane 1, under reducing conditions, a single species was detected which corresponds to the homeodomain monomer. When the protein was analysed under non-reducing conditions (lane 2), an additional doublet was detected which migrated in the position expected for a homeodomain dimer, which is consistent with the ultracentrifugation analysis. It was surprising, however, that the putative dimer form migrated as a doublet, even in the presence of SDS. Several lines of evidence suggest that both of the bands of this doublet correspond to an S–S-linked dimer which can exist in two different stable conformations.

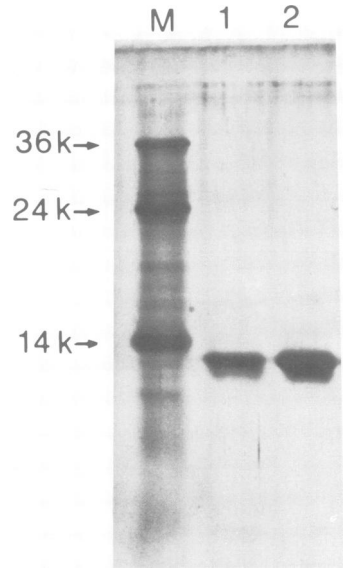

Fig. 2. Silver-stained SDS–PAGE of purified *Antp* homeodomain. Homeodomain peptide preparations were analysed on a 17% SDS–polyacrylamide gel. **Lane 1**: 2 μg peptide purified for DNA-binding studies. **Lane 2**: 10 μg peptide purified for NMR studies. Impurities present at the top of **lanes 1** and **2** are due to contamination in the sample buffer. Procedures are described under Materials and methods. M marker lane; protein mol. wt markers were purchased from Sigma: 14-kd bovine milk α-lactalbumin; 24-kd bovine pancreas trypsinogen; 36-kd rabbit muscle glyceraldehyde-3-phosphate dehydrogenase.

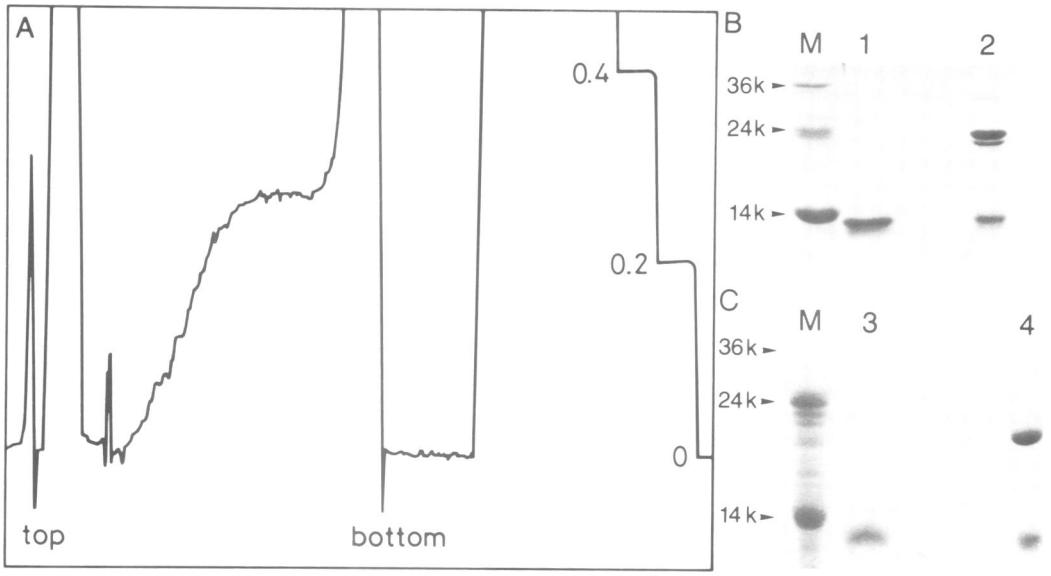

Fig. 3. Sedimentation and SDS−PAGE analysis of *Antp* homeodomain. (**A**) Sedimentation velocity profile of *Antp* homeodomain in 25 mM Tris−HCl (pH 7), 100 mM NaCl, 2 mM DTT. The scan was taken after 224 min. Rotor speed 60 000 r.p.m.; temperature 9°C. Protein concentration was 0.7 mg/ml corresponding to 8.2×10^{-5} M. Absorption calibration is shown on the right (A_{280}). Top and bottom show orientation of the cell. (**B** and **C**) Coomassie blue-stained 17% SDS−polyacrylamide gel, and Coomassie blue-stained 6 M urea/17% polyacrylamide gel, respectively. **Lanes 1** and **3** 20 μg *Antp* homeodomain in the presence of 500 mM β-mercaptoethanol in the loading buffer. **Lanes 2** and **4** 20 μg *Antp* homeodomain under non-reducing conditions. M marker lane; markers were the same as in Figure 2.

Firstly, as shown in Figure 3C (lane 4), denaturation of the protein with 6 M urea in the absence of reducing agents converts the two bands into a single species, which migrates slightly faster than either species of the doublet but significantly slower than the monomer. Secondly, alkylation of the protein with *N*-ethylmaleimide (NEM) prior to gel electrophoresis carried out under reducing (2 mM DTT) or non-reducing conditions almost completely abolished both of the doublet bands, yielding the monomer form (data not shown). Thirdly, both dimer species are still detectable after treatment at 100°C for 2 min in the presence of 0.1% SDS, again suggesting the presence of a covalent linkage. Since both bands of the doublet are susceptible to reduction with DTT and β-mercaptoethanol, and are sensitive to NEM treatment, they must both correspond to S−S-linked homeodomains. These results suggest that dimerization of purified homeodomains *in vitro* may not reflect a biologically significant dimerization. This possibility will be investigated further by replacing the homeodomain cysteine with a serine by site-directed mutagenesis and assaying its properties both *in vitro* and *in vivo*.

DNA-binding properties of the Antennapedia homeodomain

Previous genetic and *in situ* hybridization experiments have shown that the establishment of the pattern of expression of the segment polarity gene *engrailed* (*en*) depends upon the activity of the homeobox-containing segmentation gene *fushi tarazu* (*ftz*) (Howard and Ingham, 1986). The putative gene regulatory function of homeo proteins as well as the occurrence of a helix-turn-helix motif in the homeodomain (see Otting *et al.*, 1988) suggests that this *ftz*-dependent regulation may occur at the level of transcription and involve the binding of *ftz* via its homeodomain to specific *cis*-regulatory elements upstream of the *en* gene. To assess whether or not the *ftz* protein indeed binds to specific DNA sequences, mobility shift and immunoprecipitation assays were performed with cloned *en* DNA fragments and purified *ftz* protein which had been synthesized in *E.coli* (Krause and Gehring, unpublished). These experiments suggested that an *Eco*RI−*Cla*I restriction fragment that maps −2.3 kb to −0.95 kb upstream of the transcription initiation site contains one or more *ftz* binding sites. DNase I protection assays show that *ftz* protein protects three regions (designated BS1 to BS3) within this fragment (Figure 4, lanes 1−6). If binding of the *ftz* protein to these three sites were mediated by the homeodomain via the helix-turn-helix motif, one would predict that the *Antp* protein would bind to the same sequence elements because of the complete amino acid sequence identity of helix 3 (the putative recognition helix) in these two homeo proteins. Footprint analyses performed with a partial *Antp* protein (see Figure 1) indicate that this is indeed the case (Figure 4, lanes 7−12). BS1, BS2 and BS3 are all protected at approximately the same *Antp* protein concentration as those needed for protection by the *ftz* protein. DNase I digestions performed in the presence of the purified *Antennapedia* homeodomain show that this peptide protects the same DNA sequence elements from DNase I attack (Figure 4, lanes 13−17) as those protected by the full length *ftz* or the partial *Antp* proteins. This clearly indicates that the DNA-binding domains of the *ftz* and *Antp* proteins lie within their homeodomains and suggests that the structure adopted by the isolated homeodomain is similar to the structure the domain adopts in the context of the longer polypeptide chains.

Inspection of the DNA sequence of regions protected from DNase I attack (Figure 5) shows that each binding site contains at least one copy of the sequence motif ANNN-NCATTA. This motif has also been found in several other regions upstream of the *ftz* gene that were protected from

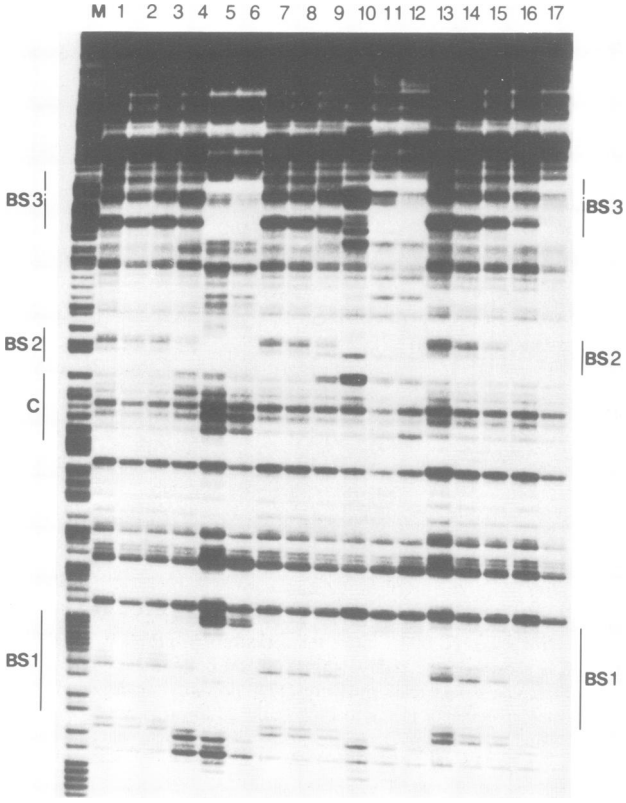

Fig. 4. Location of DNA-binding sites of homeodomain containing proteins and of purified homeodomain by footprinting. Increasing amounts of purified protein synthesized in *E.coli* and purified to near homogeneity were incubated with an *Eco*RI−*Cla*I fragment that has been 3′-end-labelled at the *Eco*RI site. After partial digestion with DNase I, nucleic acids were purified and separated on a 6% sequencing gel. The proteins used were full-length *fushi tarazu* (**lanes 2−6**), a partial *Antennapedia protein* (**lanes 8−12**), and *Antennapedia* homeodomain (**lanes 14−17**). The amounts of protein used were 40 ng (**lanes 2, 8** and **14**), 200 ng (**lanes 3, 9** and **15**), 1 μg (**lanes 4, 10** and **12**), 5 μg (**lanes 5, 11** and **17**) and 25 μg (**lanes 6** and **12**). No protein was added to the binding reaction loaded in **lanes 1, 7** and **13**. M shows a G+A reaction of the *Eco*RI−*Cla*I fragment. BS1 to BS3 identify three binding sites protected from DNase I digestion. BS3 is a bipartite binding site as indicated in Figure 5. C shows the position of the control oligo used in the shift assay.

DNase I attack by *ftz* protein or *Antennapedia* homeodomain (data not shown).

In order to study the interaction of the purified homeodomain with the identified binding sites, an oligodeoxynucleotide containing the footprinted region BS2 (oligo BS2) and a control oligodeoxynucleotide of equal size (oligo C: see Figures 4 and 5) were prepared as described in Materials and methods. The ability of the homeodomain to retard the mobility of these oligodeoxynucleotides was studied by electrophoresis in non-denaturing gels.

Binding of the homeodomain to oligo BS2 can be readily detected as revealed by the retarded electrophoretic mobility of the labelled DNA fragment (Figure 6A, compare 0 nM to 100 nM). No such specific retarded DNA−protein complexes are formed with oligo C in the same concentration range. At high protein concentrations (≥400 nM), severely retarded complexes are formed with both oligo BS2 and oligo C (Figure 6A and B). To estimate the apparent dissociation constant [K_d (app)] for homeodomain binding to oligo BS2 under our conditions, the decrease in free DNA concentration

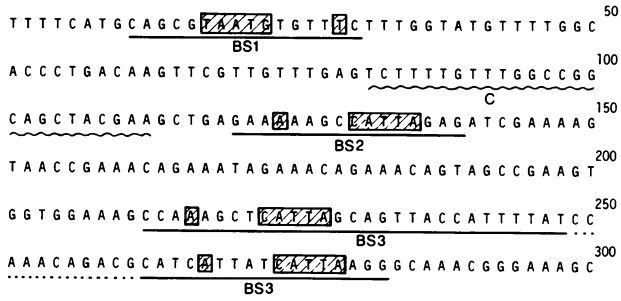

Fig. 5. Nucleotide sequence of binding sites. The horizontal lines beneath the sequence indicate the boundaries of the DNase I-protected region BS1 to BS3. The conserved nucleotides in each ANNNNCATTA motif are boxed. The position of the control oligo (C) is indicated by a wavy line. Nucleotide 1 is located 51 nucleotides downstream of the *Eco*RI site (see Materials and methods).

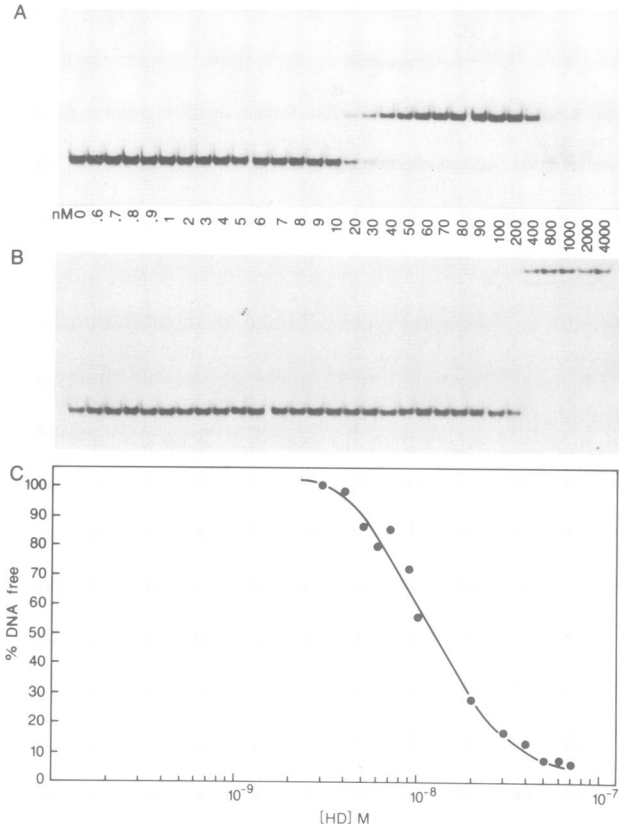

Fig. 6. Mobility shift assay of homeodomain−DNA binding. Binding reactions were carried out as described in Materials and methods. Increasing amounts of homeodomain were incubated with equivalent amounts of oligo BS2 (**A**) or oligo C (**B**) and separated on 15% native polyacrylamide gels. (**C**) Quantitation of mobility shift assays. The amount of free DNA present after addition of increasing amounts of homeodomain was quantitated by densitometry of the gels shown in Figure A. The concentration required for half-saturation of BS2 is 1.2×10^{-8} M.

as a function of the homeodomain input was quantitated by densitometry. Because the concentration of DNA in the binding reaction is $< 2 \times 10^{-10}$ M, the concentration of free homeodomain corresponds approximately to the total concentration of homeodomain within the range where binding occurs and K_d (app) thus approximates the protein concentration required for half-maximal binding. For oligo BS2, the K_d (app) was found to be 1.2×10^{-8} M (Figure

6C). This value is ~2-fold lower for a 26-bp oligodeoxynucleotide containing BS1 (data not shown).

Discussion

In order to determine the structural and functional properties of the homeodomain, we have expressed a polypeptide of 68 amino acids encoding the entire *Antp* homeodomain along with one additional methionine at the amino terminus in *E.coli* using the T7 expression system of Studier and Moffatt (1986). The homeodomain was purified to homogeneity. On the basis of sedimentation studies and PAGE we conclude that, under reducing conditions, the homeodomain exists as a monomer in solution. No significant amount of dimerized homeodomain was observed in ultracentrifugation studies up to a concentration of 8.2×10^{-5} M homeodomain, indicating that if dimerization occurs, the dissociation constant must be $< 10^{-4}$ M. Under non-reducing conditions the single cysteine (in the turn between helices 2 and 3) can form a disulphide bridge leading to dimerization which presumably is an artefact and can be prevented by the addition of 2 mM DTT. It should be noted that the homeodomains of several other proteins such as *ftz*, lack a cysteine residue at that position.

The NMR analysis described in the accompanying paper (Otting *et al.*, 1988) demonstrates the presence of three helices within the homeodomain. These studies revealed a striking structural similarity with the helix-turn-helix motif found in many prokaryotic DNA-binding proteins (e.g. Pabo and Lewis, 1982; Anderson *et al.*, 1987; Zuiderweg *et al.*, 1984). The presence of the helix-turn-helix motif fulfils predictions based upon DNA-sequence comparisons (Gehring, 1987a,b).

The functional properties of the isolated homeodomain were investigated by DNA-binding studies. Previous studies using either fusion proteins or full-length polypeptides expressed in *E.coli* showed that homeotic proteins bind to specific DNA sequences *in vitro* (Desplan *et al.*, 1985; Cho *et al.*, 1988; Hoey and Levine, 1988). We have identified three *ftz* binding sites (BS1 to BS3) ~2.0–2.3 kb upstream of the *en* gene. These sites are also recognized by a partial *Antp* protein and the isolated *Antennapedia* homeodomain. As both the *ftz* protein and the *Antennapedia* homeodomain peptide were purified to homogenity, this demonstrates clearly that the *in vitro* DNA-binding properties of the *ftz* and *Antp* homeotic proteins are retained by the isolated homeodomain.

The sequence motif ANNNNCATTA occurs at least once in each binding site. Closer inspection of the nucleotide sequence protected on both DNA strands indicates that both BS1 and BS2 are protected over 16 bp from DNase I attack and that the CATTA motif is at equivalent positions within the protected region (Figure 5 and data not shown). The same holds true for several additional *ftz* binding sites upstream of the *ftz* gene (data not shown). Although other consensus recognition sequences derived from homeo protein-binding sites *in vitro* do not share the ANNNNCATTA motif (Cho *et al.*, 1988; Hoey and Levine, 1988), they do contain a ATTA core sequence. Strikingly, the CATTA motif (which corresponds to a TAATG motif on the opposite strand) is also found in the herpex simplex virus immediate-early promoter-specific TAATGARAT element (Whitten and Clemens, 1984) and is present in the binding site for nuclear factor III (Pruijn *et al.*, 1986). Whether any of the factors interacting with these elements do so via a homeodomain can only be determined upon cloning of the corresponding genes.

The availability of the *Antennapedia* homeodomain peptide also allowed us to estimate the apparent dissociation constant for the binding to an oligodeoxynucleotide containing BS2. The value of 1.2×10^{-8} M was obtained under conditions where the free, unbound homeodomain exists as a monomer in solution. Experiments performed with the monomeric amino-terminal DNA-binding domain of λ repressor gave a K_d for O_R1 binding of 3×10^{-8} M (Sauer *et al.*, 1979). Recent experiments with the purified C-terminal domain of the yeast repressor α2 revealed that the concentration required to half-saturate the STE 6 operator was ~1.2×10^{-7} M (Sauer *et al.*, 1988). These results suggest that the binding site BS2 might be useful to characterize in more detail the interaction of the homeodomain with its target DNA sequence. This is also suggested by mobility shift experiments that showed that *Drosophila Antp* protein also forms very specific complexes with oligodeoxynucleotide BS2 while no such complexes are obtained with oligodeoxynucleotide C (data not shown).

The results presented in this paper show that the homeodomain is indeed the DNA-binding domain of homeotic proteins and that the structure as determined by NMR studies (see Otting *et al.*, 1988) is functional. This is the first demonstration of a helix-turn-helix motif in a eukaryotic DNA-binding domain and highlights the conservation of structure–function relationship in evolution. *In vitro* and *in vivo* studies using site-directed mutagenesis to determine the functional contacts between the protein and the DNA are in progress.

Materials and methods

Construction of Antp homeodomain expression plasmid pAop2

The strategy of constructing pAop2 is described in Results and involves standard cloning procedures (Maniatis *et al.*, 1982). The structure of the final plasmid was verified by DNA sequencing using the dideoxy chain terminator method (Sanger *et al.*, 1977).

Purification of the Antp homeodomain and homeotic proteins

All polypeptides were expressed in *E.coli* BL21(DE3) lysogen (Studier and Moffatt, 1986). For purifying the *Antp* homeodomain, freshly transformed cells were grown up to OD_{600} ~0.8 at 37°C in LB-medium (10 g Bacto-Tryptone Difco, 5 g Bacto yeast extract Difco, 5 g NaCl in 1 l H_2O). Cultures were induced for 5 h with 0.125 mg IPTG/ml. All of the following operations were carried out at 0–4°C and were monitored by SDS gel electrophoresis. Cells were harvested by centrifugation and washed with 50 mM phosphate buffer, pH 7.5. Cells were then resuspended in sucrose buffer (50 mM phosphate buffer, pH 7.5, 10% sucrose, 2 mM EDTA, 400 mM NaCl, 1 mM DTT, protease inhibitors, see below) and lysed with a French Press at a pressure of 1000 p.s.i. The lysate was centrifuged at 16 000 *g* in a GSA rotor for 10 min. Material precipitated at 0.8% Polymin P was removed from the supernatant by dropwise addition of a stock solution of 10% Polymin P, 50 mM phosphate buffer, pH 7.5, to the crude extract. After stirring slowly for 15 min, the extract was centrifuged at 16 000 *g* in a GSA rotor for 10 min and the supernatant was collected. This Polymin P fraction was directly loaded onto a BioRex A-70 column previously equilibrated in 400 mM NaCl in buffer A (50 mM phosphate buffer, pH 7.5, 10% glycerol, 1 mM DTT, protease inhibitors). The column was extensively washed with 500 mM NaCl buffer A and then eluted with a linear gradient of 500–900 mM NaCl in buffer B (50 mM phosphate buffer, pH 7.5). The peak fractions were pooled and adjusted to a final concentration of 500 mM NaCl. Minor remaining impurities present in the pooled Biorex fraction were removed by loading it onto a Pharmacia mono S 10/10-FPLC column which was then eluted with a linear gradient of 600–900 mM NaCl in buffer B. Peak fractions were desalted using gel filtration (Pharmacia NAP-25 columns) or ultra filtration (Amicon-cells with

Diaflo ultrafiltration membrane YM2). 2–4 mg of *Antp* homeodomain were purified per litre of induced culture. Protein concentration was determined using the Bio-Rad dye reagent with bovine γ-globulin as standard, according to the method of Bradford (1976). Protein preparations used for DNA-binding studies were purified the same way except that 25 mM Tris–HCl buffer was used instead of 50 mM phosphate, and all buffers contained 1 mM DTT. Protease inhibitors were added to buffers where indicated just before use at the following final concentrations: leupeptin 2 μg/ml, pepstatin 2 μg/ml, aprotinin 2 μg/ml, benzamidine-HCl 0.1 mM, antipain 1 μg/ml, bacitracin 1 μg/ml, soybean trypsin inhibitor 5 μg/ml.

For SDS–PAGE (Laemmli, 1970) 17% acrylamide gels were used. For urea-containing SDS–PAGE, acrylamide solutions were adjusted to 6 M urea. For modification with NEM, the protein was incubated for 30 min at room temperature with 5 mM NEM in 25 mM Tris–HCl, pH 7.0, 100 mM NaCl, 2 mM DTT. Protein samples were dissolved in sample buffer with or without reducing agent and boiled for 2 min. The *ftz* protein was purified as described by Krause *et al.* (1988). A partial *Antp* protein (amino acids 135–364, see Figure 1) was also expressed in a T7 system and purified to near homogeneity using similar methods (B.Dalle Carbonare *et al.*, in preparation).

Analytical ultracentrifugation studies

A Beckman model E analytical ultracentrifuge equipped with UV absorption optics and a photoelectric scanner was used. Sedimentation velocity experiments were performed at 8°C and 20°C in 12 mm double sector cells at a rotor speed of 60 000 r.p.m. The absorbance A was measured at 280 nm. Protein solutions were in the range of $A_{280} = 0.4$. For dilute protein solutions, 30 mm double sector cells were employed at 52 000 r.p.m. The observed sedimentation coefficients s were corrected in the customary manner to water at 20°C. A value of $\nu = 0.73$ cm^3/g was used for the partial specific volume. Sedimentation equilibrium experiments were carried out at ~10°C and 20°C in 12 mm or 30 mm double sector cells at rotor speeds of 32 000 r.p.m. and 44 000 r.p.m. Cells of 12 mm pathlength were filled with 0.11 ml of protein solution ($A_{280} = 0.1–0.3$). 30 mm cells were filled with 0.3 ml of protein solution ($A_{280} = 0.04$). Molecular weights were evaluated from ln$A$ versus r^2 plots, where r is the distance from the rotor centre. Frictional coefficients, f, were calculated in the usual manner. The axial ratios a/b of ellipsoids of revolution, as hydrodynamic equivalents, were estimated from the f/f min ratios according to Perrin's tables (Tanford, 1961). A value of 0.5 g water/g protein was assumed for the degree of hydration.

Oligodeoxynucleotides

The oligodeoxynucleotides were synthesized by the phosphoramidite method on an Applied Biosystem 380 B DNA synthesizer and were purified with oligodeoxynucleotide purification cartridges (Applied Biosystems). The duplex DNA molecules were annealed by heating and slow cooling and were radioactively labelled with [α-^{32}P]dATP using Klenow polymerase according to standard procedures (Maniatis *et al.*, 1982). Oligo BS2 was prepared by annealing two complementary oligodeoxynucleotides of sequence AGCTGAGAAAAAGCCATTAGAGA and GCTTCTCTAATGGCTTT-TTCTCA and oligo C by annealing oligodeoxynucleotides of sequence TCTTTTGTTTGGCCGGCAGCTACG and TTCGTAGCTGCCGGCC-AAACAAAAG.

DNase I protection assays

Protein was pre-incubated with 2 μg of poly(dI-dC) competitor DNA in 100 μl of 1 × binding buffer (20 mM Tris–HCl, pH 7.6, 5 mM $MgCl_2$, 75 mM KCl, 5 μg BSA, 1 mM DTT, 10% glycerol) for 15 min on ice. Binding reactions were then started by the addition of 25 ng of ^{32}P-labelled DNA. Following incubation for 30 min on ice, 2.5 mM $CaCl_2$ and 0.1 μg/ml DNase I were added and the incubation was continued for 5 min. Reactions were stopped by the addition of 100 μl of 10 mM Tris pH 7.6 containing 0.4% SDS and 20 mM EDTA. Mixtures were heated at 65°C for 15 min, and the samples were extracted with equal volumes of phenol–chloroform (1:1) and chloroform. The aqueous phase was ethanol-precipitated in the presence of 2.5 μg of sheared *E.coli* carrier DNA.

The DNA used in the footprint assay was an *Eco*RI–*Cla*I restriction fragment of ~1.35 kb from the *Drosophila* engrailed gene locus. The *Eco*RI site is located ~2.3 kb upstream of the transcription initiation site and was 3′-end-labelled with [α-^{32}P]dATP and Klenow polymerase (Maniatis *et al.*, 1982). Sequencing reactions were performed according to Maxam and Gilbert (1980). Samples were separated on 8.3 M urea/6% polyacrylamide gels.

Mobility shift assays

Binding reactions for all mobility shift assays (Fried and Crothers, 1981; Garner and Revzin, 1981; see also Carey, 1988) were carried out on ice and the samples were separated on native 0.5 × TBE (Maniatis *et al.*, 1982) 15% polyacrylamide gels at 4°C for 4–6 h at 300 V. The gels were pre-electrophoresed at 300 V for 2 h. Binding reactions were performed in 20 μl of the binding buffer used for footprint assays, except that poly(dI-dC) and $MgCl_2$ were omitted. The homeodomain was appropriately diluted in the binding buffer 15 min before addition of 75 pg of oligodeoxynucleotide. After 1 h of incubation, the samples were loaded onto the gel without adding any tracking dye. After electrophoresis, gels were covered with Saran wrap and exposed for autoradiography without an intensifying screen at −80°C. For quantification a Camag electrophoresis scanner was used.

Acknowledgements

We are indebted to Ariel Lustig and Jürgen Engel for carrying out the ultracentrifugations and interpreting the data. Bruno Dalle Carbonare kindly provided the *Antennapedia* polypeptide and Henry Krause the *fushi tarazu* protein. We are most grateful to Leslie Pick for advice on protein purification and to Erika Wenger-Marquardt for processing of the text. This work was supported by the Kanton of Basel and the Swiss National Science Foundation (projects 3.613.087 and 3.198.85).

References

Anderson,J.E., Ptashne,M. and Harrison,S.C. (1987) *Nature*, **326**, 846–852.
Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
Carey,J. (1988) *Proc. Natl. Acad. Sci. USA.*, **85**, 975–979.
Carrasco,A.E., McGinnis,W., Gehring,W.J. and DeRobertis,E.M. (1984) *Cell*, **37**, 409–414.
Cho,K., Goetz,J., Wright,C., Fritz,A., Harwicke,J. and DeRobertis,E.M. (1988) *EMBO J.*, **7**, 2139–2149.
Desplan,C., Theis,J. and O'Farrell,P. (1985) *Nature*, **318**, 630–635.
Fried,M. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.
Garner,M.M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047–3059.
Gehring,W.J. (1987a) In Firtel,R. and Davidson,E., *Molecular Approaches to Developmental Biology*. Alan R. Liss, Inc., NY, pp. 115–129.
Gehring,W.J. (1987b) *Science*, **236**, 1245–1252.
Hoey,T. and Levine,M. (1988) *Nature*, **332**, 858–861.
Howard,K. and Ingham,P. (1986) *Cell*, **47**, 949–957.
Krause,H.M., Klemenz,R. and Gehring,W.J. (1988) *Genes Dev.*, **2**, 1021–1036.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Laughon,A. and Scott,M.P. (1984) *Nature*, **310**, 25–31.
Levine,M., Rubin,G.J. and Tijan,R. (1984) *Cell*, **38**, 667–674.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–525.
McGinnis,W., Levine,M.S., Hafen,E., Kuroiwa,A. and Gehring,W.J. (1984a) *Nature*, **308**, 428–433.
McGinnis,W., Garber,R.L., Wirz,J., Kuroiwa,A. and Gehring,W.J. (1984b) *Cell*, **38**, 403–409.
Ohlendorf,D.H., Anderson,W.F. and Matthews,B.W. (1983) *J. Mol. Evol.*, **19**, 109–114.
Otting,G., Qian Yan-qui, Müller,M., Affolter,M., Gehring,W.J. and Wüthrich,K. (1988) *EMBO J.*, **7**, 4305–4309.
Pabo,C.O. and Lewis,M. (1982) *Nature*, **298**, 443–447.
Pabo,C.O. and Sauer,R.T. (1984) *Annu. Rev. Biochem.*, **53**, 293–321.
Pruijn,G.J.M., Van Driel,W. and van der Vliet,P.C. (1986) *Nature*, **322**, 656–659.
Rosenberg,A.H., Lade,B.N., Chui,D.-S., Lin, S.-W., Dunn,J.J. and Studier,F.W. (1987) *Gene*, **56**, 125–135.
Sanger,F., Nicklen,S. and Coulsen,A.R. (1977) *Proc. Natl. Acad. Sci. USA.*, **74**, 5463–5467.
Sauer,R.T., Pabo,C.O., Meyer,B.J., Ptashne,M. and Backman,K.C. (1979) *Nature*, **279**, 396–400.
Sauer,R.T., Smith,D.L. and Johnson,D. (1988) *Genes Dev.*, **2**, 807–816.
Scott,M.P. and Weiner,A.J. (1984) *Proc. Natl. Acad. Sci. USA.*, **81**, 4115–4119.
Shepherd,J.C.W., McGinnis,W., Carrasco,A.E., DeRobertis,E.M. and Gehring,W.J. (1984) *Nature*, **310**, 70–71.
Studier and Moffatt (1986) *J. Mol. Biol.*, **189**, 113–130.
Tanford,C. (1961) *Physical Chemistry of Macromolecules*. John Wiley Inc., NY and London.
Whitten,J.L. and Clemens,J.B. (1984) *Nucleic Acids Res.*, **12**, 2061–2079.
Zuiderweg,E.R.P., Billeter,M., Boelens,R., Scheek,R.M., Wüthrich,K. and Kaptein,R. (1984) *FEBS Lett.*, **174**, 243–247.

Received on August 29, 1988; revised on September 28, 1988