Several regions of Antennapedia and thyroid transcription factor 1 homeodomains contribute to DNA binding specificity
(homeobox/recognition helix/NK-2 gene)

GIUSEPPE DAMANTE AND ROBERTO DI LAURO*
European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, Federal Republic of Germany

Communicated by Walter J. Gehring, March 21, 1991 (received for review December 20, 1990)

ABSTRACT The DNA binding specificities of the homeodomains contained in thyroid transcription factor 1 and Antennapedia have been compared. The two homeodomains recognize different DNA sequences, despite the similar amino acid sequences of their recognition helices. Mutations that make the recognition helix of thyroid transcription factor 1 identical to the one contained in Antennapedia have no effects on the binding specificity of thyroid transcription factor 1. The exchange of other segments between these two homeodomains allows the identification of the regions responsible for the observed DNA binding specificities. These results indicate that amino acid residues outside of the recognition helix play an important role in the determination of the DNA binding specificities of these two homeodomains.

The homeodomain (HD) is a conserved protein motif that is capable of sequence-specific DNA recognition (1, 2). This interaction has been proposed to depend on the presence of a helix-turn-helix (HTH) motif, first recognized by the structural similarity between HDs and prokaryotic repressors (3, 4). Structural studies have shown that the HTH motif is present in the Antennapedia (Antp) and Engrailed HDs (5–8). DNA binding and transcriptional activation experiments carried out with some Drosophila HD proteins suggested that the second helix of the HTH motif (the recognition helix) is responsible for DNA sequence discrimination. The swap of the ninth amino acid residue of the recognition helix between pairwise combinations of Paired, Bicoid, and Fushi-tarazu HDs was sufficient to exchange the DNA binding specificity of these HDs (9, 10). In contrast, the Deformed and Ultrabitorax HDs, which contain an identical recognition helix, confer distinct regulatory properties to the Deformed protein when assayed in the Drosophila embryo. In this last experiment, however, it is not clear whether the differences in biological function are due to changes in DNA binding specificity (11).

In this paper, we present studies carried out on the HD of thyroid transcription factor 1 (TTF-1). The HD of TTF-1 cannot be grouped into any of the classes of HDs described so far and represents a member of a different class of mammalian HDs (Melanie Price and R.D.L., unpublished data) whose Drosophila counterpart is the NK-2 HD (12, 13). Although the TTF-1 HD is significantly divergent in primary sequence from the prototype Antennapedia HD, the two recognition helices show considerable similarity. Nevertheless, the two HDs recognize two quite different DNA sequences (12). We have exchanged segments between Antp and TTF-1 HDs in order to identify the regions responsible for the observed DNA binding specificities. The results presented in this paper indicate that the DNA binding specificity of these two proteins depends on amino acid residues outside of the recognition helix.

MATERIALS AND METHODS

Expression and Purification of TTF-1 and Antp HDs. Antp HD expression vector pAop2 (14) was a gift from Walter Gehring (Biozentrum der Universitat Basel). The construction of TTF-1 HD expression vector has been described (12). Both HDs were expressed in the bacterial strain BL21 and purified to homogeneity as described for the Antp HD (14). Wild-type and mutant HDs were also obtained by in vitro transcription-translation using T7 RNA polymerase and commercially available reticulocyte lysate (Promega). The in vitro-made mRNA for each HD was titrated in order to obtain similar amounts of proteins in each translation reaction. Translation products, labeled with [35S]methionine, were run on 18% polyacrylamide gels, and relative amounts of protein were quantitated by densitometry. We also verified, by oligonucleotide saturation experiments (15), that active proteins were present at similar concentrations in all translation mixtures used (data not shown).

DNA Binding Studies. [32P]labeled, double-stranded oligonucleotides C (12) and BS2 (14) were prepared by 5’ end-labeling of the top strand with polynucleotide kinase and subsequent annealing with the bottom strand. Gel-retardation assays were performed by incubating proteins and DNAs in a buffer containing 20 mM Tris-HCL, pH 7.6/75 mM KCl/0.25 mg of bovine serum albumin per ml/1 mM dithiothreitol/10% glycerol for 1 hr. When in vitro-translated proteins were used in the binding assay, 0.15 mg of poly(dI-dC) per ml was added in the incubation buffer. Protein-bound DNA and free DNA were separated on native 10% polyacrylamide gels at 4°C for 2–3 hr at 300 V. The gel was exposed to x-ray film and the bands were quantitated by densitometric scanning of the autoradiogram. The binding activity of each mutant is the mean of the values obtained in three independent experiments.

Construction of Mutants. Mutants of the M series contain point mutations and were obtained by polymerase chain reaction (PCR) using the method of Ho et al. (16). Swapping mutants were obtained using the overlap-extension method (17). Mutants were cloned in the vector pT7.7 and expressed by in vitro transcription-translation as described above.

RESULTS

TTF-1 and Antp Recognize Two Different DNA Sequences. The amino acid sequences of the TTF-1 and Antp HDs used in these studies are shown in Fig. 1a. A comparison of the

Abbreviations: TTF-1, thyroid transcription factor 1; Antp, Antennapedia; HD, homeodomain; HTH, helix-turn-helix.
*To whom reprint requests should be addressed.
primary structure of the two HDs shows a considerable divergence, except for those residues conserved in most or all of the HDs isolated up to now. More importantly, 7 of the 10 residues in the recognition helix are identical, including the glutamine at position 50, which has been implicated as the chief determinant of the DNA binding specificity of other HDs. Nevertheless, the HDs from TTF-1 and Antp, synthesized in vitro, showed absolute specificities for oligonucleotides C and BS2, respectively (12). To explore a wider range of protein concentrations, we used purified proteins obtained from overproducing bacterial strains. Under these conditions, TTF-1 HD binding to oligonucleotide BS2 and, conversely, Antp HD binding to oligonucleotide C can be observed (Fig. 1b), but the affinity of this interaction is similar to that observed for binding of the HDs to nonspecific DNA. This is most clearly indicated by the competition experiment shown in Fig. 1c.

**Effects of Mutations in the Recognition Helix of TTF-1.** To reveal the residues important in the determination of the differential binding specificity we mutagenized TTF-1 and Antp HDs in the attempt to switch their DNA binding specificity. The mutant proteins, obtained by *in vitro* translation in reticulocyte lysates, were then challenged with either a TTF-1 (oligonucleotide C) or an Antp (oligonucleotide BS2) binding site. We first concentrated on the recognition helix, because some nonconservative substitutions occur from TTF-1 to Antp (Fig. 2). In mutant M2, residues 12, 43, and 45 of TTF-1 HD were changed to those present in the corresponding position in the Antp HD. These mutations generate a HD consisting of the Antp recognition helix surrounded by the residues of the TTF-1 HD. Surprisingly, the TTF-1(M2) HD maintains the binding specificity of the wild-type TTF-1 HD. In addition, no significant difference in the affinity for the TTF-1 binding site was observed between
TTF-1 and TTF-1(M2). This first mutant shows that the information for the differential DNA binding displayed by Antp and TTF-1 must reside outside of the recognition helix. Two additional mutants were then constructed in order to test the role, if any, of the TTF-1 recognition helix in DNA binding (Fig. 2). Changing the glutamine at position 50 to alanine (mutant M3) results in reduced binding activity, suggesting that this glutamine residue, even if not responsible for the differential binding, may be important for DNA recognition. The substitution of tryptophan and phenylalanine at positions 48 and 49, respectively, by alanine (mutant M4) causes a null DNA binding phenotype. Trp-48 and Phe-49 are invariant residues in HDs and are thought to be essential for the proper folding of the entire HD structure (6, 7). The null phenotype of mutant M4 suggests that these two residues have a similar role also in the TTF-1 HD.

**Swap of Segments Outside of the Recognition Helix Between TTF-1 and Antp.** The DNA binding behavior of the mutants described above clearly rules out the recognition helix of TTF-1 and Antp HDs as responsible for the observed differential DNA recognition. We next investigated whether exchanging regions outside helix III could switch DNA binding specificity between TTF-1 and Antp. Mutant templates were constructed by PCR and the corresponding proteins were obtained by translation in reticulocyte lysates. We first swapped the carboxyl-terminal region (helix IV) between TTF-1 and Antp, generating the two mutants SW1 and SW2. Mutant SW1 (TTF-1,II,III/ AntpIV), when compared to TTF-1, shows a decrease in affinity for oligonucleotide C and an increase in affinity for oligonucleotide BS2 (Fig. 3). The results obtained using mutant SW4 (TTF-1,II,III/ AntpIV), which does not show any increase in the affinity to oligonucleotide BS2 when compared to SW1 (TTF-1,II,III/ AntpIV) (Fig. 3), further support a greater relevance of helix IV compared to helix III in the determination of the TTF-1 and Antp HD DNA binding specificities. Conversely, mutant SW2 (Antp,II,III/TTF-1 IV), when compared to Antp HD, shows an increase in the binding activity to oligonucleotide C and a decrease in the binding to oligonucleotide BS2 (Fig. 3).

An important aspect of the hybrid HDs obtained after exchange of helix IV is their intermediate behavior. SW1 binds oligonucleotide BS2 better than TTF-1 but not as well as Antp. Conversely, SW2 binds oligonucleotide C better than Antp but not as well as TTF-1. This intermediate behavior suggested that other regions of the HDs could be important for DNA binding specificity. To test this hypothesis we exchanged the amino-terminal region of TTF-1 and Antp. Mutant SW11, in which the amino-terminal region of TTF-1 substitutes for the corresponding region of Antp HD (TTF-1/ AntpIV), was not informative, as this protein is no longer able to bind either target site (Fig. 3). Since the only difference between SW4 and SW11 is the presence of Antp HD helix II, the result suggests an incompatibility between the amino-terminal region of TTF-1 and helix II of Antp but not vice versa, as demonstrated by the behavior of the reverse mutant SW20. Mutant SW20 (Antp/ TTF-1,II,III), when compared to TTF-1 HD, shows an increase in binding to oligonucleotide BS2 and a decreased binding to oligonucleotide C, suggesting that also the amino terminus of HD plays a role in DNA binding specificity. Again, an intermediate behavior is observed, which, when taken together with the binding behavior of mutants SW1 and SW2, suggests that the amino and carboxyl termini of these two HDs are important for DNA binding specificity. Mutant SW21 (AntpIV/ TTF-1,II,III), in which both the amino- and carboxyl-terminal regions of the Antp HD substitute for the corresponding regions of the TTF-1 HD, strongly supports this notion, since it binds very poorly to oligonucleotide C and almost as efficiently as Antp to oligonucleotide BS2, thus displaying an unambiguous Antp phenotype.
**FIG. 3.** Relative binding activity of hybrid HDs. (a) Structures of swap (SW) mutants. The helical structures of the HDs are shown at the top. Numbers indicate the amino acid present at the left border of the swapped elements. In each construct sequences coding for the TTF-1 HD are shown as a black line; sequences coding for Antp HD are shown as a thicker, shadowed line. (b) Relative binding activities of SW mutants.

All proteins used in this experiment were synthesized in reticulocyte lysates.

**DISCUSSION**

It has been established that the sequence-specific DNA recognition operated by TTF-1 (12) and Antp (14, 18) rests mostly, if not entirely, on their HDs. In the case of TTF-1, differences were found in DNase I footprints obtained with either the entire protein or the HD. Those differences were ascribed to the different sizes of the proteins, resulting in a different steric hindrance to DNase I cleavage. Methylation interference experiments, carried out with the isolated TTF-1 HD and the entire protein, revealed no differences (unpublished observations), again suggesting that most of the binding properties of TTF-1 derive from its HD. A few other HDs tested also showed the capability to recapitulate most of the binding properties of the proteins from which they were obtained (19–21). POU HDs, instead, require an additional domain, the POU-specific domain, for high-affinity DNA binding (22, 23). Perhaps as a consequence of this different mode of binding to DNA, the nature of residue 9 in the recognition helix in the Pit-1 gene HD seems to affect transcriptional activation rather than DNA binding activity (23).

A frequently reported aspect of the DNA recognition by HDs is promiscuous DNA binding specificity, as demonstrated by the observations that (i) the same HD can recognize a variety of different DNA sequences that do not contain a clear common consensus (19, 24) and (ii) distantly related HDs can recognize, in vitro, the same DNA sequence (20, 25). It should be stressed, however, that in both cases the degree of affinities involved varies widely. Contrasting with this attribute of degeneracy in DNA recognition, some HD proteins are unable to recognize the DNA sequence bound by other HDs. In these cases, the DNA binding specificity has been demonstrated to depend on the nature of residue 50, which is in the second helix of the HTH motif (9, 10). This finding is sometimes generalized to state that the recognition helices of HDs dictate their DNA binding specificities (26, 27).

The HDs of TTF-1 and Antp display two different DNA binding specificities, although they both have a glutamine at position 50. In this study we have investigated the structural basis for the observed differential DNA binding. One important question is whether the differential binding of the two HDs studied, for oligonucleotides C and BS2, reflects only a quantitative difference in affinity, which would still allow a specific interaction of the two HDs for both DNA sequences. The competition experiment shown in Fig. 1C demonstrates that the TTF-1 HD binds oligonucleotide BS2 to the same extent as nonspecific DNA, suggesting that no specificity, or very little, is involved in this interaction. The same considerations apply to the binding of the Antp HD to oligonucleotide C. Additional information derived from the competition experiment is the relatively small differential affinity (≈100) of the two HDs for specific versus nonspecific binding sites, which was noted earlier for the Antp HD (15).

Surprisingly, mutations that changed the recognition helix of TTF-1 to make it identical to the one of Antp had no effect on sequence-specific DNA recognition, indicating that the residues important for discrimination must reside outside of the recognition helix. This prediction has been verified by analysis of the swap mutants, which have indicated that the specificity of binding depends on residues present at the two ends of the HDs studied. Among the swap mutants, SW11 displays an interesting null phenotype. The explanation of this behavior may be in the close proximity, in SW11 HD, of residues 18–28 of TTF-1 HD with helix II of Antp, two regions with a high density of positive charges, which may result in an interference with the proper packing of helices I and II in this mutant.

The existence of contacts at the amino and carboxyl termini of HDs was first proposed on the basis of a comparison between the HDs and prokaryotic DNA binding proteins containing a HTH motif (28). Recently, several studies have demonstrated that such contacts are established. In the Antp–BS2 complex, investigated by NMR, residues in helices I and IV, as well as in the loop between helices I and II, make, or are likely to make, DNA contacts (29). Interestingly, several of these residues are different between Antp and TTF-1 and thus are good candidates for the determination of DNA binding specificity. Also the crystal structure of the Engrailed HD–DNA complex shows that the amino and carboxyl termini of the protein are in good position to establish contacts with DNA (7). Finally, a biochemical study of the DNA contacts established by the wild-type and mutant Fushi-tarazu HDs shows that the amino terminus of this protein contacts the minor groove of DNA (30).

Although the structural studies described above show that HDs contact their target sites with residues distributed throughout their length, our results indicate that contacts established by residues outside the recognition helix are likely to be crucial for DNA sequence discrimination by the
two HDs studied here and may be generally applicable to other HD proteins. Our studies also indicate clearly that the nature of the amino acid residues in helix III of HDs is not sufficient to precisely predict their DNA binding specificities.

We thank W. Gehring for the Antp expression vector and for communicating results before publication. The purified Antp HD was a gift of M. Müller. We also thank I. Mattaj, D. Doboule, R. De Francesco, M. Price, H. Francis-Lang, and H. Stunnenberg for comments on the manuscript. G.D. was supported by a European Economic Community fellowship.