The Plant Cell, Vol. 3, 747–748, August 1991

ATTS, a New and Conserved DNA Binding Domain

LETTER TO THE EDITOR

A variety of structural motifs have been identified in transcriptional regulatory proteins that are responsible for their binding to specific DNA sequences. These motifs include cysteine-rich metallocoordination domains (e.g., zinc "finger" domains; Berg, 1989; Churchill et al., 1990; Pan and Coleman, 1990), basic regions adjacent to leucine dimerization domains (bZIP; Vinson et al., 1989; Busch and Sassone-Corsi, 1990; Hu et al., 1990), basic regions adjacent to helix-loop-helix domains (bHLH; Murre et al., 1989; Benezra et al., 1990), helix-turn-helix domains, found in classical bacterial transcription factors (Pabo and Sauer, 1984), homeobox proteins, and POU domains (Sturm and Herr, 1988; Qian et al., 1989; Gehring et al., 1990). However, some regulatory proteins that are known or presumed to bind to specific DNA sequences lack identified DNA binding motifs, including the AbaA developmental regulatory protein of *Aspergillus nidulans* (Mirabito et al., 1989; Sewall et al., 1990). Xiao et al. (1991) recently reported the sequence of TEF-1, a cell type-specific SV40 enhancer-binding protein. We noted that TEF-1 contains a sequence motif with significant similarity to sequences present in the *A. nidulans* AbaA protein and the yeast protein TEC1p, a regulator of Ty1 transcriptional activity (Laloux et al., 1990). We propose that this motif defines a novel DNA binding domain that we term ATTS (**A**baA, **T**EC1p, **T**EF-1 **s**equence).

Figure 1 shows the regions of highest sequence similarity between the proteins, spanning ~70 amino acid residues. This region contains two highly conserved core sequences, which are underlined in Figure 1. There is little or no sequence similarity between the proteins outside of this region. The consensus sequence derived from alignment of the homologous regions does not resemble previously identified DNA binding domains. Searches of the Swiss-Prot, PIR, GenBank, and EMBL data bases revealed no additional sequences having significant similarity.

As shown in Figure 1, the consensus sequence is somewhat basic, containing 12 basic residues within 44 residues of the carboxyl terminus. Four of the 5 acidic residues occur within the 20 residues of the amino terminus. The predicted secondary structure consists of one α-helix within the ATTS motif, in the amino-terminal half, based on the Chou-Fasman algorithm (Chou and Fasman, 1978). Two β sheets are also predicted in the carboxyl-terminal half of the ATTS motif, and these overlap the two highly conserved core sequences. The predicted helix is separated from the sheets by two to five predicted β turns.

Xiao et al. (1991) suggested that the DNA binding motif of TEF-1 for its cognate GT-IIC and Sph-I-Sph-II enhanson sequences resides in the region between Pro-26 and Ala-98, encompassing the ATTS region (Figure 1). Laloux et al. (1990) showed that the amino terminal 289 residues of TEC1p, also encompassing the ATTS region, are sufficient for genetic complementation of a *tec1* deletion mutation. Although AbaA has not been demonstrated to be a DNA-binding protein, it contains a potential leucine zipper dimerization domain and activates transcription of sporulation-specific genes (Mirabito et al., 1989), which is consistent with such a function. Thus, it is highly probable that AbaA, TEC1p, and TEF-1 are DNA-binding proteins and that the ATTS motif is directly responsible for this activity. The high degree of evolutionary conservation in the ATTS motif suggests that it should not be difficult to identify related sequences in other species. It will be of interest to

```
                                 α-HELIX                                     β-SHEET               β-SHEET

AbaA      132 tGkDgEpVWS DElEDAFQqA LeanPPmGRR K..wSERGKs YGRNELIAe YIyklTGKrR TRKQVSSHlQ VL
TEF-1      27 IdnDaEgVWS pDIEqsFQEA LaiyPPcGRR KIilSDeGKm YGRNELIAr YIKlrTGKtR TRKQVSSHIQ VL
TEC1p     127 IGcDk...WS EkVEEAFlEA LrlimknGtt KI..kiRnan FGRNELIsl YIKhkTnefR TkKQISSHIQ V.

Consensus     IG-D-E-VWS ---E-AFQEA L---PP-GRR KI--S-RGK- YGRNELIA- YIK--TGK-R TRKQVSSHIQ VL
```

Figure 1. Alignment of ATTS domains.

Sequences were aligned by using the GAP and LINEUP programs of the University of Wisconsin GCG analysis package (Devereux et al., 1984). The consensus sequence was generated by using the PRETTY program. Secondary structure predictions were generated by using the PEPTIDESTRUCTURE program. Two regions of very high sequence similarity are underlined. The α-helix and β sheets are indicated above the sequence with the dotted line representing positional differences between the individual sequences. GenBank accession numbers for these sequences are J04850 (*abaA*), M32797 (*TEC1*), and M63896 (*TEF-1*).

determine whether ATTS has been as widely exploited in evolution as other DNA binding domains and to determine how it interacts with DNA.

Alex Andrianopoulos and William E. Timberlake
Departments of Genetics and Plant Pathology
University of Georgia
Athens, GA 30602

REFERENCES

Benezra, R., Davis, R.L., Lockshon, D., Turner, D.L., and Weintraub, H. (1990). The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. Cell **61,** 49–59.

Berg, J.M. (1989). DNA binding specificity of steroid receptors. Cell **57,** 1065–1068.

Busch, S.J., and Sassone-Corsi, P. (1990). Dimers, leucine zippers and DNA-binding domains. Trends Genet. **6,** 36–40.

Chou, P.Y., and Fasman, G.D. (1978). Empirical predictions of protein conformation. Annu. Rev. Biochem. **47,** 251–276.

Churchill, M.E.A., Tullius, T.D., and Klug, A. (1990). Mode of interaction of the zinc finger protein TFIIIA with a 5S RNA gene of *Xenopus*. Proc. Natl. Acad. Sci. USA **87,** 5528–5532.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. **12,** 387–395.

Gehring, W.J., Müller, M., Affolter, M., Percival-Smith, A., Billeter, M., Qian, Y.Q., Otting, G., and Wüthrich, K. (1990). The structure of the homeodomain and its functional implications. Trends Genet. **6,** 323–329.

Hu, J.C., O'Shea, E.K., Kim, P.S., and Sauer, R.T. (1990). Sequence requirements for coiled-coils: Analysis with λ repressor-GCN4 leucine zipper fusions. Science **250,** 1400–1403.

Laloux, I., Dubois, E., Dewerchin, M., and Jacobs, E. (1990). *TEC1*, a gene involved in the activation of Ty1 and Ty1-mediated gene expression in *Saccharomyces cerevisiae*: Cloning and molecular analysis. Mol. Cell. Biol. **10,** 3541–3550.

Mirabito, P.M., Adams, T.H., and Timberlake, W.E. (1989). Interaction of three sequentially expressed genes control temporal and spatial specificity in *Aspergillus* development. Cell **57,** 859–868.

Murre, C., McCaw, P.S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in the immunoglobulin enhancer binding, *daughterless*, *MyoD*, and *myc* proteins. Cell **56,** 777–783.

Pabo, C.O., and Sauer, R.T. (1984). Protein-DNA recognition. Annu. Rev. Biochem. **53,** 293–321.

Pan, T., and Coleman, J.E. (1990). GAL4 transcription factor is not a "zinc finger" but forms a $Zn(II)_2Cys_6$ binuclear cluster. Proc. Natl. Acad. Sci. USA **87,** 2077–2081.

Qian, Y.Q., Billeter, M., Otting, G., Müller, M., Gehring, W.J., and Wüthrich, K. (1989). The structure of the *Antennapedia* homeodomain determined by NMR spectroscopy in solution: Comparison with prokaryotic repressors. Cell **59,** 573–580.

Sewall, T.C., Mims, C.W., and Timberlake, W.E. (1990). *abaA* controls phialide differentiation in *Aspergillus nidulans*. Plant Cell **2,** 731–739.

Sturm, R.A., and Herr, W. (1988). The POU domain is a bipartite DNA-binding structure. Nature **336,** 601–604.

Vinson, C.R., Sigler, P.B., and McKnight, S.L. (1989). Scissor-grip model for DNA recognition by a family of leucine zipper proteins. Science **246,** 911–916.

Xiao, J.H., Davidson, I., Matthes, H., Garnier, J-M., and Chambon, P. (1991). Cloning, expression, and transcriptional properties of the human enhancer factor TEF-1. Cell **65,** 551–568.