# Is EIAV Tat Protein a Homeodomain?

Recently two publications (*1*) reported RNA-binding capacity of the *Drosophila* homeotic bicoid protein bcd (*2*). We would like to point out that the reverse discovery, namely that a protein turned out to be a homeodomain-like structure after it was discovered as an RNA binding factor, happened in our laboratory.

The HIV-1 *trans*activator protein (h1-tat) is considered a tat response region (TAR) RNA binding protein, and the equine infectious anemia virus (EIAV) tat protein (e-tat) was inferred to play a role similar to the h1-tat protein (*3*). We recently determined the three-dimensional features of both these proteins by nuclear magnetic resonance (NMR) spectroscopy (*4*) and found that e-tat contains helical secondary structure as a limit structure. This limit structure was found to be dramatically stabilized by the addition of trifluoroethanol (tfe) (*5*). The secondary structure of the e-tat protein in tfe showed clearly a helix-loop-helix-turn-helix binding motif similar to the homeodomain protein motif (*6*, *7*). Comparison between helix motif proteins sequences, including bcd, and e-tat showed some resemblance, but no striking homologies, as shown by the alignment of the bcd and e-tat sequences as determined with the ClustalV program (*8*) and corrected for a continuous helix from Ala 28 to Lys 37 in *bcd* (*7*) (Fig. 1). Thus, e-tat and *bcd* are not closely related on the sequence level (12% identity and 44% similarity).

On the level of secondary structure, however, the aligned sequences of *bcd* and e-tat showed unexpected homology (Fig. 1). The recognition helix of the homeodomain proteins, helix III from the $NH_2$-terminus, corresponds to the basic sequence region of e-tat considered to be responsible for TAR recognition. In addition to the similarity of the helix patterns, both proteins show a strictly conserved turn between helices II and III. Unexpectedly, four out of six amino acid residues responsible for contact to DNA in *bcd* are identical in e-tat (Arg 5-Arg 5, Lys 50-Lys 57, Asn 51-Asn 58, Arg 54-Arg 61), one is nearly identical [Ile 47-Leu 54; the same substitution was observed for the fungal *b2* homeotic protein (*7*)], and only Arg 3 of bcd does not have a direct counterpart in e-tat, although it may well be that Arg 4 of e-tat plays a role similar to Arg 3 of the homeotic proteins. Frequently, in homeodomains, residue 4 is basic, making the homology to e-tat even more striking.

In addition to these conserved amino acids, the highly conserved pattern of hydrophobic amino acids in homeotic proteins, 31-*xxx***HH***xx***H***xxxxxx***H***xx***H***xx*-50, where ***H*** denotes a hydrophobic residue and *x* is either Gly, Cys, or Ala, is also conserved in e-tat (boxed residues in the sequence alignment). This pattern is clearly not matched in h1-tat.

Formation of a homeodomain-like recognition helix would not be possible in the basic region of bovine immunodeficiency virus tat (b-tat) and h1-tat protein because of the lack of a "spacer" of four amino acids, Asp-Ala-Ser-Leu, between the core and basic regions that is present in e-tat (*9*) (Fig. 2). These amino acids extend the length of the e-tat basic region helix to a size similar to that of the homeodomain recognition helix. These amino acids are necessary for the identical positioning of the basic sequence region in homeodomain and e-tat proteins. b-tat and h1-tat proteins are indeed suspected to obtain extended conformations even on RNA binding (*10*).

On the basis of this resemblance, we are currently testing the DNA-binding abilities of the EIAV Tat protein. Preliminary results show specific affinity of this protein to DNA sequences from the viral long terminal repeat with a dissociation constant much lower than the dissociation constant of the e-tat and e-TAR complex.

***Paul Rösch***
***Dieter Willbold***
*Department of Biopolymers,*
*University of Bayreuth,*
*95440 Bayreuth, Germany*
*e-mail: paul.roesch@uni-bayreuth.de*

```
                50          60
h1-tat FITKGLGISYG----RKKRRQRRRPSQGG
e-tat  FL-RSLGIDYLDASLRKKNKQRLKAIQQG
         40         50         60
```

**Fig. 2.** Core and basic regions of h1-tat and e-tat proteins as aligned with the Clustal V (*8*) program. e-tat contains an additional "spacer" of four amino acids as indicated. Spacer and the missing Thr-residue in the core region may be necessary for the correct arrangement of helices in e-tat.

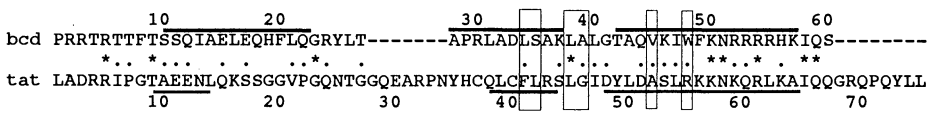


**Fig. 1.** Clustal V (*8*) alignment of bcd and e-tat proteins. Helices in the structures (*5*, *7*) are indicated by bars above and below the sequence, respectively. Conserved hydrophobic pattern in homeodomain proteins (boxed amino acids) is well conserved in e-tat.

## REFERENCES

1. J. Dubnau and K. Struhl, *Nature* **379**, 694 (1996); R. Rivera-Pomar, D. Niessing, U. Schmidt-Ott, W. J. Gehring, H. Jäckle, *ibid.*, p. 746.
2. T. Berleth *et al.*, *EMBO J.* **7**, 1749 (1988).
3. R. M. Stephens, D. Derse, N. R. Rice, *J. Virology* **64**, 3716 (1990); S. Noiman *et al.*, *Virology* **176**, 280 (1990).
4. D. Willbold, R. Rosin-Arbesfeld, H. Sticht, R. W. Frank, P. Rösch, *Science* **264**, 1584 (1994); P. Bayer, M. Kraft, M. Westendorp, R. W. Frank, P. Rösch, *J. Mol. Biol.* **247**, 529 (1995).
5. H. Sticht *et al.*, *Eur. J. Biochem.* **218**, 973 (1993); H. Sticht, D. Willbold, A. Ejchart, R. Rosin-Arbesfeld, P. Rösch, *Eur. J. Biochem.* **225**, 855 (1994).
6. S. E. Ades and R. T. Sauer, *Biochemistry* **34**, 14601 (1995); Y. Q. Qian *et al.*, *Cell* **59**, 573 (1989).
7. A. Laughon, *Biochemistry* **30**, 11357 (1991).
8. D. G. Higgins, *Methods Mol. Biol.* **25**, 307 (1994).
9. K. A. Jones and B. M. Peterlin, *Annu. Rev. Biochem.* **63**, 717 (1994).
10. J. D. Puglisi, L. Chen, S. Blanchard, A. D. Frankel, *Science* **270**, 1200 (1995); X. Ye, R. Kumar, D. J. Patel, *Chem. Biol.* **2**, 827 (1995); A. U. Metzger *et al.*, *FEBS Lett.* **384**, 255 (1996).

29 March 1996; accepted 16 April 1996