Homeotic gene *Antennapedia* mRNA contains 5′-noncoding sequences that confer translational initiation by internal ribosome binding

Soo-Kyung OH,[1] Matthew P. Scott,[2] and Peter Sarnow[1,3]

[1]Department of Biochemistry, Biophysics and Genetics, [2]Department of Microbiology and Immunology, University of Colorado Health Sciences Center, Denver, Colorado 80262 USA; [2]Departments of Developmental Biology and Genetics, Stanford University School of Medicine, Stanford, California 94305-5427 USA

The *Antennapedia* (*Antp*) homeotic gene of *Drosophila melanogaster* has two promoters, P1 and P2. The resulting *Antp* mRNAs contain 1512-nucleotide (P1) and 1727-nucleotide (P2) 5′-noncoding regions, composed of exons A, B, D, and E (P1) or exons C, D, and E (P2), respectively. Multiple AUG codons are present in exons A, B, and C. We have found that 252-nucleotide exon D, common to mRNAs from both transcription units and devoid of AUG codons, can mediate initiation of translation by internal ribosome binding in cultured cells. Many mRNAs in *Drosophila* contain long 5′-noncoding regions with apparently unused AUG codons, suggesting that internal ribosome binding may be a common mechanism of translational initiation, and possibly its regulation, in *Drosophila*.

[*Key Words*: Internal ribosome binding; homeotic gene; *Antennapedia* mRNA; dicistronic mRNA; RNA transfection]

Received May 18, 1992; revised version accepted July 7, 1992.

The homeotic genes required for segments in *Drosophila melanogaster* to have different developmental fates are expressed in some segment primordia but not in others. Each segmented part of the embryo expresses a different homeotic gene, or combination of them (for review, see Duncan 1987; Mahaffey and Kaufman 1988). Transcription of the homeotic genes is tightly regulated by the activities of a large number of regulators, including many of the segmentation genes (Scott and Carroll 1987; Ingham 1988; Irish et al. 1989). Transcription of a homeotic gene in a location where the gene is normally silent can lead to transformations of segments (Frischer et al. 1986; Schneuwly et al. 1987a,b). Thus, transcriptional regulation is clearly involved in the spatially regulated activities of homeotic genes. The sequence organization of certain homeotic mRNAs suggests that there may be an additional level of regulation.

Transcription of the homeotic *Antennapedia* (*Antp*) gene is initiated at two promoters (P1 and P2), producing transcripts that differ in their 5′-noncoding regions (5′ NCR) (Laughon et al. 1986; Schneuwly et al. 1986; Stroeher et al. 1986). Figure 1 illustrates that P1-initiated transcripts contain a 1512-nucleotide 5′ NCR derived from sequences in non-protein-coding exons A, B, D, and E, whereas P2-initiated transcripts contain a 1727-nucleotide 5′ NCR derived from exons C, D, and E. The noncoding exons A and B contain 8 AUG codons preceding the translation initiator AUG codon located in exon E; similarly, exon C contains 15 upstream AUG codons. All of these upstream AUG codons are followed by very short open reading frames that could potentially encode peptides ranging from 6 to 44 amino acids. Noncoding exon D and the 5′ half of exon E, common to mRNAs initiated at both P1 and P2, are devoid of AUG codons (Laughon et al. 1986; Stroeher et al. 1986). It is remarkable that the large number of total upstream AUG codons in these 5′-noncoding regions does not prevent translation of the main open reading frame. According to the scanning model for translational initiation (Kozak 1989, 1991), translational initiation of *Antp* mRNA should be very inefficient.

If the scanning mechanism were used for translational initiation of *Antp* mRNA, derived from the P2 promoter, for example, the ribosomal 43S ternary complex would bind at the 5′ end of the mRNA and would scan 1730 nucleotides of the 5′ NCR, bypassing 15 AUG codons, to initiate protein synthesis at the sixteenth AUG codon. Several of the 15 upstream AUG codons are in a favorable context to initiate protein synthesis in *Drosophila* cells (Table 3, below; Cavener 1987). In addition, RNA structures in this long leader may render the scanning of the ribosomal subunits inefficient.

There are precedents for translational initiation in mRNAs with multiple AUG codons in their 5′ NCRs, by both ribosomal reinitiation and internal ribosome binding mechanisms. For example, yeast GCN4 mRNA con-