Structural analysis of Doc transposable elements associated with mutations at the white and suppressor of forked loci of Drosophila melanogaster

Alan Driver*, Simon F. Lacey, Tim E. Cullingford, Andrew Mitchelson, and Kevin O'Hare

Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AZ, UK

Summary. DNA sequences from two spontaneous mutations of Drosophila melanogaster associated with insertion of a Doc transposable element have been cloned. In white-one, the element is inserted in the white locus close to where transcription initiates. In a lethal allele of suppressor of forked, su(f)S2, the element is inserted within the transcription unit in the protein coding region. Four other Doc elements have been cloned from a wild-type strain. Doc is a member of the class of transposable elements known as retroposons, which includes the D. melanogaster F, G, Jockey, and I elements. There is no sequence homology between the ends of the Doc element. The 3' or right end terminates with a polyadenylation signal sequence followed by a stretch of oligo-A. The length of the oligo-A varies between elements, and a duplication of variable size is found as a direct repeat flanking inserted Doc elements. Members of the family are conserved at the 3' end, but may be truncated at the 5' or left end. These structural features suggest a mechanism of transposition via an RNA intermediate.

Key words: Transposable element – Mutation – Retroposon

Introduction

The mutation white-one (w1) abolishes pigmentation in the eye of the fruitfly Drosophila melanogaster and was isolated by Morgan (1910). This was first mutation to be characterised in D. melanogaster, initiating the study of genetics in this organism (Morgan and Bridges 1916). In this paper we describe the cloning of w1, and show that it is associated with the insertion of a previously described transposable element, the Doc element. We also present structural analysis of the suppressor of forked [su(f)] locus, su(f)S2 (Schalet 1986).

Materials and methods

Drosophila strains. white-one from Bowling Green (Morgan 1910); suppressor of forked-S2/FM7 from A. Schalet (Schalet 1986).

Results

Cloning and analysis of the w1 allele

Zachar and Bingham (1982) analysed the structure of the w1 allele by DNA blotting and found it to be associated with an insertion in the proximal region of the white (w) locus. Our own DNA blotting experiments were in general agreement with this and showed that the insertion lacked BamHI sites (data not shown). We therefore prepared an EMBL4 library of BamHI fragments from w1 and screened with a probe from this region of w. Several clones which contained identical BamHI fragments were purified. This fragment was cloned into pBR322 to give the recombinant plasmid. Comparison with the wild-type map showed there was a 4.9 kb insert between the HindIII site at +3.2 and the BamHI site at +4.4 (see Fig. 1 for map). The co-ordinate system used is that of Levis et al. (1982).

To characterise better the position and nature of the insert, the DNA sequences of the junctions between the insert and the flanking w sequences were determined (Fig. 2). Transcription of w is from left to right for the orientation of the element shown in Fig. 1 and the sequences reported correspond to the sense strand of w. The sequence at the right end has a stretch of oligo-A preceded by two polyadenylation signal sequences (AATAAA). Because of the similarity of this region to the 3' ends of mRNAs, we will refer to this as the 3' end of the element. This sequence is identical to that we reported from a partial clone of white-honey (w0), a derivative of w1 (Zachar and Bingham 1982; O'Hare et al. 1984). The sequence at the left end of the w1 insertion (5' end of the element) shows

Recombinant DNA techniques. Lambda libraries were made using the vector EMBL4 digested with BamHI and SAI. The insert DNA was unfractionated BamHI fragments or size selected Sau3A1 partial digest products purified from gels made with low gelling temperature agarose. After ligation and packaging, the phage were plated on the Escherichia coli strain Q359 to eliminate non-recombinants. Further rounds of phage purification and growth were carried out using the E. coli strain C600. DNA sequencing was by standard dideoxy techniques on single stranded DNA prepared from M13 recombinants containing restriction enzyme fragments as inserts. Probes for DNA blotting and phage purification were made either by nick-translation of plasmid recombinants or primed synthesis using the universal primer on M13 recombinants. For details and references, see Berger and Kimmel (1987).
no obvious homology, either inverse or direct, to the sequence at the right end (3' end of the element). A comparison of these sequences shows that the inserted element is bounded by direct duplication of a sequence from the w locus. The duplication could be as large as 13 bp (AATGTCCGCCTTC) or as small as 11 bp (TGTCGCCTTC), as the first two A residues could be part of the oligo-A tract or part of the duplicated sequence. For the insertion in w°, the 13 bp sequence present in the wild-type w locus at position +3700 to +3688 is now present at both ends of the insertion. Nucleotide positions are those of O'Hare et al. (1984).

The map matches that of the first Doc element isolated by Bender et al. (1983) from the Bithorax complex. The sequences match those reported by Schneuwly et al. (1987) for two Doc elements associated with an inversion between cytological locations 84B and 84D in the mutant Ant t 36 of the Antennapedia complex.

**Isolation and analysis of other Doc elements**

In DNA blotting experiments using probes from the su(f) locus, an insertion, which appeared to be a Doc element, was found associated with the spontaneous lethal mutation su(f)s2 (A. Mitchelson and K. O'Hare, unpublished results). A library was made using DNA from the heterozygous stock and screened using probes from the wild-type su(f) locus. Clones of the mutant gene were then identified by counterscreening with pw 1°. The insertion in su(f)s2 has the same map as that of the Doc in w° (Fig. 1). The junction sequences between the Doc element and the flanking su(f) sequences were determined for su(f)s2 and com-

---

**Fig. 2. DNA sequences of the termini of Doc elements.** The DNA sequences are aligned and dashes used to represent gaps introduced to maximise the alignment. The Doc84B and Doc84D sequences are from Schneuwly et al. (1987). The element sequences are shown as upper case letters, the flanking duplications as underlined upper case letters and the rest of the flanking sequences as lower case letters. Asterisks show point polymorphism at the 3' ends. The length of the oligo-A at the 3' end of the different elements is indicated by the subscript