Rearrangements Involving Repeated Sequences Within a *P* Element Preferentially Occur Between Units Close to the Transposon Extremities

Frédéric Pâques,¹ Bruno Bucheton and Maurice Wegnez

Laboratoire d'Embryologie Moléculaire et Expérimentale, Unité Associée 1134 du Centre National de la Recherche Scientifique, Université Paris XI, 91405 Orsay Cedex, France

Manuscript received January 20, 1995

Accepted for publication October 23, 1995

ABSTRACT

In a previous report we described rearrangements occurring at a high rate (30% of the progeny of dysgenic flies) within a cluster of 5S genes internal to a *P* element. These events were characterized as precise amplifications and deletions of 5S units. Here we analyze recombination events within *P* elements containing two repeated arrays of 5S genes flanking a central *white* gene. Deletions (50%) and duplications (3%) of the *white* gene together with various amounts of flanking 5S genes were observed. These recombinations occur preferentially between the most external 5S units of *P* transposons. Such rearrangements could be favored by interactions between the proteins bound to the *P* terminal sequences.

P-*M* hybrid dysgenesis in *Drosophila melanogaster* is characterized by the transposition of *P* elements (for reviews, see ENGELS 1989; RIO 1990). In addition to transposition, hybrid dysgenesis also induces a number of genetic recombination events including chromosomal rearrangements, crossing over in males, and *P* excisions (ENGELS 1989). Two types of excisions arise in a dysgenic context: imprecise excisions, which remove various extents of *P* and flanking sequences and precise excisions, which occur by conversion on a transposon-free template (ENGELS *et al.* 1990; GLOOR *et al.* 1991). These events are probably initiated by cutting of the transposon, which gives rise to a double-strand gap. *P* elements probably then transpose by a *cut-and-paste* pathway, while their original insertion sites are restored by gene conversion. Similar results have been found for the *Tc1* transposon in *Caenorhabditis elegans* (PLASTERK 1991; PLASTERK and GROENEN 1992).

Crossing over is often associated with gene conversion and is thought to be a consequence of Holliday junction cutting (HOLLIDAY 1964; MESELSON and RADDING 1975; SZOSTAK *et al.* 1983). However, gene conversions without crossing over have been described at the *MAT* locus and at other sites in *Saccharomyces cerevisiae* (KLAR and STRATHERN 1984; KLEIN 1984; WHEELER *et al.* 1990). In the case of *P*-induced conversions on a transposon-free template, no crossing over occurs between either allelic or ectopic sites (ENGELS *et al.* 1990; GLOOR *et al.* 1991). Another homologous recombination process, *Single-Strand Annealing* (SSA), involves homologous sequences flanking a double-strand break

(LIN *et al.* 1984, 1990; CARROLL *et al.* 1986; MARYON and CARROLL 1991a,b; SUGAWARA and HABER 1992). This process is nonconservative because it results in a loss of sequences. Like *P*-induced conversions, SSA is not associated with crossing over.

HASTINGS (1988) and ENGELS *et al.* (1990) proposed a modification of the *Double-Strand Break Repair* model of SZOSTAK *et al.* (1983) to explain gene conversion without crossing over. In this model, the conversion structure is not resolved by Holliday junction cutting; instead annealing takes place between the 3'-extending ends. This model thus combines conversion and SSA in a single pathway. ENGELS *et al.* (1990) also proposed that imprecise *P* excisions would be a byproduct of gene conversion between sister chromatids. Neosynthesized strand matching could occur between short complementary sequences of the 3'-extending ends. More recently, NASSIF *et al.* (1994) proposed that the neosynthesis associated with *P*-induced conversions could occur according to the *Bubble Migration* model of FORMOSA and ALBERTS (1986). This model assumes that during neosynthesis, the new DNA strand is continuously removed from the template. Such a replicative process, associated with gene conversion, could result in two complementary single-stranded DNA, which could match.

In a previous report, we have described rearrangements occurring at a high rate (30% of the progeny of dysgenic flies) in a dysgenic context within a tandemly repeated cluster of *D. teissieri* 5S genes internal to a *P* element (PÂQUES and WEGNEZ 1993). These rearrangements were characterized as precise amplifications and deletions of 5S units. Recently, KURKULOS *et al.* (1994) showed that two direct repeats of 276 bp within a *P* element promoted the deletion of intervening sequences, at a frequency of 3%. THOMPSON-STEWART *et*

Corresponding author: Maurice Wegnez, Laboratoire d'Embryologie Moléculaire et Expérimentale, Université Paris XI, Bâtiment 445, 91405 Orsay Cedex, France.

¹ Present address: Rosenstiel Center, Brandeis University, 415 South St., Waltham, MA 02254-9110.

FIGURE 1.—Constructions and strains. (A) A polymer of tandemly repeated 5S genes was constructed *in vitro* as described in MATERIALS AND METHODS. This polymer, inserted into Bluescript (pCA4), contains seven 5S units: six complete units (376 bp) and one truncated unit (216 bp). The truncated unit will be in all cases counted as one unit when determining the size of 5S arrays. The two spotted boxes of the *HinPI* monomeric 5S fragment used for the construction correspond to the 5S coding region. Open arrowheads indicate the orientation of 5S genes. (B) The polymer of pCA4 was used to build CAR1, CAR2, and CAR3 transposons (see MATERIALS AND METHODS). These transposons were then used to transform a *white* strain of *D. melanogaster*. CAR1 in G13.2 and CAR3 in Z25.F are on the X chromosome, while CAR2 in H38.3 is on chromosome III. The *P* ends are symbolized by the large (5') and small (3') thick arrows, respectively. The thin arrow indicates the orientation of the *white* gene, represented by an open box. Sequences surrounding insertion sites are drawn as thin lines. (C) Structure of the rearranged CAR1 transposon of the strain E22 (see the RESULTS). This rearranged transposon contains a duplication of the *white* gene and an internal 5S array comprising eight 5S units. The probes used for the molecular characterizations are shown as numbered open boxes: 1, 5S probe; 2, 5' *P* probe; 3, 3' *P* probe; 4, *white* probe. Restriction sites: A, *AseI*; B, *BamHI*; E, *EcoRI*; EV, *EcoRV*; H, *HindIII*; Hi, *HinPI*; P, *PstI*; S, *SalI*; Sa, *Sau3A*; Sm, *SmaI*; Sc, *ScaI*; X, *XhoI*.

al. (1994) observed dysgenesis-induced deletions and duplications within a tandemly repeated array carrying a *P* insertion. All these observations fit the model of ENGELS *et al.* (1990).

In this study, we introduced a *P* element containing two repeated arrays of *D. melanogaster* 5S units flanking a central *white* marker gene into a dysgenic context. As previously, homologous recombination events were induced at a very high frequency. Deletions (50%) and duplications (3%) of the *white* gene together with various amounts of flanking 5S genes were observed. However, rearrangements occurred preferentially between the most external 5S units. This suggests the existence of molecular interactions between the proteins involved in hybrid dysgenesis that would be located in the vicinity of the two *P* transposon extremities. These protein interactions could also promote matching of the neosynthesized strands during the conversion process.

MATERIALS AND METHODS

Constructions: Restriction fragments were recovered from agarose gels using the GeneClean kit. Ligations were made with the T4 DNA ligase. BRL competent DH5 α cells were utilized for transformation. Sequencing was done according to the Sanger method with the Stratagene sequenase kit.

pCA4: A tandem array of 5S genes was constructed *in vitro*. We started with a previously cloned *D. melanogaster* 5S dimer (SAMSON and WEGNEZ 1988) that was digested with *HinPI*. The *HinPI* fragment, corresponding to a single 5S unit (Figure 1A), was polymerized *in vitro* with the T4 DNA ligase, and polymers were cloned at the *Clal* site of Bluescript. One clone containing a 5S tetramer, pCA1 (Figure 1A), was mapped with *EcoRV*, *ScaI* and *Sau3A*. This analysis demonstrated that the four 5S units were in the same orientation. This clone was partially digested with *EcoRV* and the two largest fragments containing 5S units were recovered and ligated. Ligation products were cloned at the *EcoRV* site of Bluescript. A clone containing six complete 5S units (376 bp *EcoRV* fragments) and a 216-bp *EcoRV/HinPI* 5S fragment was obtained. This clone, pCA4 (Figure 1A), was mapped and shown to contain

FIGURE 2.—Experimental design to study rearrangements within CAR transposons. Females of the G13.2, Z25.F and E22 strains, carrying their transposon on the X chromosome, were crossed, as indicated on the left, with males carrying the $P[ry^+, \Delta 2-3]$ element. F_1 dysgenic males (boxed genotype) were crossed with *white* females, and the Sb^+ F_2 progeny was checked for eye phenotypes. F_1 male progeny from the G13.2 strain was also crossed with females carrying a compound X chromosome (XX/Y). Females of the H38.3 strain, carrying their transposon on chromosome III, were crossed as indicated on the right.

all 5S units arranged in the same orientation. All subsequent constructions were made with this 5S array.

CAR1: The *EcoRI*-*SaII* fragment of pCA4, containing the complete 5S array, was recovered. It was then inserted, after filling the protruding ends with the T4 DNA polymerase, into the *SmaI* site of pCA4. We checked that the two 5S arrays of this clone were in the same orientation by sequencing their distal extremities. We thus inserted, at the *EcoRI* site located between these two 5S arrays, an *EcoRI* fragment comprising

the *mini-white* gene from Caspew18 (PIRROTTA 1988). The *Bam*HI-*SaII* fragment containing the two 5S arrays and the intervening *mini-white* gene finally was inserted at the *Bam*HI-*XhoI* site of Carnegie 4 (RUBIN and SPRADLING 1983) to give the CAR1 transposon (Figure 1B).

CAR2: The construction strategy was that followed for CAR1. However, a dimer of the initial 5S array was first inserted at the *SmaI* site of pCA4. Mapping with *EcoRV* and *SaI* showed that all 5S arrays of CAR2 were in the same orientation (Figure 1B).

FIGURE 3.—Molecular characterization of transposons from the G13.2 w progeny. Genomic DNA (10 flies for each strain) was digested with *EcoRI* and *PstII*, fractionated on a 0.8% agarose gel, and hybridized with three probes. Lane 1, G13.2 control; lanes 2–12, w strains. (A) Hybridization pattern with the 5' *P* probe (probe 2, Figure 1). The presence

CAR3: The *EcoRI*-*Sall* fragment of pCA4, containing the 5S array, was cloned at the *EcoRI*-*XhoI* site of Casper 4 (PIRROTTA 1988) (Figure 1B).

Germ line transformation: The *w1118* strain was transformed with *CAR1*, *CAR2* and *CAR3* as described by RUBIN and SPRADLING (1982). The *Sb P[ry⁺Δ2-3](99B)/TM6* strain was used for dysgenic crosses (ROBERTSON *et al.* 1988). All crosses were performed at 25°C.

DNA preparation and molecular analyses: Fly DNA extractions were done as described by PÂQUES and WEGNEZ (1993). Restriction digests were fractionated on 0.8 or 1% agarose gels and DNA transferred to Hybond-N membranes (Amersham), or fractionated on a 5% polyacrylamide gel and blotted as described by PRÉAT (1990). The different probes used are shown in Figure 1.

RESULTS

New eye phenotypes appear at a high frequency in dysgenic progeny of transformed flies carrying a *P* element containing the *white* gene inserted between two 5S gene clusters: The *CAR1* transposon contains two *D. melanogaster* 5S arrays flanking the *white* marker gene (Figure 1). We obtained the G13.2 transgenic strain by transforming *w1118* with *CAR1*. The G13.2 eye phenotype (red) was found to be perfectly stable in 15,900 surveyed G13.2 males. We then looked for eye color changes within the dysgenic progeny of G13.2 flies. Such changes would be indicative of rearrangements between the two 5S arrays resulting in deletion or duplication of the *white* gene.

G13.2 females, carrying the *CAR1* transposon on the X chromosome, were mated with males possessing the *P[Δ2-3]* transposase source. As shown in Figure 2, individual F₁ dysgenic males were mated with *white* females. The G13.2 X chromosome was then recovered only in F₂ females. The presence of males with a red eye phenotype in the F₂ progeny is thus due to the transposition of the *P* element. Their frequency gives an estimate of the transposition rate of *P* elements carrying an intact

of a 1-kb fragment (arrow) indicates that the 5' *P* end and flanking sequences of the transposon are not modified. In some strains, additional *EcoRI*-*PstI* hybridizing fragments (lanes 2, 4, 6, 7, 9, 10 and 12) can be explained by the presence of new transposons at ectopic sites. The two weak signals ~4 kb in lanes 1 and 9 are due to residues of prior hybridization with *white* (not shown). (B) Hybridization pattern with the 3' *P* probe (probe 3, Figure 1). The presence of a 280-bp fragment (arrow) indicates that the 3' *P* end and flanking sequences are not modified. As in A, the pattern demonstrates the presence of new transposons at ectopic sites. In all strains, a weak signal ~1 kb probably corresponds to a non-specific hybridization with fragments visible by ethidium bromide staining (not shown). (C) Hybridization pattern with the 5S probe (probe 1, Figure 1). Sizes of the fragments hybridizing with the 5S probe (arrows): 650 bp, lanes 2, 6, 7 and 8; 1.03 kb, lanes 7 and 11; 1.40 kb, lane 10; 1.77 kb, lanes 3, 9 and 12; 2.50 kb, lanes 1, 4 and 9. In all strains, weak signals ~1 kb are visible and probably correspond to non-specific hybridization with the fragments that also give signals with the 3' *P* probe in B. The stronger signals in lanes 7 and 11 would be due to specific hybridization with the 5S probe.

FIGURE 4.—Molecular characterization of bright red strains carrying a rearranged *CAR1* transposon at the original site. Ten bright red strains carrying a single transposon were analyzed. Restriction fragments were fractionated on a 0.8% agarose gel and hybridized with the 5S probe. Lane 1, G13.2 control; lanes 2–11, bright red strains. (A) Analysis of *SaII* fragments. The two signals ~5 kb in G13.2 lane are due to the presence of one *SaII* site within the *white* gene (see structure of *CAR1*, Figure 1). These two signals are conserved in eight of the 10 analyzed strains (grey arrows). Persistence of these signals indicates that the two external 5S arrays as well as the *P* ends and their flanking sequences were not rearranged and that the *white* genes are in the same orientation. These eight strains also display one additional signal corresponding to a *SaII* fragment of variable size (6.6–9.5 kb, lanes 2, 3, 5–10). One of the two shared *SaII* fragments is missing in the strain analyzed in lane 11 (upper arrow). Both are missing in the strain analyzed in lane 4. In all strains, large *SaII* fragments (>20 kb) give strong signals corresponding to the endogenous 5S cluster. (B) Analysis of *EcoRI* fragments. The two signals ~2.5 kb in G13.2 lane correspond to the two *EcoRI* fragments of *CAR1* containing the 5S arrays (see structure of *CAR1*, Figure 1). These two signals are conserved in eight of the ten analyzed strains (grey arrows). These eight strains also display one additional signal corresponding to an *EcoRI* fragment of variable size (black arrows). One of the two shared *EcoRI* fragments is missing in the strain analyzed in lane 11 (upper arrow). In this strain, the presence of several additional signals demonstrates the presence of a new transposon at an ectopic site. Both of the two shared *EcoRI*

white marker. This value, 5.6% ($4.5 \times 5/4$), is an underestimation of the total transposition rate because rearranged newly inserted transposons, which have lost the *white* gene, also appear with a high frequency (see below).

As shown in Figure 2, three eye phenotypes were observed among F_2 females. The red eye phenotype, observed in 48.2% of the females, is that expected for the presence of one intact G13.2 X chromosome. The white (*w*) phenotype characterizing 48.0% of the females could be explained by the loss of or damage to the *CAR1 white* marker. The bright red phenotype found in 3.8% of the females could be due to a dosage effect (the presence of more than one copy of the *white* marker) or to a position effect (affecting a new inserted transposon).

We also crossed F_1 dysgenic males with females carrying a compound X chromosome (XX/Y females in Figure 2) to establish strains carrying the X chromosome from the G13.2 strain. The three kinds of phenotypes were observed among F_2 males with frequencies similar to those observed when crossing with *w* females (52.0, 42.5 and 5.5% instead of 48.0, 48.2 and 3.8% for *w*, red and bright red phenotypes, respectively).

Molecular analysis of the *CAR1* transposon of 135 strains issued from *w*, red and bright red flies was done. All *w* strains we characterized are derived from independent F_1 dysgenic males, as well as all red and all bright red strains. Thus, the molecular events which were observed are independent events.

Deletion of *white* in transposons carrying the gene inserted between two 5S clusters occurs at a very high rate in dysgenic context: As pointed out above, the *w* phenotypes could be explained by a deletion of the *CAR1 white* marker gene. To test this hypothesis, we analyzed *w* male progeny from 36 independent stocks. As shown for 11 strains in Figure 3, various rearrangements of *CAR1* were observed. These rearrangements are characterized by the removal of the *white* gene but also by significant deletions of the surrounding 5S genes (Figure 5A). In a single case, one *w* strain was shown to contain a newly inserted nonrearranged *CAR1* transposon. The *w* phenotype of this strain can be explained by an insertion of *CAR1* into the heterochromatin or at a silenced site or by a point mutation. According to the structure of *CAR1* transposons and of their surrounding sequences, rearrangements can be classified into three categories.

Deletions by homologous recombination: In 32 out of 36 analyzed *w* strains, the rearrangement was internal to the *P* ends. A single 5S array was observed, displaying various sizes. As those sizes differ by 370–380 bp (the length of one 5S unit) increments, it can be inferred

fragments are missing in the strain analyzed in lane 4. In all strains, large *EcoRI* fragments (± 20 kb) give strong signals corresponding to the endogenous 5S cluster.

A) w phenotype: 52.0%
36 strains characterized

Deletions by homologous recombination

Deletions by illegitimate recombination

Deletions by homologous recombination and rearrangement affecting flanking sequences

B) red phenotype: 42.5%
84 strains characterized

in situ deletion + insertion of an intact transposon at a new site

C) bright red phenotype: 5.5%
15 strains characterized

intact CAR1 *in situ* + insertion of an intact transposon at a new site

Duplication of *white* in transposons carrying the gene inserted between two 5S clusters: The bright red phenotypes (Figure 2) could be due to an increase of the *white* gene copy number, either by the transposition of CAR1 or by a duplication of the gene within the transposon. Molecular analyses of 15 independent strains demonstrated that both phenomena occurred.

Transposition of *CAR1*: In five strains, intact *CAR1* transposons were shown to be present at their original site and also at an ectopic site. The bright red phenotypes thus can be attributed to the presence of two transposons (Figure 5C).

Internal duplication of the *white* gene within *CAR1*: A single transposon, larger than *CAR1*, was detected in the 10 remaining strains. Extensive molecular analyses were performed on these strains. Analysis of *SaII* and *EcoRI* digests with a 5S probe is shown in Figure 4. In nine strains, rearranged *CAR1* transposons at the original site contain two *white* genes and three 5S arrays (Figure 5C). The 3' flanking sequence of the transposon is modified in one of these strains. The two external 5S arrays of rearranged transposons are the same as those of *CAR1*, and the duplicated *white* genes are in the same orientation. Size variations of the new internal 5S array are due to different numbers of gene units. We demonstrated, for the E22 strain, that all 5S units of the transposon are in the same orientation.

Complex rearrangement: Rearrangements of *CAR1* in one strain are very complex (Figure 4, lane 4; and Figure 5C). The 3' flanking sequences of *CAR1* are conserved, indicating that the rearrangements occurred at the original site. The transposon includes two *white* sequences in inverted orientation, separated by an 8-kb segment comprising a complex array of 5S units. The 5' *P* end has been lost. The 5S arrays at the 3' side of both *white* sequences have been rearranged (five and six gene units, respectively, instead of seven). This is the only detected case of intra-array rearrangement affecting *CAR1* 5S genes.

Lack of intra-array rearrangements in transposons containing two 5S clusters surrounding the *white* gene: Rearrangements within one of the two 5S clusters of *CAR1* would not lead to a modification of the eye phenotype. Such events were expected to occur at a high frequency, given our previous results from transposons carrying a single repeated array of *D. teissieri* 5S genes (PÂQUES and WEGNEZ 1993). Molecular analysis of 84 red strains (48 independent crosses with *w* females and 36 independent crosses with XX/Y females, Figure 2) showed that such events are very rare when two 5S clusters are present in the same transposon. All results are summarized in Figure 5B.

No rearrangements were observed in 78 transposons located at their original site and in three transposons at ectopic sites. The only detected rearrangements occurred in three strains and affected only their 5' flanking sequences (Figure 5B). The 5S arrays of the *CAR1* transposons thus were not affected by hybrid dysgenesis context in all 84 analyzed red strains.

Similar results were obtained with *CAR2*. This transposon contains two 5S arrays of different sizes (Figure 1B). *CAR2*, inserted on the third chromosome of the H38.3 strain, was introduced into a dysgenic context (Figure 2). Analysis of 46 independent F₂ Sb⁺ red

FIGURE 6.—Synopsis of Z25.F progeny molecular analyses. The grey squares facing each transposon structure indicate the number of times it was found. Symbols used for 5S genes, *white*, *P* and original flanking sequences are those used in Figure 1. Restriction sites: E, *EcoRI*; H, *HindIII*; P, *PstI*.

strains demonstrated that all these strains contained a single transposon, with no rearranged 5S arrays.

CAR3, which contains a single *D. melanogaster* 5S array, was used as a control (Figure 1B). This transposon was inserted into the X chromosome of one transgenic strain (Z25.F). Three eye phenotypes (*w*, red, and bright red) were also obtained in F₂ Sb⁺ progeny by crossing dysgenic *Sb CAR3* males with *w* females, but with significantly different frequencies (Figure 2). Analysis of 21 independent red derived strains demonstrated that the number of gene units in the 5S cluster of *CAR3* had changed in four strains (Figure 6). No rearrangements affect sequences flanking the 5S array. In two cases, a size increase of the 5S gene array was found, while in two other cases the size was reduced. The 5S intra-array rearrangement thus occurs with a high frequency (15.5% of total progeny) when the transposon contains a single 5S cluster.

Rearrangements within *P* are mainly promoted at the transposon extremities: The structure and size distribution of deleted *CAR1* transposons indicate that recombination preferentially occurs between the most external 5S genes of *CAR1* (Figure 5). To test this apparent bias in the deletion breakpoint distribution on a larger scale, we put the transposon of the E22 strain into a dysgenic context (Figure 2). This strain is one of the bright red strains obtained with the G13.2 transposon. The transposon contains two *white* genes and three 5S arrays in the same orientation (Figure 1C). Homologous recombination between repeated sequences internal to the transposon is thus expected to yield different phenotypes. Rearrangements between *white* genes or between one external 5S array and the internal 5S array would remove only one *white* gene. Such events would revert the phenotype to the red phenotype of the G13.2 strain. Rearrangements between the two external 5S arrays would remove both *white* genes, and result in a *w* phenotype.

As expected, the *w* and red phenotypes were obtained in the progeny of dysgenic flies carrying the E22 transposon (Figure 2). However, the rate of reversion to the red phenotype is only 0.3% (4/1275), as the rate of *w* flies is 40.5%. Two red flies were crossed, and their progeny molecularly analyzed. As expected, in these

strains, an inter-5S array rearrangement had removed one copy of the *white* gene. Deletion of the two *white* genes then occurs with a frequency 135 times greater than deletion of a single copy.

The parental bright red phenotype was also observed, as well as a new, reinforced phenotype, dark red (Figure 2). This new phenotype is probably due to an increase of the *white* gene copy number, either by transposition, or by further duplication within the *P* element.

DISCUSSION

In a previous work, we brought a *P* transposon carrying a single tandem array of *D. teissieri* 5S genes into a dysgenic context. Rearrangements were found to occur at a very high rate, leading to precise amplifications and deletions of 5S units within the cluster (PÂQUES and WEGNEZ 1993). Here we analyze, in the same genetic conditions, the rearrangements that occurred within transposons containing the *white* gene inserted between two tandem arrays of *D. melanogaster* 5S genes.

Deletions and duplications of *white* within CAR1 transposons: As shown in Figure 2, eye phenotypes within the dysgenic progeny of G13.2 flies can be grouped within three classes. We obtained w flies at a very high frequency (48–52%), while bright red flies account for 3.8–5.5% of the progeny. All analyzed CAR1 transposons of w flies were characterized by a complete deletion of the *white* gene, and most of these rearrangements were produced by homologous recombination between two units of the surrounding 5S clusters (Figure 5A). Deletion of the *white* gene was also observed in three of the 84 red strains. In these three cases, the red phenotype is due to the transposition of an intact CAR1 transposon to a new site (Figure 5B). The frequency of *white* deletion events can thus be calculated as $52.0\% + 3/84 \times 42.5\% = 53.5\%$. Duplication of the *white* gene together with various numbers of 5S genes was observed in 3.3% of G13.2 dysgenic progeny ($5.5\% \times 9/15$, Figure 5C).

The rearrangements affecting the *white* gene of CAR1 cannot be explained by unequal exchanges between sister chromatids, as there are much more deletions than duplications. Rather, they strongly support the model of ENGELS *et al.* (1990). Figure 7 shows how a gene conversion interrupted by matching of the 3'-extending ends could yield deletions as well as duplications of the CAR1 *white* gene. Dysgenesis-induced rearrangements between repeated sequences internal to a *P* element (PÂQUES and WEGNEZ 1993; KURKULOS *et al.* 1994) or surrounding a *P* insertion (THOMPSON-STEWART *et al.* 1994) have previously been reported and can be explained in a similar way. It must be pointed out that fork slippage during replication would result in the same kind of rearrangement. However, to explain our results, we have then to assume extraordinary high rates of slippages, which makes this explanation unlikely.

Do *P* ends interact during the rearrangement process? The CAR1 rearrangements we observed in this study differ markedly from those we obtained with a transposon carrying a single 5S array. Amplifications and deletions were found to occur within a single *D. teissieri* 5S array internal to a *P* element with a rate of ~30% (PÂQUES and WEGNEZ 1993). The same events were also observed with the CAR3 transposon containing one single array of *D. melanogaster* 5S genes. Rearrangements affecting only one of the two 5S clusters of CAR1 would not lead to a new eye phenotype, and could be detected when analyzing red flies of the G13.2 dysgenic progeny. As shown in Figure 5B, we did not observe such events among 84 independent strains. Rearrangements within individual 5S arrays of the CAR2 transposon were similarly not obtained in 46 red progeny of the H38.3 strain.

One structural feature that could bias the rearrangements within CAR1 is the presence of repeated 5S sequences in the vicinity of both transposon extremities (Figure 1). Analysis of the 5S cluster size in rearranged transposons gives interesting clues in favor of this hypothesis. First, deleted CAR1 transposons carrying a small number of 5S units significantly outnumber transposons with several 5S units (Figure 5A). Second, rearrangements within the E22 transposon seem to involve preferentially repeated sequences close to *P* ends. This CAR1 derived transposon, as shown in Figure 1C, is characterized by a direct duplication of the *white* gene and by the presence of three 5S arrays in the same orientation. Rearrangements removing one of the two *white* copies are very rare (0.3%). Such rearrangements could be the result of homologous recombination involving either the 5S array separating both *white* copies and one of the external 5S arrays, or the two *white* genes. On the contrary, rearrangements removing both *white* genes are much more frequent (40.5%).

These results suggest that matching of complementary neosynthesized strands occurs early in the conversion process. This could be explained by the *Bubble Migration* model proposed by FORMOSA and ALBERTS (1986) and NASSIF *et al.* (1994). This model assumes that, during the gene conversion process, neosynthesized strands are continuously unwound from the template, while DNA synthesis continues. With the CAR1 and CAR2 transposons, and the transposon of the E22 strain, complementary 5S sequences would be available very early during a conversion process. Their frequent matching would result in a high frequency of deletions. However, one prediction of this model is that large duplications would be less frequent than small ones. We observed the opposite situation: the internal 5S cluster in rearranged CAR1 transposons carrying a *white* duplication always contain more than seven 5S genes. If duplications could yield internal 5S arrays of random sizes, the probability of getting 9 arrays out of 9 containing more than seven 5S genes would be $(6/13)^9 =$

FIGURE 7.—Model for recombination in a dysgenic context of the CAR1 transposon. Excision of the transposon occurs on one chromatid in G2, and flanking sequences are single-stranded by an exonuclease (A, B). The double strand gap initiates a gene conversion on the sister chromatid (C) and *P* sequences are neo-synthesized. Matching of 3' extending ends can occur efficiently as soon as homologous sequences are available (D). Interactions between *P* ends favor matching between sequences close to them. The result then can be a deletion (G). If neo-synthesis goes further, many possibilities of homologous matching exist (E). *P* end interactions would then induce duplications (H) or deletions (G). Branch migration might also yield Holliday junctions (ellipses) during the neo-synthesis step (F). Resolution of these junctions could produce nonrearranged transposons (I).

9.10^{-4} . Thus, large duplications seem to be significantly favored.

Interactions between *P* extremities during the rearrangement process could explain all our results. It can be hypothesized that, because of such interactions, early matching during the repair process could occur preferentially between 5S units close to *P* ends (Figure 7, D and G). This would result in a large number of deletions characterized by a small number of 5S units. However, if extension proceeds on both sides to give nearly complete CAR1 sequences, including *P* extremities (Figure 7E), recombination then could occur between 5S units close to these newly synthesized *P* extremities, leading to a duplication of the *white* gene with an internal large 5S cluster (Figure 7H). Resolution of Holliday junctions also could occur (Figure 7F), leading to nonrearranged CAR1 transposons (Figure 7I). However, our experimental system does not allow to detect such events.

The model shown in Figure 7, which explains the small number of 5S units within deleted CAR1 transposons and the large number of 5S units within the internal cluster of rearranged transposons carrying a *white* duplication, also accounts for the absence of rearrangements within individual 5S arrays of CAR1 and CAR2. If *P* end interactions favor rearrangements between sequences close to them, direct repeats in the vicinity of these ends would probably act synergistically, explaining the structure of our rearranged CAR1 and CAR2 transposons. When direct repeats are not close to both *P* ends, as in the case of the CAR3 transposon and in our previous experiments (PÂQUES and WEGNEZ 1993), *P* site-specific interactions would not be strong enough to overcome the high frequency of intra-array rearrangements.

Preferential rearrangements between sequences close to *P* ends could be a general feature of hybrid dysgenesis. It has been shown that nearly precise excision, leaving <40 bp of *P* ends, could be observed relatively frequently (RIO *et al.* 1986; SEARLES *et al.* 1986; TSUBOTA and SCHEDL 1986). Precise excision has also been observed in the absence of transposon-free template (RIO *et al.* 1986; TSUBOTA and SCHEDL 1986), and could be due to recombination between the 8-bp direct repeats surrounding the transposon.

Molecular requirements for *P* end interactions and single-strand matching: All enzymes involved in the *P* element rearrangement process are not yet known. The transposase and another protein, IRBP (inverted repeat binding protein), have been shown to bind *P* ends (RIO and RUBIN 1988; KAUFMAN *et al.* 1989). One key parameter to consider is the specificity of the DNA-protein interactions. It is known that the transposase recognizes and binds two 20-bp segments in each *P* end. This enzyme also displays a high nonspecific affinity for DNA, so that it can bind much larger sequences *in vitro*, and probably *in vivo* (KAUFMAN *et al.* 1989). After having cut the *P* element, the transposase could remain on flank-

ing sequences. It could also reoccupy available *P* end sequences when strand exchange occurs during conversion. DNA strand invasion and specific binding of protein complexes to *P* end would thus be simultaneous events.

The largest deletion that could be obtained by homologous recombination, considering the sequences at both sides of 5S clusters, would leave a fragment of 216 bp, representing two thirds of one 5S gene (Figure 1). This type of deletion, generating one 275-bp *EcoRI/PstI* fragment, was not observed among the 34 rearranged transposons produced by homologous recombination between 5S genes (Figure 5A). Only one recombination event involving the truncated 5S gene (216 bp) located at the 3' end of one neosynthesized CAR1 transposon and the first complete 5S gene of the other neosynthesized strand occurred to produce the *white* duplication surrounding an internal 5S cluster comprising 13 gene units (Figure 5C). Homologous matching thus involved at most 216 bp. Several rearrangements involved at most 592 bp, *i.e.*, the combined size of the truncated 5S gene and of one complete 5S gene. This is the case for the four duplications containing an internal 5S array of 12 units, and for the 16 deletions yielding a 5S array of 2 units (Figure 5 A and C).

Homologous matching efficiency thus dramatically increases between 216 and 592 bp. KURKULOS *et al.* (1994) observed dysgenesis-induced deletions within a *P* element carrying two 276-bp direct repeats separated by a 6-kb segment. However, the rate of the deletions affecting the internal segment was only 3%. The presence of larger direct repeats in the CAR1 transposon could explain the very high rate of rearrangements we obtained. In yeast single-strand annealing, the MEPS (minimal efficiently processing segment; SHEN and HUANG 1986) was estimated to range between 63 and 89 bp (SUGAWARA and HABER 1992). The site-specific protein interactions that seem to occur in the vicinity of *P* ends could introduce a strong bias in the size requirements for an efficient pairing.

What unites rearrangements occurring within unstable tandem arrays in various organisms? Instability of tandemly repeated sequences is a widespread phenomenon. Tandem arrays internal to *P* elements are rearranged at a high rate when placed in a dysgenic context (PÂQUES and WEGNEZ 1993; KURKULOS *et al.* 1994; this study). The *P* element also can induce deletions and duplications of tandemly repeated sequences when inserted within such arrays (THOMPSON-STEWART *et al.* 1994). Minisatellites, in humans, are also known to be unstable (JEFFREYS *et al.* 1990; ARMOUR and JEFFREYS 1992). Duplications and deletions within a single array, transfer of sequences between alleles and complex recombination events have been observed (JEFFREYS *et al.* 1994). No accompanying crossing over were observed. Interallele transfers, which do not correspond to se-

quence replacements but rather to sequence insertions, could be due to gene conversions. These insertions would occur by realignments during initiation of the conversion process. We propose that matching of the neosynthesized strands would explain deletions, some of the complex events, and the lack of crossing over occurring in this system. Thus, the same molecular process could generate the rearrangements observed within tandemly repeated arrays in *Drosophila* *P* elements as well as in human minisatellites.

We thank A. BUCHETON, M. COBB, C. COLAS DES FRANCS, J. HABER, J. K. MOORE, A. NICOLAS and J.-L. ROSSIGNOL for their critical comments on the manuscript.

LITERATURE CITED

- ARMOUR, J. A. L., and A. J. JEFFREYS, 1992 Biology and applications of human minisatellite loci. *Curr. Opin. Genet. Dev.* **2**: 850–856.
- CARROLL, D., S. H. WRIGHT, R. K. WOLFF, E. GRZESLUK and E. B. MARYON, 1986 Efficient homologous recombination of linear DNA substrates after injection into *Xenopus laevis* oocytes. *Mol. Cell. Biol.* **6**: 2053–2061.
- ENGELS, W. R., 1989 *P* elements in *Drosophila melanogaster*, pp. 437–484 in *Mobile DNA*, edited by D. E. BERG and M. M. HOWE. American Society of Microbiology, Washington, DC.
- ENGELS, W. R., D. M. JOHNSON-SCHLITZ, W. B. EGGLESTON and J. SVED, 1990 High-frequency *P* element loss in *Drosophila* is homolog dependent. *Cell* **62**: 515–525.
- FORMOSA, T., and B. M. ALBERTS, 1986 DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. *Cell* **47**: 793–806.
- GLOOR, G. B., N. A. NASSIF, D. M. JOHNSON-SCHLITZ, C. R. PRESTON and W. R. ENGELS, 1991 Targeted gene replacement in *Drosophila* via *P* element-induced gap repair. *Science* **253**: 1110–1117.
- HASTINGS, P. J., 1988 Recombination in the eukaryotic nucleus. *Bio-Essays* **9**: 61–64.
- HOLLIDAY, R., 1964 A mechanism for gene conversion in fungi. *Genet. Res.* **5**: 282–304.
- JEFFREYS, A. J., R. NEUMANN and V. WILSON, 1990 Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell* **60**: 473–485.
- JEFFREYS, A. J., K. TAMAKI, A. MACLEOD, D. G. MONCKTON, D. L. NEIL *et al.*, 1994 Complex gene conversion events in germline mutation at human minisatellites. *Nature Genet.* **6**: 136–145.
- KAUFMAN, P. D., R. F. DOLL and D. C. RIO, 1989 *Drosophila* *P* element transposase recognizes internal *P* element DNA sequences. *Cell* **59**: 359–371.
- KLAR, A. J. S., and J. N. STRATHERN, 1984 Resolution of recombination intermediates generated during yeast mating type switching. *Nature* **310**: 744–748.
- KLEIN, H. L., 1984 Lack of association between intrachromosomal gene conversion and reciprocal exchange. *Nature* **310**: 748–753.
- KURKULOS, M., J. M. WEINBERG, D. ROY and S. M. MOUNT, 1994 *P* element-mediated *in vivo* deletion analysis of *white-apricot*: deletions between direct repeats are strongly favored. *Genetics* **136**: 1001–1011.
- LIN, F. L., K. SPERLE and N. STERNBERG, 1984 Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Mol. Cell. Biol.* **4**: 1020–1034.
- LIN, F. L. M., K. SPERLE and N. STERNBERG, 1990 Repair of double-stranded DNA breaks by homologous DNA fragments during transfer of DNA into mouse L cells. *Mol. Cell. Biol.* **10**: 113–119.
- MARYON, E., and D. CARROLL, 1991a Involvement of single-stranded tails in homologous recombination of DNA injected into *Xenopus laevis* oocytes nuclei. *Mol. Cell. Biol.* **11**: 3268–3277.
- MARYON, E., and D. CARROLL, 1991b Characterization of recombination intermediates from DNA injected into *Xenopus laevis* oocytes: evidence for a nonconservative mechanism of homologous recombination. *Mol. Cell. Biol.* **11**: 3278–3287.
- MESELSON, M. S., and C. M. RADDING, 1975 A general model for genetic recombination. *Proc. Natl. Acad. Sci. USA* **72**: 358–361.
- NASSIF, N., J. PENNEY, S. PAL, W. R. ENGELS and G. B. GLOOR, 1994 Efficient copying of nonhomologous sequences from ectopic sites via *P*-element-induced gap repair. *Mol. Cell. Biol.* **14**: 1613–1625.
- PÂQUES, F., and M. WEGNEZ, 1993 Deletions and amplifications of tandemly arranged ribosomal 5S genes internal to a *P* element occur at a high rate in a dysgenic context. *Genetics* **135**: 469–476.
- PIRROTTA, V., 1988 Vectors for *P*-mediated transformation in *Drosophila*, pp. 437–456 in *Vectors: A Survey of Molecular Cloning Vectors and Their Uses*, edited by R. L. RODRIGUEZ and D. T. DENHARDT. Butterworths, Boston, MA.
- PLASTERK, R. H. A., 1991 The origin of footprints of the *Tc1* transposon of *Caenorhabditis elegans*. *EMBO J.* **10**: 1919–1925.
- PLASTERK, R. H. A., and J. T. M. GROENEN, 1992 Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following *Tc1* excision. *EMBO J.* **11**: 287–290.
- PRÉAT, T., 1990 High resolution Southern analysis of genomic DNA using heat denatured acrylamide gels. *Nucleic Acids Res.* **18**: 1073.
- RIO, D. C., 1990 Molecular mechanisms regulating *Drosophila* *P* element transposition. *Annu. Rev. Genet.* **24**: 543–578.
- RIO, D. C., and G. M. RUBIN, 1988 Identification and purification of a *Drosophila* protein that binds to the terminal 31-base-pair inverted repeats of the *P* transposable element. *Proc. Natl. Acad. Sci. USA* **85**: 8929–8933.
- RIO, D. C., F. A. LASKI and G. M. RUBIN, 1986 Identification and immunochemical analysis of biologically active *Drosophila* *P* element transposase. *Cell* **44**: 21–32.
- ROBERTSON, H. M., C. R. PRESTON, R. W. PHILLIS, D. M. JOHNSON-SCHLITZ, W. K. BENZ *et al.*, 1988 A stable genomic source of *P* element transposase in *Drosophila melanogaster*. *Genetics* **118**: 461–470.
- RUBIN, G. M., and A. C. SPRADLING, 1982 Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**: 348–353.
- RUBIN, G. M., and A. C. SPRADLING, 1983 Vectors for *P* element-mediated gene transfer in *Drosophila*. *Nucleic Acids Res.* **11**: 6341–6351.
- SAMSON, M. L., and M. WEGNEZ, 1988 Bipartite structure of the 5S ribosomal gene family in a *Drosophila melanogaster* strain, and its evolutionary implications. *Genetics* **118**: 685–691.
- SEARLES, L. L., A. L. GREENLEAF, W. E. KEMP and R. A. VOELKER, 1986 Sites of *P* element insertion and structures of *P* element deletions in the 5' region of *Drosophila melanogaster* *Rp11215*. *Mol. Cell. Biol.* **6**: 3312–3319.
- SHEN, P., and H. V. HUANG, 1986 Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* **112**: 441–457.
- SUGAWARA, N., and J. E. HABER, 1992 Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. *Mol. Cell. Biol.* **12**: 563–575.
- SZOSTAK, J. W., T. L. ORR-WEAVER, R. J. ROTHSTEIN and F. W. STAHL, 1983 The double-strand-break repair model for recombination. *Cell* **33**: 25–35.
- THOMPSON-STEWART, D., G. H. KARPEN and A. C. SPRADLING, 1994 A transposable element can drive the concerted evolution of tandemly repetitive DNA. *Proc. Natl. Acad. Sci. USA* **91**: 9042–9046.
- TSUBOTA, S., and P. SCHEDL, 1986 Hybrid dysgenesis-induced revertants of insertions at the 5' end of the *rudimentary* gene in *Drosophila melanogaster*: transposon-induced control mutations. *Genetics* **114**: 165–182.
- WHEELER, C. J., D. MALONEY, S. FOGEL and R. S. GOODENOW, 1990 Microconversion between murine *H-2* genes integrated into yeast. *Nature* **347**: 192–194.