Supragenic loop organization: mapping in *Drosophila* embryos, of scaffold-associated regions on a 800 kilobase DNA continuum cloned from the 14B-15B first chromosome region

Patrick Surdej, Claude Got, Roland Rosset and Raymond Miassod*
Laboratoire de Génétique et Biologie Cellulaires, Centre National de la Recherche Scientifique, Case 907, 13288 Marseille, Cedex 9, France

Received May 2, 1990; Revised and Accepted May 25, 1990

ABSTRACT

The supragenic loop organization of the *Drosophila* genome was investigated on a 800 kilobase (kb) DNA continuum from the 14B-15B first chromosome region. Nuclear scaffolds from 0–18 hr embryos were prepared with Laemmli's low-salt, detergent procedure and digested with restriction enzymes. Scaffold-associated regions (SARs) were mapped by probing Southern transfers of total, scaffold-associated and free DNA with a set of 70 recombinant phages overlapping the investigated genomic region. In all, 85 restriction fragments showed association to scaffolds. 12 of them were present in the majority of scaffolds. They bore strong SARs organizing the DNA molecule as consecutive loops with sizes ranging from 15 to 115 kb. 44 were present in only a fraction of scaffolds. They contained weak SARs subdividing the basic loops into smaller ones. 29 additional restriction fragments were present in a very small fraction of scaffolds. The position of SARs with respect to transcribed regions was investigated. Strong SARs appeared to be located on untranscribed DNA and to frame transcription units. In contrast, at least some weak SARs were shown to comap with transcribed regions or to reside within characterized transcription units. Statistical analyses established that strong and weak SARs were periodically positioned on the DNA continuum and that there was a potential contact point between scaffolds and the DNA continuum every 11 kb, or multiples thereof. Implications for SAR role(s) are discussed.

INTRODUCTION

An increasing amount of data suggests that the structural organization of chromosomes plays a role in the dynamics of the eucaryotic genome. The first line of evidence comes from studies on replication. For example, efficient *in vitro* replication has been shown to occur in sequentially-reconstituted nucleus-like structures only after some steps of the assembling process were completed (1). Furthermore, in *Drosophila* transformed with chorion genes, a genome region prone to a stage- and tissue-specific amplification process, transgene amplification was obtained at the expected developmental stage and tissue, but at a level significantly different from that of the resident chorion genes (2). Finally, a direct relationship between replicon size and loop size has been established for a number of eucaryotes (3). Moreover, for *Xenopus* rDNA (4), the hamster dihydrofolate reductase (*dhfr*) region (5,6) and parts of yeast chromosomes (7) loop limits proved to coincide with potential replication origins or replication termination sites. These data indicate that features of genome structure and replication are related.

A second line of evidence comes from comparative studies on genome structure and genome transcription. For instance, the boundaries of the nuclease sensitive domains of chicken lysozyme (8), human β-globin (9) and human interferon (10) genes have been demonstrated to comap with the limits of the loop domains encompassing these genes. Furthermore, the position effect, i.e. the loss of part of the transcriptional control of the activity of genes randomly integrated in eucaryotic genomes, was minimized, if not cancelled, when transgene constructions included enough of the normal 5′- and 3′- surrounding sequences (11–14) or were *in vitro* engineered so as to frame coding sequences with DNA from scaffold-associated regions (SARs, 15–16). This suggests that SARs may bear positional information required for control of transcription.

Investigations on SARs further substantiated this idea. In the *Drosophila* genome, SARs have been extensively studied by Laemmli and coworkers for the histone (*His*) gene cluster, one of the heat-shock protein (*Hsp*) clusters, and the alcohol dehydrogenase (*Adh*), fushi-tarazu (*Ftz*) and *Sgs-4* genes (17–21). These authors have also reported an analysis of SARs on the 320 kb DNA cloned in the '*rosy-Ace*' region from the *Drosophila* second chromosome (22). For yeast, several SARs have been investigated (7). For vertebrates, SARs have been studied in the regions of *Xenopus* rDNA (4), chicken lysozyme (8), mouse Kappa (κ) and heavy chain immunoglubulin (23,24), human hypoxanthin-phosphoribosyl transferase (25), β-interferon

* To whom correspondence should be addressed

(10), *dhfr* (5,26) and globin (9) genes. These studies led to the repeated observation that SARs are not randomly distributed with respect to genetic unit transcribed parts, tissue-specific control elements, enhancers and topoisomerase II cleavage sites (27). Moreover, interspecific conservation of SARs has been documented (7,8,10,20, 23,28).

For the obvious reason of cloned probe availability, a number of these studies concerned only transcribed genome regions (with surrounding sequences), sometimes highly specialized and not exceeding 15−50 kb in length, except for the globin gene clusters (9), a *dhfr* amplicon (5) and the '*rosy-Ace*' region (22). Therefore, the possibility that proposals on SARs are currently biased towards specialized transcribed genome regions and that information on SARs at a supragenic level is not yet available had to be considered. Consequently, it was of interest to investigate SARs on a long cloned DNA located in an apparently non-specialized genome region.

We have analyzed the loop organization of a 800 kb DNA continuum that we recently cloned from the *Drosophila* first chromosome 14B−15B region (29). Our data, in part confirmed by studies from other laboratories, demonstrated the presence, on this DNA stretch, of numerous transcription units, with specific expression patterns. A number of them are involved in various cellular functions or control processes. The present paper reports the mapping of SARs on this DNA continuum and an analysis of their positioning, both along the DNA continuum and with regard to transcribed regions or to characterized transcription units. Data are discussed in terms of the possible various roles played by SARs.

MATERIALS AND METHODS

SAR mapping

SARs were mapped as follows. 0−18 hr aged *Drosophila* embryos were collected from mass population cages (Oregon R strain). Nuclei were purified as described (30). Nuclear scaffolds were prepared with Laemmli's low-salt, 3′,5′-diiodosalicylate method (LIS procedure, 17) and digested with Eco RI+Hind III or Eco RI+Hind III+Bam HI. Scaffold bound DNA (P-DNA) and unbound DNA (S-DNA) were separated, purified, and their amounts estimated by UV measurements. They accounted for 25% and 75% of total DNA (T-DNA), respectively. Defined amounts of T-, P- and S-DNA were then electrophoresed on 0.8% and 2% agarose gels, transferred to Amersham nylon membranes and probed with recombinant DNA labeled by the random primer method, according to standard procedures (31, 32).

Because of potential P- and S-DNA cross-contamination filters were probed with a *Drosophila* genomic DNA fragment for which SARs had been mapped. The *Adh* gene was selected. DNA from plasmid *pAdh* hybridizes with a 1.65 kb Eco RI-Hind III fragment and a 2.6 kb Hind III-Eco RI fragment of genomic DNA (33). From previous data (18) the 1.65 fragment (5′-part of the *Adh* gene, plus upstream sequences including a SAR) and the 2.6 fragment (3′-part of the *Adh* gene, plus downstream sequences not interacting with scaffold) were expected to be restricted to P- and S-DNA, respectively. This was indeed observed for all nuclear scaffold preparations we prepared (illustrated in Fig. 2). In one occasion, a preparation during which the rinse steps were erroneously increased showed contamination of S-DNA with P-DNA. This preparation was not used.

Mapping of transcribed regions

Transcribed regions were delimitated by probing Southern transfers of DNA clones, digested with combinations of restriction enzymes, with labeled, random hexanucleotide-primed, single strand DNA copies of poly(A)$^+$ RNAs from 0−18 hr aged embryos, as shown in more detail elsewhere (29).

Statistical analyses

The DNA continuum was subdivided into 852 consecutive units 0.94 kb long (the average size of restriction fragments generated by the 7 restriction enzymes for which maps had been drawn). Numeric distributions characterizing the various parameters under analysis (expressed regions, scaffold-associated fragments, fragments bearing strong SARs and fragments bearing weak SARs) were generated. For this, the status of every unit was noted as 1 (presence of the parameter) or 0 (absence). Two statistical analyses were performed. First, the numeric distribution for expressed parts was compared to that for strong SARs and to that for weak SARs. This gave an indication of whether the parameters were associated or mutually exclusive. Second, the numeric distribution specifying the position of all scaffold-associated fragments (or of weak or strong SARs) was compared to theoretical periodical ones. This comparison gave informations on the possible regular positioning of these parameters on the DNA continuum. Theoretical distributions were constructed as follows. For instance, a theoretical distribution pattern of a 12 unit period was obtained by giving value 1 to unit rank 1, value 0 to the next 11 units, then value 1 to unit rank 13, and so on. Since starting points for both distributions were arbitrary, comparisons were done repeatedly after sliding the theoretical distribution along the experimental one, by one unit increment, up to the period value. Also, regions 0−235 and 235−805 of the DNA continuum (see Fig. 6) were separately investigated, since they are separated by a small gap. A computer was used to generate the theoretical distributions and to perform the comparisons. The statistical criteria used were the classic χ^2 test, and a second test developed in our laboratory (34).

RESULTS

SAR mapping

Nuclei were purified from 0−18 hr *Drosophila* embryos. Nuclear scaffolds were prepared and digested with Eco RI+Hind III, or Eco RI+Hind III+Bam HI, or Hae III or DNase I. Fig 1 shows the distribution, on agarose gels, of restriction fragments present in P- and S- DNA. They ranged from less than 0.2 to more than 11 kb, and, as expected, the size range depended on the type of digestion. Whatever the enzyme, fragments were larger in P-DNA. The size difference was particularly large for digestion with Hae III. Discrete bands were visible in P-DNA digested with restriction enzymes. The 4.5 band generated by digestion with Eco RI+Hind III or Eco RI+Hind III+Bam HI (see white arrows in Fig.1) was cloned and demonstrated to be ribosomal DNA (C. Brun, P. Surdej and R. Miassod, unpublished results).

Southern transfers of P- and S-DNA that had been digested with Eco RI+Hind III, or Eco RI+Hind III+Bam HI, and fractionated on agarose gels (equal loads) were hybridized successively with DNA from a set of recombinant phages and hybridization signals were compared. Alternatively, gels were loaded with equal amounts of T- and P-DNA and with a 3-fold amount of S-DNA, to take into account the final P- and S-DNA

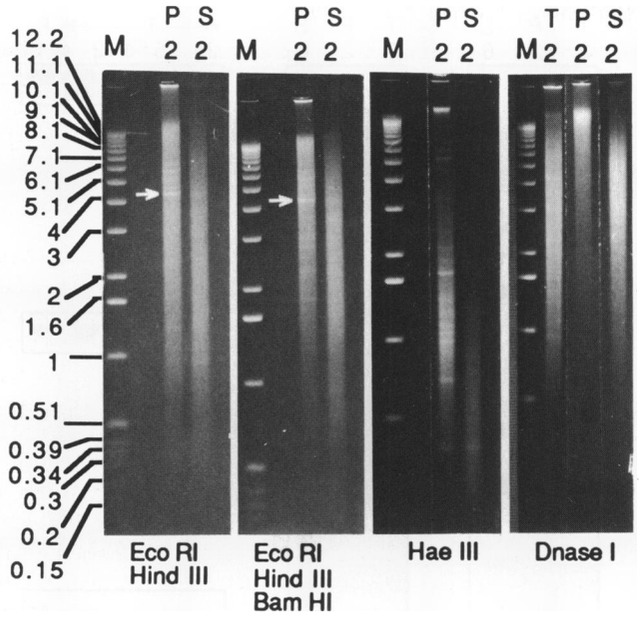

Fig. 1: Fractionation, on 0.8% agarose gels, of scaffold-bound and free DNA from *Drosophila* nuclear scaffolds digested with various enzymes. Scaffolds were prepared, digested with enzymes and fractionated on gels as reported in Materials and Methods. Enzymes used for digestion of scaffolds are indicated under the lanes. Digestion with DNase I (100 units) was allowed for 15 s, at 25°C, and stopped by adding EDTA. Letters and numbers above the lanes indicate the nature and the amount (μg) of the analyzed fractions. T, P and S stand for T-, P- and S-DNA, respectively. M corresponds to size markers. Molecular weights are indicated on the left.

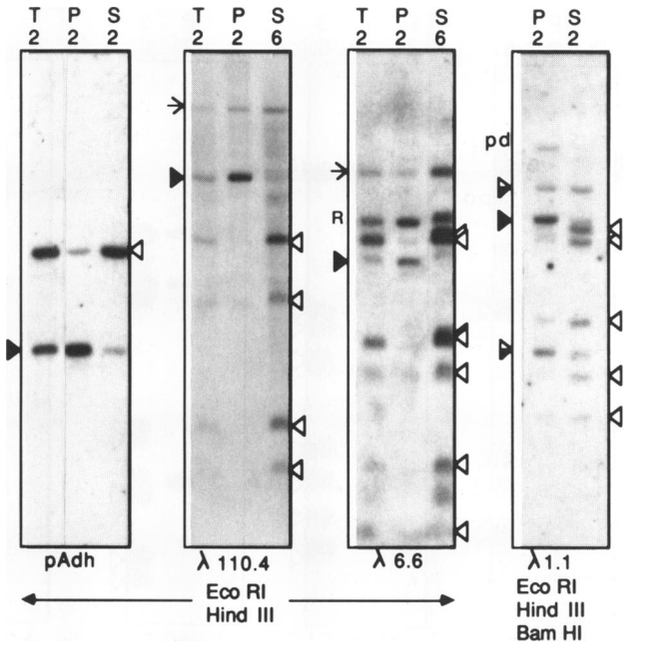

Fig. 2: Strong SARs present on the DNA continuum: autoradiograms of Southern transfers of T-, P- and S-DNA hybridized with recombinant phage DNA and showing regions of the DNA continuum associated to 100% scaffolds. Scaffolds were prepared, digested with enzymes and fractionated on gels. Southern transfers were hybridized with cloned DNA, as reported in Materials and Methods. Letters and numbers above the lanes indicate the nature and the amount (μg) of the analyzed fractions. T, P and S stand for T-, P- and S-DNA, respectively. Probes (maps drawn in Fig.6) and restriction enzymes used to digest scaffolds are specified under the lanes. pd indicates partial digestion. R refers to an extra-band corresponding to a genomic fragment located outside the DNA continuum. Filled triangles point to fragments restricted to P-DNA (strong SARs). Half-filled triangles show fragments enriched in P-DNA, but distributed in less than 50% P-DNA (weak SARs). Arrows refer to fragments present in P-DNA, but not enriched, and considered as not bearing SARs. Empty triangles show fragments restricted to S-DNA. Exposure was for 2−6 days, with an intensifying screen.

yields. In this way, comparison of signal heights between T- and P-DNA, and between P- and S-DNA, allowed, for each restriction fragment, an estimation of its enrichment in P-DNA and of its partitioning between P- and S-DNA, respectively. Data are presented in Figs 2−5. For 3 restriction fragments hybridization signals were restricted to P-DNA. This is shown in Fig. 2 (see filled triangles). 9 additional fragments were highly enriched in P-DNA and partitioned mostly in this DNA fraction. Illustrative data are presented in Fig. 3A (see filled triangles). Hybridization of Southern transfers bearing increased amounts of T-DNA and a fixed amount of P-DNA (illustrated in Fig. 3B) allowed an estimation of the partitioning of these fragments between P- and S-DNA. This was of the order of 75% in P-DNA. SARs born by these 12 SARs were operationally designated as strong SARs.

44 restriction fragments were enriched in P-DNA but their partitioning in P-DNA did not exceed 50%. This is documented in Fig. 4 and also in Figs 2−3 (see half-filled triangles). SARs carried by these 44 fragments were operationally designated as weak SARs. 29 additional fragments showed significant signals in P-DNA, but without enrichment (Fig. 5 and arrows in Figs 2−4). They were not considered as bearing SARs. All other restriction fragments were restricted to S-DNA (see empty triangles in Figs 2−4). Repeats of the experiment with 3 distinct scaffold preparations gave reproducible results. Fragments free of interaction with the scaffold never gave signal on P-DNA lanes and fragments enriched in P-DNA were always the same, although small variations of partitioning were observed, from one scaffold preparation to another. Average sizes and size ranges for these various fragments are reported in Table I. Size ranges were similar, but average sizes had a tendency to increase with the strength of binding.

The positioning of the various restriction fragments on the DNA continuum physical map is indicated in Fig. 6. Fragments belonging to both groups were scattered on the DNA continuum.

This SAR mapping was probably only an estimation. On the one hand, it is conceivable that a SAR-containing fragment might in fact bear more than one SAR, since the size of restriction fragments issued from digestion of scaffolds (see Table I) was above that usually reported for SARs (a few hundred bases). For example, in region 210−225, digestion of scaffolds with Eco RI+Hind III suggested the presence of one SAR in a 15 kb long fragment, whereas a more complete digestion with Eco RI+Hind III+Bam HI showed the presence of 2 distinct SARs. Also, the restriction enzymes used to digest scaffolds might have splitted a SAR between 2 contiguous restriction fragments. If this breakage destroyed attachment, then cleavage might have led to an under-estimation. Alternatively, if attachment was conserved on both sides of the cleavage site, an over-estimation of the SAR number might have resulted. This could apply to several regions where distinct SAR-containing fragments are contiguous or separated by free fragments less than 1 kb long.

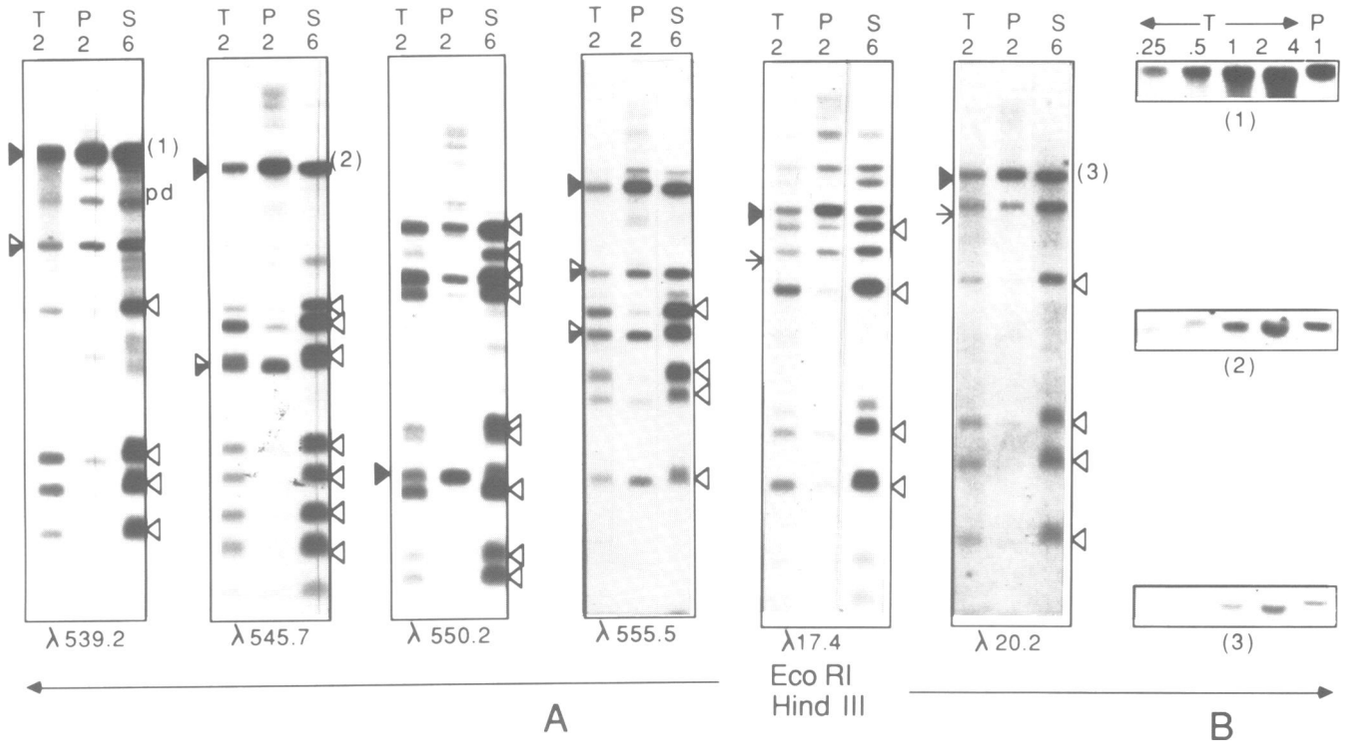

Fig. 3: Strong SARs present on the DNA continuum: autoradiograms of Southern transfers of T-, P- and S-DNA hybridized with recombinant phage DNA and showing regions of the DNA continuum enriched in P-DNA and associated to more than 50% scaffolds. A: as in Fig. 2, except that filled triangles point to fragments for which hybridization signals were present in more than 50% scaffolds (strong SARs). Fragments numbered (1) to (3) are those submitted to the semi-quantitative analysis reported in part B. B: Southern transfers bearing various amounts of T-DNA and a fixed amount of P-DNA were hybridized with fragments of cloned DNA overlapping the genomic fragments numbered (1) to (3) in part A. Only parts of autoradiograms of interest are shown. Scaffold preparations used in parts A and B were distinct. Exposure was for 2–8 days, with an intensifying screen.

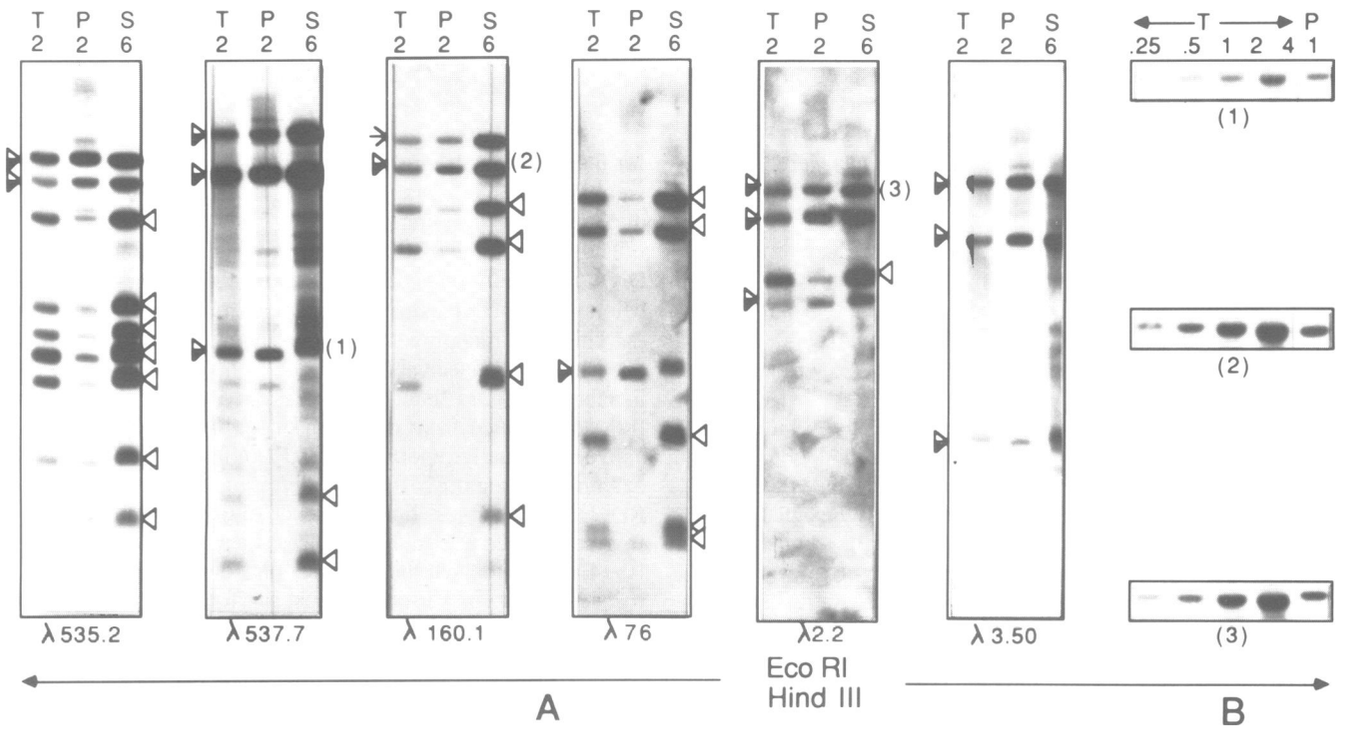

Fig.4: Weak SARs present on the DNA continuum: autoradiograms of Southern transfers of T-, P- and S-DNA hybridized with recombinant phage DNA and showing regions of the DNA continuum enriched in P-DNA and associated to less than 50% scaffolds. As in Fig. 3.

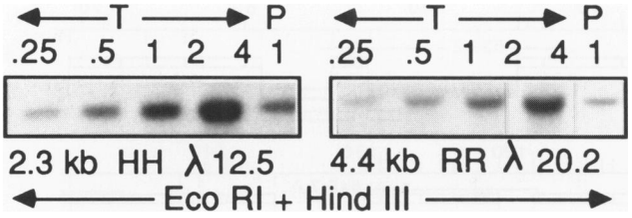

Fig. 5: Autoradiograms of Southern transfers of T- and P-DNA hybridized with recombinant phage DNA and showing regions of the DNA continuum present in P-DNA, but with no enrichment. As in Fig. 2B. Fragments used as probes are specified under the Fig. (R: Eco RI; H: Hind III).

Table I. Average sizes and size ranges (in kb) of fragments interacting or not with scaffolds.

	Free DNA fragments	Scaffold-associated DNA fragments		
		no SARs	weak SARs	strong SARs
Average sizes	1.6±1.3	3.0±1.7	4.0±2.2	5.5±1.7
Size ranges	0.05−6.55	0.65−7.1	1.0−8.1	1.2−9.0

Table II. Comparative examination of the distributions of SARs and of transcribed regions.

Compared distributions	Probability (if the distributions were at random), for the coincidence to be		
	lower	equal	higher
Transcribed parts + strong SARs	0.9216	0.0308	0.0476
Transcribed parts + weak SARs	0.2050	0.0509	0.7441

Comparative examination of the distribution of SARs and transcribed regions

We previously detailed a procedure for identifying the transcribed parts of long stretches of cloned DNA (29). We applied it to 0−18 hr embryos (data not shown). Fig. 6 shows the limits of transcribed parts to the nearest restriction site for the 7 enzymes for which maps had been drawn. The distributions of transcribed parts and of SARs were compared. We observed that strong and weak SARs were present both in densely transcribed and untranscribed regions. Also, no obvious correlation was noticed between the loop size and the level of expression. Lastly, 113 out of the 158 coding sequence-containing restriction fragments did not harbor SARs. Therefore, attachment to scaffolds and transcription are not systematically associated.

Limits of transcribed regions and of SARs were too imprecise to tell whether SARs are transcribed sequences, or not. However, a statistical analysis of the data (Table II) established that the probability for strong SARs not to be transcribed was high, above 0.92.

Positioning of SARs with regard to transcription unit limits

Because of the large size of the DNA continuum investigated and of the relative paucity of genetic data, the limits of transcription units have not yet been established, with the exception of the *para*, *r* and *PS 2* transcription units whose exon-intron structures have been established (35−37). However, as detailed elsewhere (29) transcription units were tentatively identified by their 3′ parts. This is schematized in Fig. 7.

With regard to strong SARs (Fig. 7A), most of the time (10 out of 12) several transcription units were located in between 2 consecutive SARs. This applied in particular to *para*, *r*, and *PS 2* and strongly suggested that strong SARs delimit loops harboring several transcription units. With regard to weak SARs (Figs 6 and 7B), sequences coding for *r* (at 415−428 on the map of Fig. 6) were restricted to S-DNA (illustrated by probe 76 in Fig. 4). However, both the *para* and *PS 2* genes (respectively at 237−282 and 510−542, on the map of Fig. 6) harbored internal weak SARs. Restriction fragments harboring weak SARs overlapped exons plus introns (see probe 160.1 for *para*, and probes 3.50 and 4.5 for *PS 2*, in Fig. 4). One of these fragments overlapped only the seventh intron of the *PS 2* gene.

Investigation of the regularity of the distribution of scaffold-associated fragments

The possibility of a regular positioning of SARs on the DNA continuum was investigated by comparing their distribution with theoretical periodical distributions of various period values, using 2 statistical criteria (detailed in Materials and Methods). In a first step, the comparison was carried out with the distribution of all scaffold-associated fragments, including those not considered as bearing SARs, for the following reasons. First, by their ability to bind to scaffolds, all these fragments made a group distinct from that of fragments not-interacting with scaffolds. Second, all these fragments shared in commun 2 properties: repeated elements present on the DNA continuum were shown to be localized in their vicinity (P. Surdej and R. Miassod, manuscript in preparation) and a subset of them comapped with sequences replicating autonomously in yeast (C. Brun, Qi Dang and R. Miassod, manuscript submitted). This statistical analysis indeed established that they were regularly positioned on the DNA molecule. Table III lists the period values for which a fitting between the data and theoretical distributions was observed. Both regions 0−235 and 235−805, when separately examined, showed a fitting for a single period value (or multiples thereof). These single values were very close to each other: 13 or 12 units, i.e. 11kb (or multiples thereof), respectively for regions 0−235 and 235−805. Table III also reports the percentage of theoretical positions fitted by the data. 36 out of the 66 positions of the 11 kb theoretical distribution were fitted with the data. For period values above 36 kb the majory of positions was fitted.

The same analysis was repeated for the distribution of strong SARs and that of weak SARs. Results are reported in Table III. For strong SARs, both in regions 0−235 and 235−805, the smallest period values for which a fitting was observed were 50−57 and 40 units (i.e. 47−54 and 38 kb), respectively for regions 0−235 and 235−805. A similar result was obtained for weak SARs in both regions: the smallest reached values were close to those observed for strong SARs. They were 63 and 36 units (i.e. 59 and 34 kb), respectively for regions 0−235 and 235−805. All these values can be considered as rough multiples of the 12−13 unit (i.e. 11 kb) period values characterizing the distribution of scaffold-attached regions. The percentage of theoretical positions fitted by experimental distributions is also reported in Table III: all theoretical positions were fitted with a SAR in region 0−235, and part of them in region 235−805.

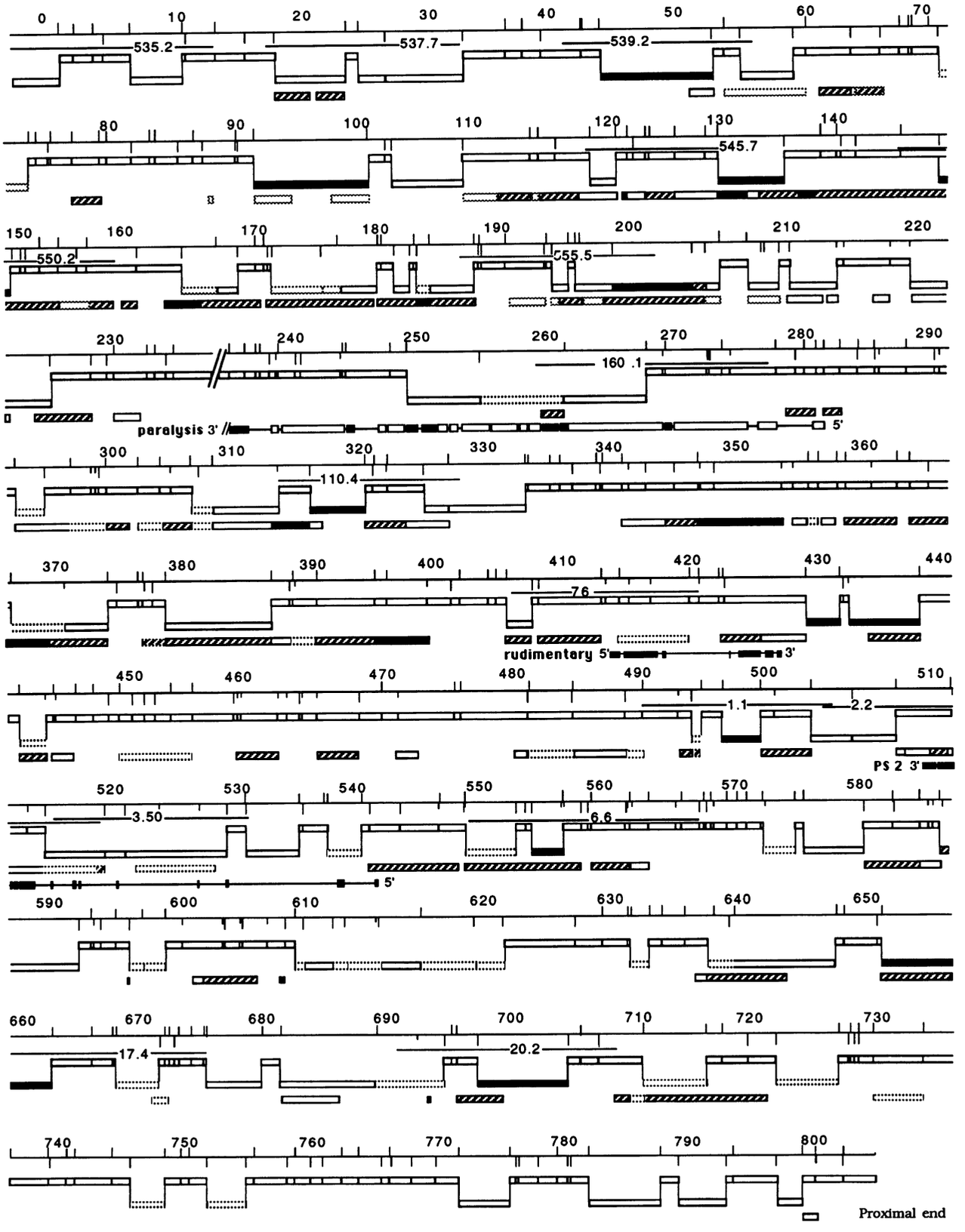

Fig. 6 : Mapping of SARs on the 800 kb DNA continuum physical map.The top line is the physical map with a scale, in kb. Long vertical bars above the line indicate Eco RI sites and long and short vertical bars below the line Hind III and Bam HI (only those investigated in this study) sites. Limits of recombinant phages (data of Figs 2–5) are shown on the next line. The next 2 lines report the limits of the free and scaffold-associated fragments. On the upper line are filled rectangles representing restriction fragments not interacting with scaffolds. On the lower line are filled and empty rectangles representing fragments bearing strong and weak SARs, respectively; rectangles bordered with a broken line are fragments shown to be associated to no more than 25% scaffolds and not considered as bearing SARs. The last line delimits transcribed regions (see Materials and Methods): filled, cross-hatched and empty rectangles are for regions transcribed at a high, moderate and low level, respectively. Broken rectangle borders indicate uncertain limits. The exon-intron structure of the *para*, *r* and *PS 2* genes are shown (redrawn from references 35 to 37, respectively).

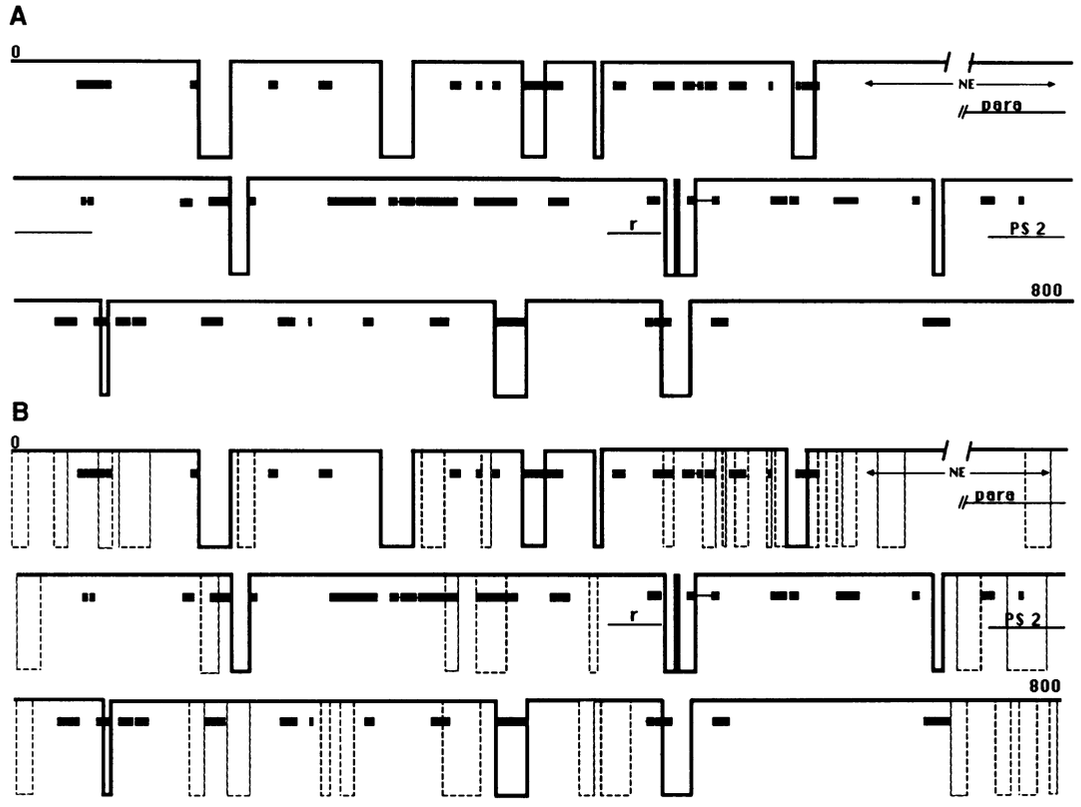

Fig. 7 : Schematic view of the looping of the 800 kb DNA continuum. A: the DNA molecule is schematized as a broken line whose upper and lower parts correspond to free regions and to regions bearing strong SARs, respectively. Individual and linked filled rectangles are 3′ parts of transcription units (29). N.E.: not examined. The *para*, *r* and *PS 2* transcription units are indicated. B: as in A, but the distribution of weak SARs (broken lines), is overimposed.

Table III. Period values of theoretical periodical distributions fitting the distributions of strong SARs, or weak SARs or all scaffold-associated regions.

Experimental parameter examined	Period values giving a fitting in			
	region 0-235		region 235-805	
All scaffold-associated regions	13	(68%)	12	(52%)
	26-27	(90%)	24	(12%)
	39-40	(100)	37	(75%)
			57-64	(80%)
			71-74	(89%)
Strong SARs	50-57	(100%)	40	(27%)
	99-100	(100%)	70-72	(38%)
Weak SARs	63	(100%)	36	(50%)
			79	(71%)

For strong and weak SARs the period value was varied from 3 to 100 units and values at a 0.999 confidence level were recorded. For all scaffold-associated fragments the period value was first varied from 3 to 15 and values at a 0.95 confidence level were recorded; then the period value was varied from 15 to 100 and values at a 0.99 confidence level were scored. Period values are expressed as units (1 unit=0.94 kb). Values in parenthesis are the percentage of theoretical positions fitted by the data.

DISCUSSION

The present study intended to look at the supragenic loop organization of a long DNA continuum from the *Drosophila* genome by mapping regions associated to nuclear scaffolds. We chose the LIS procedure and selected the approach consisting in probing Southern transfers of P- and S-DNA from 0-18 hr embryos for the following reasons. Results from previous comparative analyses of the *Drosophila His-*, *Hsp-*, *Adh-* and *Ftz*-SARs (21) and of the *Xenopus* rDNA-SAR (4) with either the LIS or the high salt procedures were in a good agreement. However, there were discrepancies for another report on the *His*-SAR (17) and for the mouse κ-SAR (23): restriction fragments adjacent to these SARs, and bearing coding sequence partitioned randomly between P- and S-DNA with the high salt procedure, whereas they were restricted to S-DNA with the LIS procedure. Authors have suggested that this random partitioning reflected attachment of transcribed regions via transcription complexes (17, 23). The alternative to the approach used in this study was the *in vitro* approach which consists in adding end-labelled cloned DNA fragments to scaffold preparations and looking for attachment. Both gave consistent results for all *Drosophila* SARs (17,18,21), the mouse κ-SAR (23), the human interferon-SAR (10) and the 3 SARs located next to the β-globin gene (9). But differences were noted for the Ψ-globin gene-SAR (9) and for the SAR positioned 3′ from the *dhfr* gene (5), or were visible on data reported for the *Drosophila Sgs*-4-SAR (17) and for the SAR located at the level of yeast ARS1 (7). We investigated *Drosophila* embryos, rather than an established cell line, since the *Drosophila* SARs located 3′ from the *Ftz* gene and 2 of the 4 SARs framing the *Adh* gene did not interact with nuclear scaffolds from an established cell line (18).

The LIS procedure involved a stabilization step of the nuclear scaffold at 37°C. This raised the question of whether part of the observed scaffold-DNA interactions could be an artefact related

to this step. However, the fact that this approach and the alternative *in vitro* one (in which the stabilization step precedes the addition of end-labelled DNA fragments) gave consistent results in a number of cases demonstrates that the observed association is not an artefactual trapping of DNA caused by the heat step. For the investigated DNA continuum, we have used the alternative *in vitro* to test association of parts of this DNA to yeast scaffolds (C. Brun, Qi Dang and R. Miassod, manuscript submitted): here also consistent results were obtained. Also, most SARs analyzed in detail have been shown to be present in A+T-rich DNA regions. The question of whether the LIS procedure allows identification of true SARs, or simply of mere interactions between scaffolds and A+T-rich DNA stretches has therefore been raised. Although the present study does not report sequence data on SARs, 2 lines of evidence suggested that the identified SARs might have particular base compositions. One was the reported observation of a size difference between P- and S-DNA, both at the level of the whole genome and of the DNA continuum. This difference might be due to the fact that digestion of the DNA molecule is hampered in the vicinity of SARs, or to the fact that restriction sites are unequally distributed on P- and S-DNA, because of peculiarities of the SAR base composition. The observation that every Eco RI and Hind III site present on the DNA continuum was accessible to digestion, as well as every investigated Bam HI site, favors the second explanation. In particular, the large size difference in the case of digestion with Hae III and the persistence of this difference in digestions with DNase I suggested a low level of C+G, and of pyrimidines, respectively. The other evidence was from a preliminary mapping of sites for restriction enzymes cleaving DNA within C+G-rich clusters (C. Brun and R. Miassod, work in progress): of 37 cleavage sites presently identified only 7 were located on SAR-containing fragments. Indeed, data from the literature do not support the hypothesis that occurrence of A+T-richness is a necessary and sufficient condition for defining SARs. This conclusion was drawn from analyses reporting that regions more than 60−70% A+T-rich do not bind to yeast (7), mouse (24), human (9) scaffolds and from an analysis (24) of the mouse heavy chain and κ gene SARs for the A-box and T-box consensus sequences defined by Gasser and Laemmli (18).

The present analysis extends that performed by Laemmli and collaborators on the '*rosy-Ace*' region and brings new informations. The investigated 800 kb DNA continuum appears to be representative of the *Drosophila* genome, euchromatin part, from several points of view. The density of transcription units, with distinct developmental patterns and roles, were those expected for a genome part devoid of obvious specialization (29). Also, the cumulated lengths of restriction fragments present respectively in S- and P-DNA, after weighting for partitioning, were exactly in the 75/25 ratio characterizing the whole genome. Lastly, as said above, the size difference between S- and P-DNA noted for the whole genome was also observed for the DNA continuum. Since the present data and those obtained in the '*rosy-Ace*' region altogether concerned 1% of *Drosophila* euchromatin, it can be suggested that the reported observations might apply to the whole *Drosophila* euchromatin.

The 800 kb DNA molecule appeared to be made of alternate scaffold-associated and free regions. Therefore, the first conclusion of our analysis is that the loop model, based on observation of whole nuclear structures, and then of specific genes or specialized groups of genes, applies at a supragenic level.

12 out of the 85 scaffold-associated restriction fragments were highly enriched in P-DNA and partitioned in P-DNA above 75%, or were restricted to this fraction. 44 additional fragments were also enriched in P-DNA, but were present in P-DNA under 50%. These restriction fragments were considered as bearing SARs. They were operationally defined as strong and weak SARs. Inspection of previous analyses shows that a low partitioning of SARs in P-DNA is not a peculiarity of the present data. Although all SARs previously analyzed in *Drosophila*, as well as the *Xenopus* rDNA-SAR (4) and the SAR localized at the level of yeast ARS1 (7), gave hybridization signals restricted to P-DNA (and can therefore be considered as strong SARs) all other SARs so far investigated in vertebrate systems were shown to partition between P-and S-DNA, sometimes at a very low value. This partitioning has led some authors to disregard regions of the human β-globin cluster or the whole α-globin gene cluster for which signals in P-DNA were less than 50% (9). On the other hand, others have investigated in detail regions of human *dhfr* amplicons for which partitioning was estimated to be at most 10−15% (5) or 20−40% (26). In the present study, every scaffold-associated restriction fragment enriched in P-DNA and partitioning in P-DNA at a ratio less than 50% was considered as bearing a weak SAR, but scaffold-associated fragments not enriched in P-DNA were considered as not bearing SARs. This was somehow arbitrary since other analyses (see further in the text) tended to establish that the various scaffold-associated fragment subclasses might be of the same nature. A second conclusion of this work is therefore that, together with strong SARs, weak SARs are present in *Drosophila* DNA, as is the case for DNA from vertebrates.

A first obvious explanation of partitioning is cross-contamination between P- and S-DNA. This cannot be retained for the present study or for the report on the '*rosy-Ace*' region, since in both cases restriction of some of the hybridization signals to either P- or S-DNA was observed. But this explanation cannot be excluded for the other analyses where no hybridization signals restricted to P-DNA were reported. A second possible explanation is that strong and weak SARs might be of a different nature. Strong SARs would bind scaffolds from every nucleus whereas weak SARs would bind scaffolds from only a subset of nuclei, depending on the status of the cell from which the nucleus originated, with regard to differenciation, or to the cell cycle, or to transcription. The observed partitioning would thus simply be due to the fact that one is looking at a composite nuclei population. Observations from the literature and from the present analysis are consistent with this hypothesis. The SAR located within the fourth intron of the human *dhfr* gene (26) does not interact with scaffolds at metaphase. Fluctuation of partitioning values for the chicken lysosyme- (8), the mouse heavy- and κ- (23,24) and the human globin gene cluster- (9) SARs, according to the tissue or the cell lines from which scaffolds were derived, are visible in the reported data. The same applies to the *Drosophila Sgs*-4-SAR and to the SAR located 3′ to the *Ftz* gene (17,18). In the present analysis, we noticed that small fluctuations of partitioning values from one scaffold preparation to another were never systematically in favor of P- or S-DNA, but depended on the SAR. This could be accounted for by changes of ratios of the various nuclei subpopulations, due to changes in the proportion of the variously-aged embryos. We also observed that distinct genomic fragments shown up by the same cloned fragments (see for instance probe 6.6 in Fig. 2) were characterized by quite different partitioning values, i.e. that a given nucleotide sequence at 2 different genomic localizations

might display various interaction with scaffolds. A third possible explanation is that strong and weak SARs might be of the same nature but, because of variable affinity, every SAR, although interacting with a (or several) proteic partner(s) in every nucleus, would be more or less prone to detach from scaffolds during the isolation procedure. Previous *in vitro* binding experiments (8,21,24), performed by Laemmli and collaborators on *Drosophila* SARs, and showing large differences between affinities of SARs for scaffolds support this hypothesis. The fact that both strong and weak SARs appeared to be regularly disposed on the DNA molecule, with similar periodicities, also speaks in favor of a commun nature.

A third conclusion of this study relates to the positioning of SARs with regard to transcription units. At variance to the proposal (27), based on available data, that SARs in *Drosophila* are found in non-transcribed regions, our analysis suggests that some weak SARs may be located within transcription units, and more precisely within introns, as is the case for vertebrates. From this point of view the *Drosophila* genome may therefore not differ from other eucaryotic genomes.

Lastly, the investigated molecule was long enough to allow statistical analyses of the positioning of SARs. These analyses established that scaffold-associated regions can be regularly positioned on the DNA continuum: they appeared to be spaced by 11 kb intervals, or multiples thereof. This result agrees with the calculation of an average attachment point every 15.2 +/− 8.8 kb, based on the knowledge of the ratio of P-DNA to T-DNA (100/25) and of the measured average size of scaffold-associated fragments from the DNA continuum (3.8 +/− 2.2kb). Depending on how partitioning is interpretated, these regular contact points between DNA and the scaffold are present in all scaffolds, or are potential contact points realized in only part of them. Strong and weak SARs, analyzed independently, still showed this preferential periodic positioning. The range of values obtained for strong and weak SARs were similar, which suggested that strong and weak SARs are of a similar nature, at least from this point of view. These values (34 to 59 kb), and in particular those characterizing the largest DNA stretch, (34−38 kb) are compatible with the size usually reported for supercoiled DNA loops. Indeed, a value of 6 +/− 0.8 μm, i.e. 36 +/− 4.8 kb, was cited in reference 3.

The possibility that SARs might be, or not, DNA sequences effectively interacting *in vivo* with proteic partner(s) has been debated (38). 4 lines of evidence obtained from the present report and from our additional work support this proposal. 1—SARs are periodically disposed on the molecule. 2—They delimit loops whose size is compatible with that of *Drosophila* supercoiled loop domains. 3—Repeated sequences present on the DNA continuum are positioned next to scaffold-associated fragments (P. Surdej and R. Miassod, manuscript in preparation). 4—Sequences from the DNA continuum able to replicate autonomously in yeast strictly comap with a subset of the scaffold-associated DNA fragments (including those bearing strong SARs, weak SARs and those considered as not bearing SARs), and association of these fragments to scaffolds is conserved from *Drosophila* to yeast (C.Brun, Qi Dang and R. Miassod, manuscript submitted).

The looping of this long DNA continuum might therefore be envisaged as follows (see the scheme in Fig. 7). Strong SARs basically organize the DNA, in the majority of nuclei, into consecutive domains with sizes ranging from 15 to 115 kb. One or several transcription units are located within loops. 44 additional SARs subdivide these loops into smaller ones in a fraction of the nuclei population, depending on their status with regard to several cellular parameters such as cell cycle, differenciation or transcription. Within each large loop, these additional SARs may thus isolate different transcription units into distinct domains. According to this view, SARs would ensure a dynamic subdividing of genome, either isolating transcription units into distinct domains, or, alternatively, grouping them into a single domain. This hypothesis accounts for loop sizes, for partitioning, for fluctuation of scaffolding with tissue origin, but it does not accomodate the fact that some SARs may be overlapped with transcribed sequences or located within introns, the fact that the density of SARs and that of transcribed regions along the DNA continuum are apparently unrelated and the observed regular positioning. A second way of looking at the looping of this long DNA stretch is to hypothesize that both strong and weak SARs essentially play a structural role in the packaging of the DNA molecule. Scaffold-attached regions that were not considered as bearing SARs, if really interacting with proteic partners *in vivo*, would also, together with SARs, participate to this process so that there would be a potential contact point every 11 kb or multiples thereof. This interpretation takes into account the loop sizes, the regular density of SARs along the DNA continuum and the observed periodicity. A third possibility is that SARs participate to the replication processes, as suggested by our observation of their comapping with sequences involved in autonomous replication in yeast (C. Brun, Qi Dang and R. Miassod, manuscript submitted). This last possibility is compatible with all data reported in the present paper but not with the fact that this comapping concerned only a SAR subclass. One is therefore led to propose that in fact these three basic roles might be played by SARs. Some would ensure the delimitation of transcription domains, others the packaging, others initiation of replication and some both, so that three looping structurations would be superimposed. It is therefore highly probable that, in addition to playing a role in the nucleation assembly of topoisomerase II, or of histone H1, as recently established by Laemmli and coworkers (39,40), SARs may have multiple other roles.

ACKNOWLEDGEMENTS

We are indebted to D. Brandli for generating computerized statistical analysis programs and to F. Bourdillon for help in examination of statistical data. Helpful comments on the manuscript by J. Pradel are acknowledged.

REFERENCES

1. Newport, J. (1987) Cell 48, 205−217.
2. Delidakis, C. and Kafatos, F.C. (1987) J. Mol. Biol. 197, 11−26.
3. Buongiorno-Nardelli, M., Micheli, G., Carri, M.T. and Marilley, M. (1982) Nature 298, 100−102.
4. Marilley, M. and Gassend-Bonnet, G. (1989) Exp. Cell Res. 180, 475−489.
5. Dijkwell, P.A. and Hamlin, J.L. (1988) Mol. Cell. Biol. 8, 5398−5409.
6. Andeli, S., Klar, A., Meuth, M. and Cedar, H. (1989) Cell 57, 909−920.
7. Amati, B.B. and Gasser, S.M. (1989) Cell 54, 967−978.
8. Phi-Van, L. and Strätling, W.H. (1988) EMBO J. 7, 655−664.
9. Jarman, A.P. and Higgs, D.R. (1988) EMBO J. 7, 3337−3344.
10. Bodes, J. and Maass, K. (1988) Biochem. 27, 4706−4711.
11. Hiromi, Y., Kuroiwa, A. and Gehring, W.J. (1985) Cell 43, 603−613.
12. Mc Nabb, S.L. and Beckendorf, S.J. (1986) EMBO J. 5, 2331−2340.
13. Citri, Y., Colot, H.V., Jacquier, A.C., Yu, Q., Hall, J.C., Baltimore, D. and Rosbach, M. (1987) Nature 326, 42−47.

14. Greaves, D.R., Wilson, F.D., Lang, G. and Kioussis, D. (1989) Cell 56, 979–986.
15. Grosveld, F., Van Assendelft, G.B., Greaves, D.R. and Kollias, G. (1987) Cell 51, 975–985.
16. Stief, A., Winter, D.M., Stratling, W.H. and Sippel, A.E. (1989) Nature 341, 343–345.
17. Mirkovitch, J., Mirault, M.E. and Laemmli, U.K. (1984) Cell 39, 223–232.
18. Gasser, S.M. and Laemmli, U.K. (1986) Cell 46, 521–530.
19. Gasser, S.M. and Laemmli, U.K. (1986) EMBO J. 5, 511–518.
20. Mirkovitch, J., Gasser, S.M. and Laemmli, U.K. (1988) Cell 200, 101–109.
21. Izaurralde, E., Mirkovitch, J. and Laemmli, U.K. (1988) J. Mol. Biol. 200, 111–125.
22. Mirkovitch, J., Spierer, P. and Laemmli, U.K. (1986) J. Mol. Biol. 190, 255–258.
23. Cockerill, P.N. and Garrard, W.T. (1986) Cell 44, 273–282.
24. Cockerill, P.N., Yen, M.H. and Garrard, W.T. (1987) J. Biol. Chem. 262, 5394–5397.
25. Sykes, R.C., Lin, O., Hirana, S.J., Framson, P.E. and Chinault, A.C. (1988) Mol. Gen. Genet. 212, 301–309.
26. Käss, E. and Chasin, L.A. (1987) J. Mol. Biol. 198, 677–692.
27. Gasser, S.M. and Laemmli, U.K. (1987) Trends in Genet. 3, 16–21.
28. Cockerill, P.N. and Garrard, W.T. (1986) FEBS Lett. 204, 5–7.
29. Surdej, P., Got, C. and Miassod R. (1990) Biol. Cell 68, 105–118.
30. Shermoen, A.W. and Beckendorf, S.K. (1982) Cell 29, 601–607.
31. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory.
32. Feinberg, A.P. and Vogelstein, B. (1984) Anal. Biochem. 137, 266–267.
33. Savakis, C., Ashburner, M. and Willis, J.H. (1986) Dev. Biol. 114, 161–169.
34. Houde, M., Tiveron, M.C. and Brégégère, F. (1989) Gene 85, 393–404.
35. Loughney, K., Kreber, R. and Ganetsky, B. (1989) Cell 58, 1143–1154.
36. Freund, J.N., Vergès, W., Schedl, P. and Jarry, B.P. (1986) J. Mol. Biol. 189, 25–36.
37. Brown, N.H., King, D.L., Wilcox, M. and Kafatos, C. (1989) Cell 59, 185–195.
38. Cook, P.R. (1988) J. Cell Sci. 90, 1–6.
39. Izaurralde, E., Käs, E. and Laemmli U.K. (1989) J. Mol. Biol. 210, 573–585.
40. Adachi, Y., Käs, E. and Laemmli U.K. (1989) EMBO J. 8, 3997–4006.