MOLECULAR AND CELLULAR BIOLOGY, June 1989, p. 2322–2331
0270-7306/89/062322-10$02.00/0

Vol. 9, No. 6

Chromatin Loop Structure of the Human X Chromosome: Relevance to X Inactivation and CpG Clusters

ALAN H. BEGGS† AND BARBARA R. MIGEON*

Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received 2 December 1988/Accepted 21 February 1989

Part of the higher-order structure of chromatin is achieved by constraining DNA in loops ranging in size from 30 to 100 kilobase pairs; these loops have been implicated in defining functional domains and replicons and possibly in facilitating transcription. Because the human active and inactive X chromosomes differ in transcriptional activity and replication, we looked for differences in their chromatin loop structures. Since the islands of CpG-rich DNA at the 5′ ends of X-linked housekeeping genes are the regions where functional differences in DNA methylation and nuclease sensitivity are found, we looked for scaffold association of these sequences after extraction of histones with lithium diiodosalicylate. Specifically, we examined the 5′ CpG islands within the hypoxanthine phosphoribosyltransferase, glucose 6-phosphate dehydrogenase, P3, GdX, phosphoglycerate kinase type 1, and α-galactosidase loci in human lymphoblasts obtained from individuals with 1 to 4 X chromosomes. Although we detected no scaffold-associated regions near these genes, we found several such regions at the ornithine transcarbamylase and blood clotting factor IX loci. Our results suggest that the CpG islands are excluded from the nuclear scaffold and that even though transcriptionally active, housekeeping genes are less likely than X-linked tissue-specific genes to be scaffold associated. In all cases, the pattern of scaffold association was the same for loci on active and inactive X chromosomes.

Sex dosage compensation in mammalian females is achieved by random inactivation of one X chromosome (reviewed in reference 26). Active and inactive X chromosomes differ in a number of characteristics, including degree of condensation (3), timing of replication (18), and transcriptional activity (19). In addition, genes that are constitutively expressed are methylated differently on the two chromosomes (24, 46, 49, 50). The functional differences in DNA methylation reside in the clusters of CpG dinucleotides, the so-called CpG islands, that are commonly found at the 5′ ends of many housekeeping genes (4). These clusters are methylated only in genes on the inactive X chromosome, and methylation helps maintain the inactive state (23). Furthermore, since these clusters are nuclease hypersensitive in active but not in inactive genes (51), the chromatin structures of the two chromosomes must differ. Finally, methylation and gene activity are coordinately regulated at three CpG clusters near the glucose 6-phosphate dehydrogenase (G6PD) locus (46), which suggests that CpG islands may help define functional domains along the chromosome.

Although the differences in structure of the homologous X chromosomes during interphase are striking, little is known about the organization of X chromatin at the molecular level. The 300-Å (30-nm) chromatin fiber is thought to be organized into 30- to 100-kilobase-pair (kbp) loops that are anchored to a protein structure variously termed the nuclear matrix (7, 8, 37, 41–43, 48), scaffold (9, 15–17, 21, 22, 25, 33, 34), or cage (10, 11, 29). This model is appealing because it provides a means by which functional domains may be defined and organized within the nucleus.

Studies of chromatin loops, using 2.0 M NaCl as the means to extract histones, have shown that newly replicated DNA preferentially associates with the nuclear matrix (29, 48). Using similar conditions, others have shown that matrices prepared from steroid-treated chick oviduct are enriched for RNA precursors (7) and ovalbumin, a gene that is tissue specific and highly transcribed (8, 37). On the other hand, when histones are extracted by lithium 3′,5-diiodosalicylate (LIS) (33), attachment to the nuclear scaffold has been shown to occur at specific regions, which implies that loop structure is constant from one cell to the next. Many of these scaffold-associated regions (SARs) have been mapped near transcriptional control signals, for example, at 5′ ends of several *Drosophila* genes (16) and adjacent to the mouse Kappa light-chain enhancer (9).

Because the nuclear matrix is the site of DNA replication (29, 48) and perhaps RNA transcription (7, 8, 37, 42), differences in chromatin loop structure might account for the differences in replication and activity of the homologous X chromosomes. Therefore, we examined the chromatin organization of specific sequences on human X chromosomes. Of particular interest were the CpG clusters; not only do clusters and SARs map to the 5′ ends of genes, but both have been implicated in defining structural and functional domains within chromatin. The estimate by Bird of 30,000 CpG clusters per haploid genome (4) places them about every 100 kbp, a distance comparable to that between the bases of chromatin loops.

Using lymphoblast cultures from an XY male, an XX female, and a female with four X chromosomes (zero, one, and three inactive Xs, respectively), we looked for SARs at CpG islands and within the bodies of several housekeeping and tissue-specific genes. We found that clustered CpG dinucleotides in the promoter regions of six X-linked housekeeping genes were not associated with the nuclear scaffold, even when the relevant gene was transcribed. Furthermore, C+G-rich regions throughout the genome (*Msp*I tiny fragments) were also not scaffold associated. With the same filters used to study C+G-rich regions, we did identify SARs within the ornithine transcarbamylase (OTC) and blood clotting factor IX genes and near the autosomal phospho-

* Corresponding author. CMSC 10-04, The Johns Hopkins Hospital, Baltimore, MD 21205.

† Present address: Division of Genetics, Childrens Hospital, The Howard Hughes Medical Institute, Boston, MA 02115.

glycerate kinase (PGK) pseudogene. For all regions examined, we found no differences between active and inactive X chromosomes.

MATERIALS AND METHODS

Cell culture. Human lymphoblast cultures were initiated and maintained as described previously (32). To aid in quantitation, DNA was labeled by addition of [^{3}H]thymidine (0.1 μCi/ml) to the culture medium for 2 days before harvest.

Isolation of nuclei and nuclear scaffold preparation. Nuclei were isolated and extracted with 25 mM LIS by the procedure of Mirkovitch et al. (33) as modified by Cockerill and Garrard (9). Briefly, we extracted histones from isolated nuclei to obtain halos consisting of chromatin loops anchored at their bases to the nuclear scaffold. After digestion with various restriction endonucleases, scaffold-associated DNA was separated from solubilized sequences by centrifugation. After purification, equal amounts of total, pellet (scaffold-associated), and supernatant (scaffold-nonassociated) DNAs were analyzed by agarose gel electrophoresis, Southern blotting, and hybridization to probes of interest. Before extraction, the nuclei were stabilized by heating at 37°C for 20 min. Restriction enzymes (Bethesda Research Laboratories, Gaithersburg, Md.) were used at 500 to 1,000 U/ml in 3- to 5-h digestions, and scaffolds were pelleted in an Eppendorf microfuge for 3 min. To avoid partial digestion, it was crucial to completely remove all of the LIS after the extraction. Therefore, we washed the pellet with digestion buffer five times over the course of 2 h. Nevertheless, digestion of chromatin was incomplete in some cases, and such partial digests allowed us to examine larger regions than was otherwise possible. By convention, we refer to LIS-extracted nuclei as scaffolds and those extracted by the high-salt method as matrices (33).

Preparation of high-salt-extracted nuclear matrices. For 2.0 M NaCl extraction, nuclei were treated as described above except that the extraction buffer substituted 2.0 M NaCl for 25 mM LIS (5 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES]-NaOH [pH 7.4], 0.25 mM spermidine, 2 mM EDTA-KOH [pH 7.4], 2 mM KCl, 0.1% digitonin, 0.1 mM phenylmethylsulfonyl fluoride, 2.0 M NaCl).

DNA analysis. Unfractionated genomic DNA from nuclei as well as DNA from pellet (P) and supernatant (S) fractions of halo preparations were purified by standard procedures (50). The amount of unfractionated DNA was quantitated spectrophotometrically. From the specific activity of incorporated [^{3}H]thymidine, we determined the amount of pellet and supernatant DNA (3 to 10 μg) to load on 1% agarose gels. For each cell preparation, equal amounts of DNA were loaded in all lanes. Southern blotting to GeneScreen Plus (Dupont, NEN Research Products, Boston, Mass.) was performed as described previously (38). Filters were hybridized to probes labeled by the random-primer method (14), using 10^6 dpm/ml in 0.9 M NaCl–45 mM Tris (pH 7.5)–1% sodium dodecyl sulfate–0.2 mg of heparin per ml–0.1 mg of boiled salmon sperm DNA per ml. All filters were washed in 0.1% sodium dodecyl sulfate–0.1× SSC (SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 65°C unless otherwise indicated and autoradiographed at −70°C with intensifying screens. To reprobe filters, either radiolabeled probes were allowed to decay or filters were stripped of remaining probe in 0.4 N NaOH, followed by neutralization in 0.1× SSC–0.1% sodium dodecyl sulfate–0.2 M Tris (pH 7.5).

Analysis and interpretation of blots. One potential artifact of this system is that solubilized fragments larger than 5 to 6 kbp may partition into both P and S fractions (22, 34). Therefore, we limited our analysis to fragments below this size. Probing of the same filter with several probes provided evidence that the lack of hybridization in lanes giving no signal was not due to variations in the amounts of DNA loaded. In some cases, autoradiographic bands were quantitated with an LKB Ultrascan XL densitometer. For fragments showing matrix association, the proportion of the sequence associated with the scaffold was calculated as follows. The areas under the peaks in P and S fractions (determined by densitometry) were adjusted by multiplying the percentage of bulk DNA in that fraction (e.g., 30 and 70% for P and S fractions, respectively). These values were then normalized so that the sum of the two fractions would equal 100%.

Probes. Human DNA probes used were as follows. pPB1.7 is an intron I probe for X-linked hypoxanthine phosphoribosyltransferase (HPRT) (50). pGDT5b is a G6PD cDNA clone provided by G. Martini (27, 36). Probe pEB9 contains all of the sequences present in pGD3 and pGD1.4 (49), including the CpG clusters of P3 and GdX genes (1, 46) and the entire body of P3. pSPT/PGK recognizes the 5′ CpG cluster of PGK-1 (24), whereas the PGK-1 cDNA hybridizes to all four human PGK-related loci (40). The α-galactosidase (GLA) cDNA probe, p502 (47), lacks 20 bp of 5′ untranslated sequence. The 1,300-bp OTC cDNA (13), containing 80% of the total coding region, includes 500 bp of 3′ untranslated region and was a gift from K. Davies. W3′GXba, from D. Anson, contains a pseudocluster (CpG < GpCs) located 170 kbp 3′ to factor IX and is derived from c4.1 (35). The factor IX cDNA and genomic probes pXI, pVIII, pIII, and pIV were generous gifts from G. Brownlee (2).

End-labeling analysis of CpG cluster fragments. LIS-extracted nuclei were digested with *Eco*RI-*Hin*dIII. The DNAs purified from P and S fractions were redigested with *Msp*I, end labeled with [^{32}P]dCTP, and separated on 1% agarose gels as described elsewhere (23). Equal amounts of ^{32}P, and hence equal numbers of end-labeled fragments, were loaded to allow direct comparison of size distributions from autoradiographs.

RESULTS

Treatment of isolated nuclei with 25 mM LIS or 2 M salt has been shown to extract histones and result in the formation of nuclear scaffolds or matrices (reviewed in references 17 and 48). Our LIS-extracted nuclei, when stained with ethidium bromide, appeared the same by fluorescent microscopy as those prepared by others using the same technique (9, 48); that is, the fluorescent nucleus was surrounded by an extensive orange halo (data not shown). Comparable to findings in previous studies (9, 16, 17, 21, 22, 33, 34), 50 to 75% of the DNA was solubilized after digestion of the halos by restriction enzymes. Contrary to what might be expected if DNA sequences were partitioning randomly, in most cases hybridization signals were predominantly in one fraction or the other. Evidence that differences in signal intensity were not the result of loading artifacts comes from the observations that signals that were entirely in the S fraction with one probe might be predominantly in the P fraction when the same filter was rehybridized with another probe. Furthermore, on some blots, some of the multiple fragments recognized by a probe partitioned to the S fractions, whereas others were predominantly in the P fractions.

CpG clusters and C+G-rich regions are not associated with the nuclear scaffold. The human HPRT gene contains a 5′

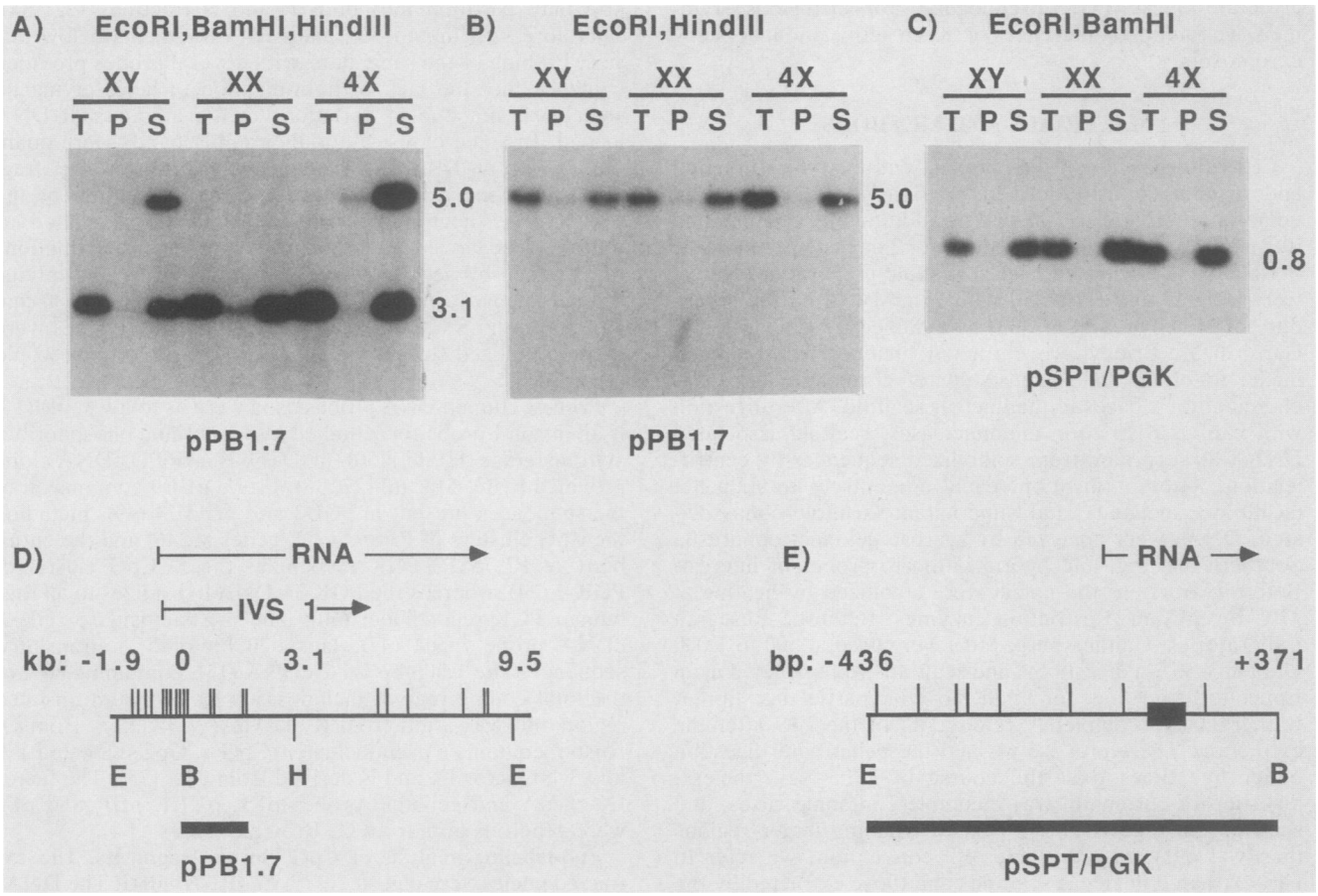

FIG. 1. Southern blots showing partitioning of CpG clusters in HPRT (A and B) and PGK (C) genes, probed with pPB1.7 and pSPT, respectively. Shown are lymphoblast DNAs purified from nucleus (T), pellet (P), and supernatant (S) fractions for a male (XY) and females with two (XX) or four (4X) X chromosomes; fragment sizes (in kilobase pairs) are indicated on the right. Maps of the 5′ regions of human X-linked HPRT (D) and PGK (E) genes show *Hpa*II (|), *Eco*RI (E), *Bam*HI (B), and *Hin*dIII (H) sites and probes. In panel C, the lanes with 4X DNA contain 7 rather than 10 μg; in panel E, the probe contain the entire 5′ CpG cluster, which is methylated differently on active and inactive X chromosomes. ▬, Exon 1.

CpG cluster that is methylated differently on active and inactive X chromosomes (50). *Eco*RI-*Bam*HI-*Hin*dIII-digested DNA probed with pPB1.7 yielded a 3.1-kbp *Bam*HI-*Hin*dIII fragment containing the 3′ half of the cluster (Fig. 1A, lanes T). In all three cell lines examined, this fragment partitioned predominantly into the S fractions and was not in the P fractions, which contained the scaffold-associated fragments. In addition, a 5.0-kbp band, resulting from partial digestion at the *Bam*HI site, was also visible in the lanes with supernatant DNA. Halo preparations digested only with *Eco*RI and *Hin*dIII confirmed that the entire CpG cluster was in the S fraction (Fig. 1B). Since there were no differences between cells with three inactive chromosomes (4X) and those without an inactive X chromosome (XY), it seems that this region does not associate with the scaffold regardless of methylation status or transcriptional activity.

We also examined the CpG clusters at the PGK-1 and GLA loci. The 807-bp *Eco*RI-*Bam*HI fragment containing the PGK-1 CpG cluster (24) partitioned predominantly to the supernatant fraction (Fig. 1C), as did two *Nsi*I fragments containing the CpG cluster at the GLA locus (5, 47) (not shown); none of these fragments were associated with the nuclear scaffold in either male, female, or 4X cells.

The 100-kbp region surrounding the human G6PD locus contains three genes, each with an associated 5′ CpG cluster (46; Fig. 2). The methylation statuses of all three clusters appear to be coordinately regulated and correlated with X-chromosome activity (46). Most of the G6PD CpG cluster was contained on a 1.5-kbp *Bam*HI fragment that we found exclusively in the S fraction of LIS-extracted halo preparations (Fig. 2A). The CpG clusters and coding sequences of the P3 and GdX genes were also in the S fractions (Fig. 2C).

We also analyzed a C+G-rich region located 170 kbp 3′ to the factor IX gene (35). In this pseudocluster, the CpGs are less than the GpCs, and several of the CpGs are methylated on both active and inactive X chromosomes (B. R. Migeon, unpublished data). Like true CpG clusters, none of the fragments from this pseudocluster were associated with the scaffold (data not shown).

Since the bulk of CpG clusters in the human genome are on autosomes, to determine whether lack of scaffold association was a general feature of C+G-rich sequences throughout the genome, we digested total genomic DNA with *Msp*I-*Eco*RI-*Hin*dIII and end labeled the fragments with ^{32}P. The small fragments (50 to 300 bp) obtained after gel separation and autoradiography (Fig. 3, lane T) were the same size as those obtained after *Msp*I or *Hpa*II digestion alone (12) and corresponded to DNA within CpG clusters (12). When DNAs fractionated and purified from *Eco*RI-*Hin*dIII digests of LIS-extracted halos were redigested with

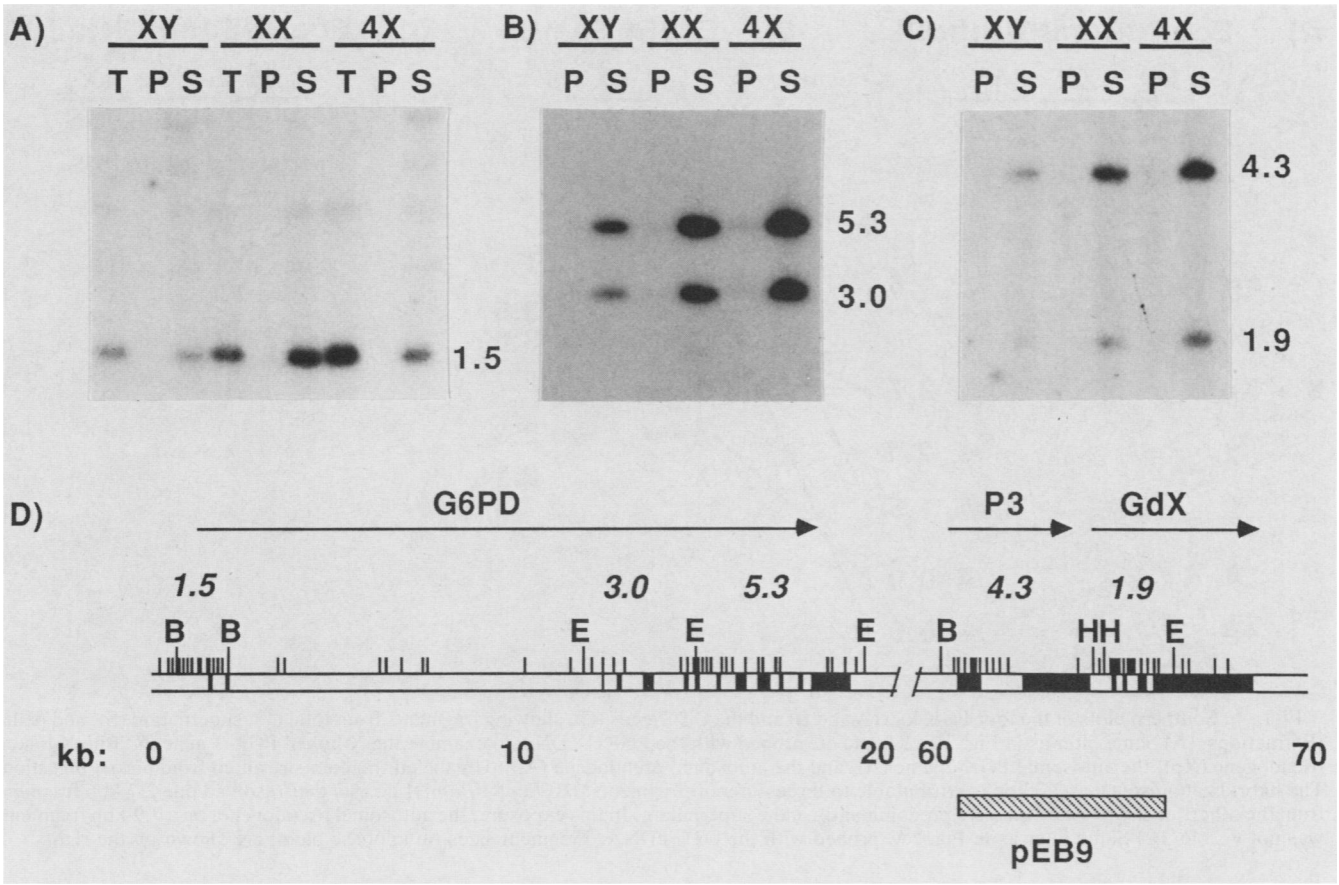

FIG. 2. Southern blots of the G6PD region. (A) Same blot as in Fig. 1A (*Eco*RI-*Bam*HI-*Hin*dIII), probed with G6PD 5′ cDNA (pGDT5b sub 10). The 1.5-kbp *Bam*HI fragment contained most of the 5′ G6PD CpG cluster. (B) *Eco*RI-*Bam*HI-*Hin*dIII digests probed with G6PD 3′ cDNA (pGDT5b sub 12). (C) Same filter as in panel B, probed with pEB9 and washed at 72°C to remove hybridization to repeated sequences. (D) Map of the G6PD region showing G6PD, P3, and GdX loci, exons (■), *Eco*RI (E), *Bam*HI (B), and *Hin*dIII (H) sites, and the approximate locations of *Hpa*II sites (|). Location of pEB9, used to analyze P3 and GdX, is indicated below the map. Sizes of relevant fragments (in kilobase pairs) are indicated above the map. S, Supernatant; P, pellet; T, total. In panel A, the reduced intensity of bands in the S lanes relative to the T lanes resulted from partial digestions; however, the lack of signal in P lanes clearly shows that these fragments were not matrix associated.

*Msp*I, the majority of tiny *Msp*I fragments were in the S fraction (Fig. 3). The depletion of tiny fragments in the P fraction is especially striking because the DNA in this fraction represented approximately five times the number of cells represented by the DNA in the S fraction for the following reasons: first, only 30% of the DNA in the halos partitioned with the P fraction; and second, although equal amounts of ^{32}P were loaded in P and S lanes, in fact the S fraction had incorporated two to three times more label than did the P fraction, as expected if the S fraction was enriched in tiny *Msp*I fragments. Furthermore, the enrichment of tiny *Msp*I fragments in the S fraction was seen even though many *Eco*RI-*Hin*dIII fragments greater than 5 to 6 kbp partitioned randomly, thus tending to randomize the distribution of *Msp*I tiny fragments. We obtained similar results for the other two cell lines (data not shown).

Search for SARs in X-linked housekeeping genes. Because the 5′ CpG clusters and significant amounts of flanking sequences in the promoter regions of the six housekeeping genes examined did not associate with the scaffold, we asked whether other regions of these genes were on the scaffold. Figure 2B shows LIS-extracted halo preparations probed with the 3′ portion of the G6PD cDNA, demonstrating that

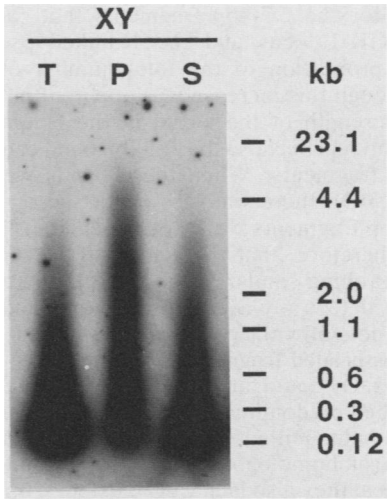

FIG. 3. End-labeling analysis of *Msp*I fragments from total (T), pellet (P), and supernatant (S) fractions of *Eco*RI-*Hin*dIII-digested halos prepared from an XY male. Size markers are shown on right.

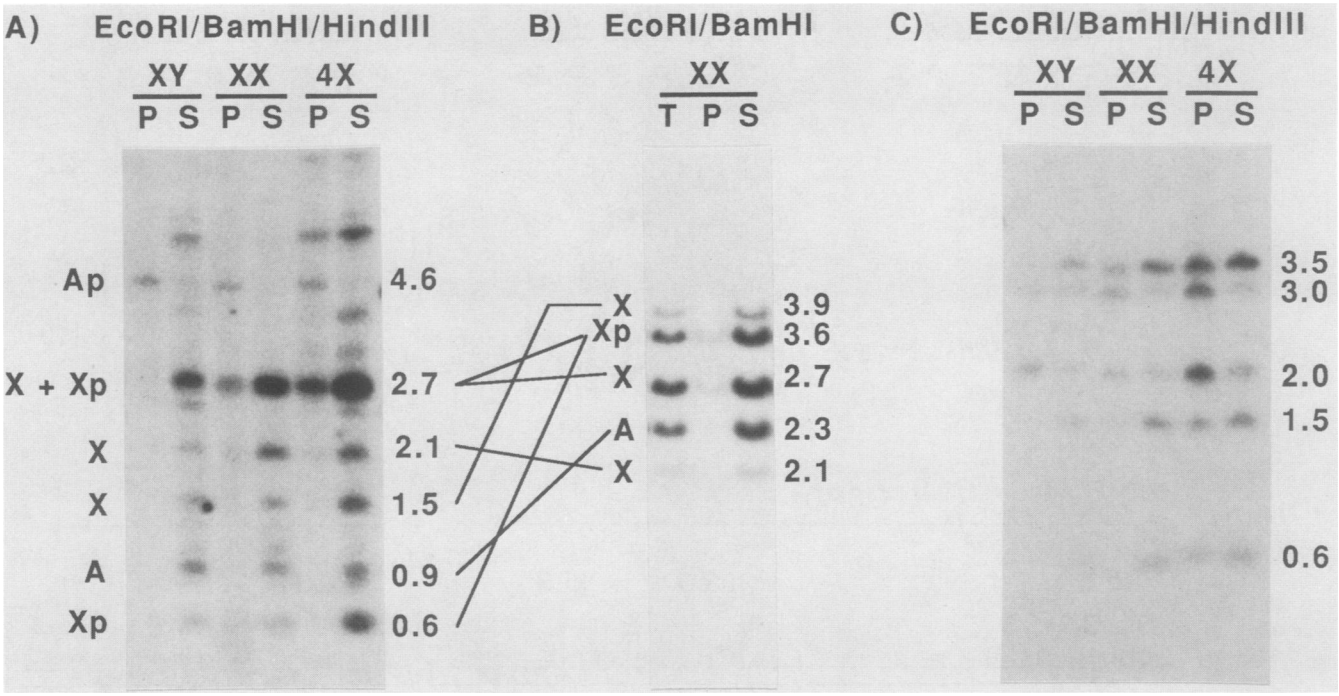

FIG. 4. Southern blots of the four PGK loci (A and B) and the OTC locus (C), showing fragments from total (T), supernatant (S), and pellet (P) fractions. (A) Same filter as in Fig. 2B, 2C, and 5C probed with the PGK-1 cDNA to examine the X-linked PGK-1 gene (X), the X-linked pseudogene (Xp), the autosomal PGK-2 gene (A), and the autosomal pseudogene (Ap). Unlabeled fragments resulted from partial digestion. The hybridization signal at 2.7 kbp is attributable to three different fragments. (B) *Eco*RI-*Bam*HI digests that resolved one 2.7-kbp fragment from the others and shows that all were predominantly in the supernatant. In this exposure, the autosomal pseudogene, on a 7.9-kbp fragment, was not visible. (C) Same filter as in Fig. 5A, probed with the OTC cDNA. Fragment sizes (in kilobase pairs) are shown on the right.

this relatively G+C-rich region (Fig. 2D) was also excluded from the scaffold.

Using a PGK-1 cDNA probe, we could identify restriction fragments derived from four discrete loci (28, 30, 31, 40, 45): the X-linked PGK-1 gene, the X-linked pseudogene, the autosomal, testis-specific PGK-2 gene, and the autosomal pseudogene (Fig. 4A and B). The only sequence enriched in the pellet was the 4.6-kbp *Hin*dIII fragment from the autosomal pseudogene. This autosomal fragment included the entire 1.3-kbp pseudogene and about 2.3 kbp of 5′ and 1 kbp of 3′ sequences and showed no X dosage effect. In contrast, the signals for the 2.7-kbp fragments that came from the X-linked PGK-1 locus and the X-linked pseudogene increased in proportion to the total number of X chromosomes; although this increase was apparent in both P and S lanes, the strength of the signal in the P lanes was more apparent than real because the 2.7-kbp band came from three hybridizing fragments. When these bands were better resolved (Fig. 4B), there was less pellet signal, showing that these 2.7-kbp fragments were not enriched in the pellet. We conclude, therefore, that there is no SAR at the X-linked PGK loci. Probing similar filters with the GLA cDNA, we also did not detect any SARs in the exons and neighboring sequences (not shown).

Scaffold-associated fragments are found in X-linked tissue-specific genes. OTC and factor IX are two X-linked genes that are transcribed predominantly in liver. Both are subject to X inactivation, but neither shows appreciable differences in methylation of homologous loci on active and inactive X chromosomes; they also lack CpG clusters (20, 38). Using an OTC cDNA, we could examine five unique fragments spread throughout the >85 kbp of DNA that encompasses this X-linked locus (Fig. 4C). Two (the 2.0- and 3.0-kbp fragments) were significantly enriched in the P fractions of all three cell lines. Although the other three fragments were not pellet enriched, they gave a significant pellet signal, in contrast to results obtained for G6PD, P3, and GdX with the same filter (Fig. 2B and C). Independent halo preparations digested with *Eco*RI-*Bam*HI gave essentially similar results (data not shown). The partitioning of the signal was the same regardless of the number of inactive X chromosomes.

Figures 5 and 6 show our analysis of the 35-kbp factor IX gene. Using pXI, we found that the promoter region including all of exon 1 and a *Kpn*I repetitive element partitioned into both fractions (Fig. 5A). Depending on the restriction digest, the proportion of fragments associated with the scaffold varied from 20 to 50%; however, the pellet was consistently more enriched with the 2.9-kbp fragments than with the 1.2-kbp fragments, which suggested that the SAR lay 5′ of the 1.2-kb fragment. With probe pVIII, which recognizes two *Nsi*I fragments containing sequences from introns 1 to 4 (Fig. 5B), we also detected some hybridization to the P fraction. Whereas less of the 4.6- and 2.8-kbp fragments were in the XY pellet than in the XX or 4X pellets (10 versus 20%), the proportion of attached fragments did not differ in the 2X and 4X females (one and three inactive Xs, respectively); therefore, the small difference between XY and XX DNAs cannot be attributed to differences in X-chromosome activity. The cDNA recognized another fragment that was highly enriched in the pellet in all cell lines (Fig. 5C). This 4.8-kbp *Eco*RI fragment contained parts of intron 4 and exon 5 and the other *Kpn*I reiterated element (Fig. 6). Using another halo preparation, we could map this SAR to a 3.7-kbp *Eco*RI-*Hin*dIII fragment (not shown); on the basis of densitometry, 75% of these fragments were associated with the scaffold. This was not true of the 7.2 kbp

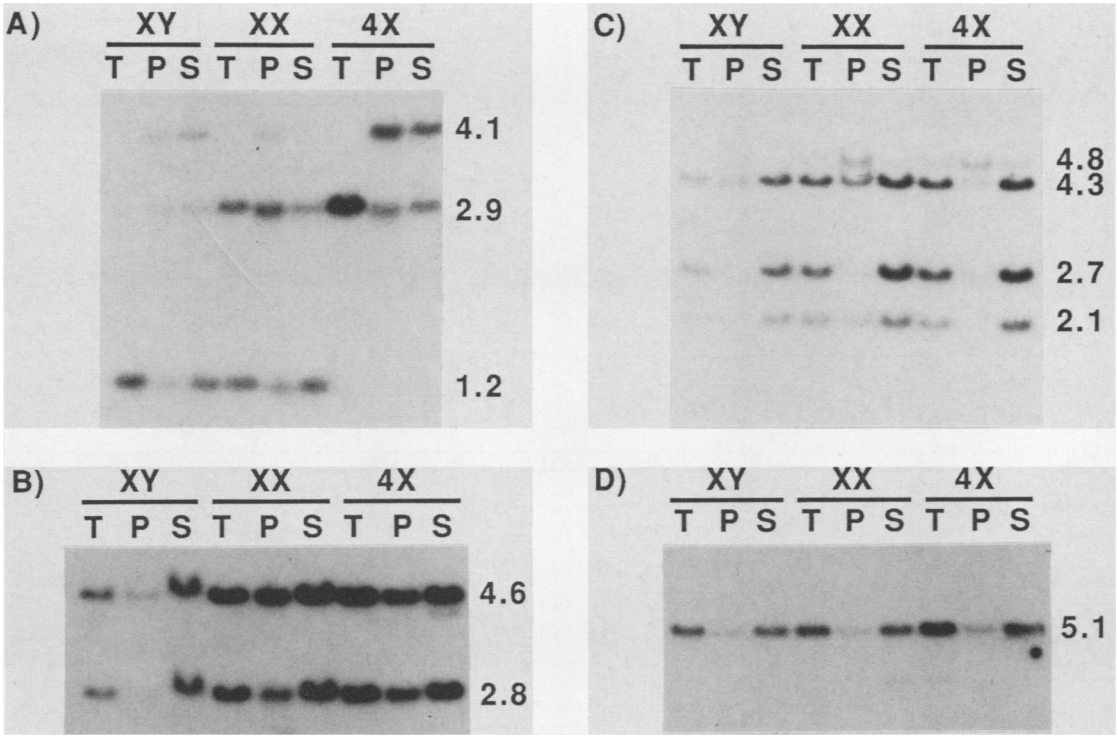

FIG. 5. Analysis of the factor IX locus. See Fig. 6 for locations of probes and origins of fragments. (A) *Eco*RI-*Bam*HI-*Hin*dIII digests (same filter as in Fig. 1A and 2A) probed with pXI to examine the 5′ portion of the gene. The XX female is heterozygous for a 1.2/2.9-kbp *Bam*HI polymorphism; the XY male and 4X female are hemizygous for the 1.2-kbp and homozygous for the 2.9-kbp alleles, respectively. The 4.1-kbp fragment resulted from partial digestion. (B) *Nsi*I digests probed with pVIII. (C) *Eco*RI-*Bam*HI digests (same filter as in Fig. 1C) probed with the cDNA and pIII. The 4.8-kbp fragment that is enriched in the pellet (P) contains exon 5, the 4.3-kbp fragment contains exons 7 and 8 and the 3′ end of the gene, and the 2.1- and 2.7-kbp fragments are homologous to pIII. The lanes with 4X DNA contained 7 instead of 10 μg. (D) *Eco*RI-*Hin*dIII digests (same as in Fig. 1B) probed with pIV to examine intron 6. T, Total fraction; S, supernatant fraction. Fragment sizes (in kilobase pairs) are shown on the right.

immediately 3′ to this region, as <10% of these fragments were in the pellet (probed by pIII and pIV; Fig. 5C and D and Fig. 6). The 4.3-kbp fragment containing the 3′ end of the gene was also predominantly in the supernatant, but some variability was observed; as in the 2X female, up to 19% of these sequences were in the P fraction.

As cleavage sites for topoisomerase II (Topo II) have been found within SARs (9, 15, 16, 22), we looked for them at the factor IX locus. We searched not only for sequences homologous to the *Drosophila* consensus (39) but also for those homologous to the chicken Topo II consensus sequence (44). We found 24 sites with a 14-of-15-bp match to the *Drosophila* Topo II sequence and 53 sites with a 17-of-18-bp match with the chicken sequence. One A+T-rich region (79.5%) between nucleotides 8380 and 8550 contained 15 of the 24 *Drosophila* consensus Topo II cleavage sites (4.6-kbp fragment; Fig. 5B); however, this fragment showed less scaffold attachment than did regions with fewer Topo II cleavage sites (e.g., 4.8-kbp fragment; Fig. 5C). Similarly, the SARs we identified did not seem to be particularly enriched in A or T boxes (16).

Although matrix association has been reported to be highly tissue specific when histones are extracted with high salt (8, 37), no tissue differences have been observed after LIS extraction (reviewed in reference 17). To determine whether the chromatin loop structure of the factor IX locus differed in expressing versus nonexpressing tissues, we prepared LIS-extracted halos from a fetal liver and hybridized this filter with the factor IX 5′-region probe pXI. The 2.9-kbp hybridizing fragments partitioned similarly in liver, where the gene was expressed, and lymphoblasts, where it was not (data not shown). Similar results were obtained when this filter was reprobed with the PGK-1 cDNA. Therefore, the patterns of association were similar in the two tissues irrespective of transcriptional activity.

Effect of high-salt extraction on chromatin loop structure. To determine whether our results (especially with respect to ubiquitously transcribed housekeeping genes) would differ if histones were extracted with high salt rather than with LIS, we repeated some of our studies on 2.0 M NaCl-extracted nuclei. *Eco*RI-*Bam*HI-*Hin*dIII digests were probed with the entire G6PD cDNA (Fig. 7A). As in LIS-extracted halos (Fig. 2A and B), none of the fragments were significantly enriched in the pellet. To examine a nonexpressed, tissue-specific locus, this filter was reprobed with two factor IX probes, pXI (Fig. 7B) and pIV (not shown); again, the results were similar for the two extraction methods; in both cases, the 5′ portion of factor IX was partially associated with the matrix, whereas intron 6 was predominantly in the supernatant.

DISCUSSION

Our studies of human genes, like those of other eucaryotes (reviewed in reference 17), demonstrate the highly structured organization of nuclear DNA. Because DNA replication seems to occur in association with the nuclear matrix, we might expect a small proportion of any given fragment to

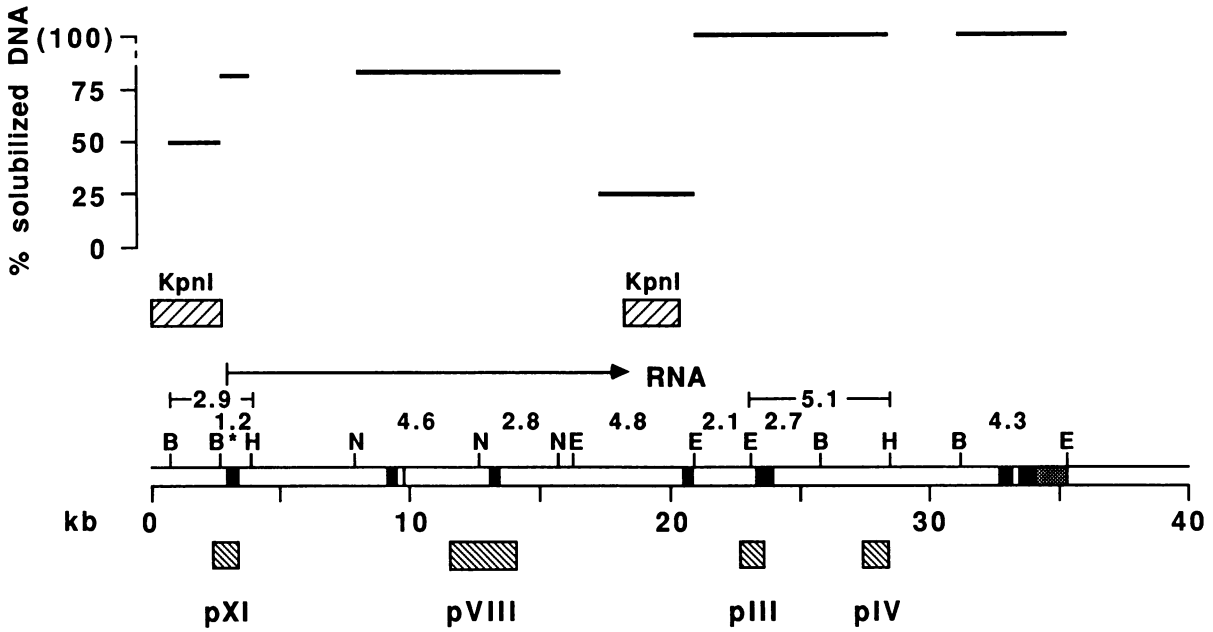

FIG. 6. Map of the factor IX gene and summary of scaffold association experiments. Shown are positions of exons (■), the 3′ untranslated region (▩), and the origin of restriction fragments shown in Fig. 5. Also shown are restriction fragment sites (*Eco*RI [E], *Bam*HI [B], *Hin*dIII [H], and *Nsi*I [N]), probes (pXI, pVIII, pIII, and pIV), and the two *Kpn*I repetitive elements. The fraction solubilized was obtained by quantitating multiple filters and pooling data for specific regions shown by the horizontal lines. The signals given by fragments not detectable in the pellet fraction were below the linear range of the film and densitometer and are arbitrarily shown as 100% solubilized.

be in P fractions. In fact, flow cytometric analysis shows that approximately 18% of our lymphoblasts are in S phase, so few cells are likely to be replicating the same sequence at any given time. Thus, it is not surprising that the partitioning of nonassociated fragments is so complete; i.e., little or no signal is seen in the P fractions. On the other hand, in regions that show scaffold attachment, the proportion of total DNA in the pellet is much more variable, ranging from 20 to 80%. Such variation may be a technical artifact, perhaps related to partial proteolytic degradation or disruption of SARs by LIS. Kas and Chasin (22) have suggested the possibility of two classes of SARs, one more susceptible to disruption than the other. Alternatively, some SARs may be attached only in a fraction of the cell population, since the scaffold attachment of the hamster *DHFR* gene is lost during mitosis (22). In addition, some variation may result from position effects such that nonattached regions in the vicinity of SARs may be trapped in the P fraction during the separation procedure (9).

CpG clusters are not on the scaffold. Many of the SARs described to date have been found near regulatory sequences (reviewed in reference 17). In *Drosophila melanogaster*, there are SARs in 5′ untranscribed regulatory regions of the histone, heat shock, alcohol dehydrogenase, and fushi tarazu genes as well as the glue protein gene *Sgs-4* (16, 33). Cockerill and Garrard (9) have also found an SAR in a

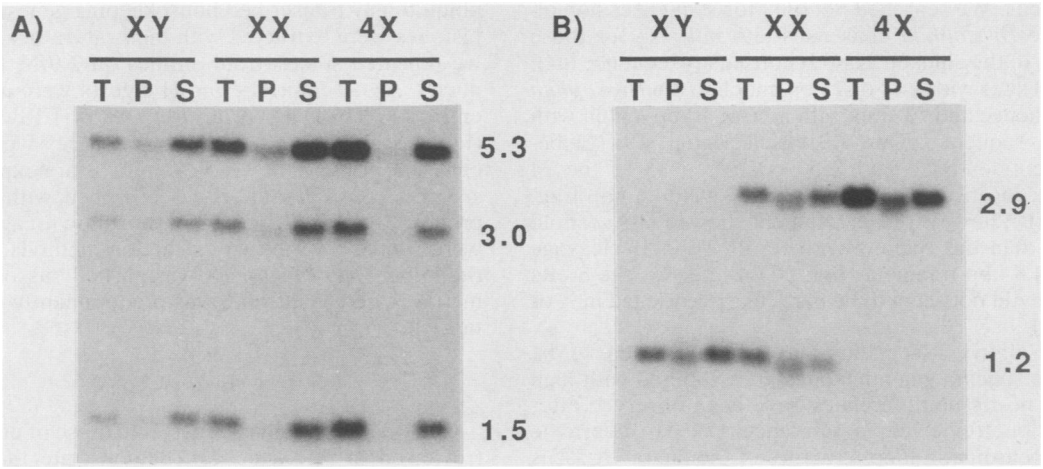

FIG. 7. Effect of high-salt extraction on matrix association of sequences at the G6PD (A) and factor IX(B) loci. Lymphoblast nuclei were extracted by the 2.0 M NaCl procedure (see Materials and Methods); halos were then digested with *Eco*RI-*Bam*HI-*Hin*dIII. (A) Probed with both 5′ and 3′ G6PD cDNAs; (B) same filter probed with factor IX pXI. Comparable LIS digests are shown in Fig. 2A, 2B, and 5A.

transcribed region next to the enhancer sequence in the mouse kappa light-chain gene. On the other hand, although implicated in transcriptional regulation of housekeeping genes, CpG islands were not associated with the scaffold in any of the six genes that we examined; these sequences partitioned completely into the non-scaffold-associated fraction. For four of these (HPRT, P3, GDX, and GLA), the sequences excluded from the scaffold included not only the island itself but also several kilobase pairs of flanking sequences on either side. It seems that the regulatory regions of housekeeping genes, at least those on the X chromosome, are scaffold independent.

Sequence specificity at SARs. Most SARs identified to date contain A+T-rich regions (reviewed in reference 17). We asked whether attachment sites could be found in C+G-rich DNA. Because CpG clusters in promoter regions of X-linked housekeeping genes seem to be excluded from the scaffold and because the tiny *Msp*I fragments in total genomic DNA are greatly enriched in the supernatant, it seems the CpG dinucleotides throughout the genome are not associated with the scaffold. Therefore, our results, together with previous studies of A+T-rich sequences, suggest that C+G-rich regions may be specifically excluded as sites for scaffold attachment.

Our analysis of the factor IX locus did not reveal any striking correlation between scaffold association and Topo II cleavage sites or other A+T-rich sequence motifs. This finding may not accurately reflect the situation in vivo because the *Drosophila* and chicken consensus sequences are degenerate and are derived from in vitro assays. On the other hand, the two tightly associated SARs at the factor IX locus include the two *Kpn*I repetitive elements in this gene. As Chimera and Musich (6) have shown that a class of *Kpn*I repeats is matrix associated in the African green monkey, the elements we have identified may belong to a subset of *Kpn*I sequences that mediate scaffold attachment.

SARs in tissue-specific genes. Not only are the regulatory regions of housekeeping genes not on the scaffold, but we observed no SARs in the exons of the G6PD, HPRT, and X-linked PGK-1 genes that we examined. Because we could not always examine the entire locus, we may have missed some attachment sites. In fact, an attachment site has been found in the middle of the constitutively expressed *DHFR* gene in hamsters (22). On the other hand, perhaps it is not merely fortuitous that by using cDNA probes we could easily detect SARs within the coding sequences of the two tissue-specific genes examined. There are SARs at the 5′ end and in the middle of the factor IX locus and at least two SARs at the OTC locus. The degrees of scaffold association vary, and these differences are reproducible in several restriction enzyme digests. Furthermore, although most sequences in the 3′ portion of factor IX are predominantly in the S fractions, there are some with detectable signals in the P fractions (Fig. 5C), compatible with the entire locus being close to the nuclear scaffold. Our data, albeit limited, raise the possibility that tissue-specific genes may be closer to the matrix than are housekeeping genes. Association with the matrix may be especially advantageous for highly transcribed genes, since it would facilitate accessibility to tissue-specific transcription factors, whereas housekeeping genes that are thought not to need tissue-speciifc signals may be located more distally on the loops in which they reside. The presence (or absence) of CpG clusters themselves may influence the size of chromatin loops and positioning of specific genes along the loop. Supporting this view, Jarman and Higgs (21) find that the human α-globin cluster (one of the few tissue-specific genes known to have a CpG islands [4]) has no SARs but that the β-globin cluster (no CpG islands) has several, some in close association with transcriptional regulatory elements. In any event, the attachment sites in the factor IX gene are the same in expressing and nonexpressing tissues, and the partitioning of the ubiquitously expressed PGK-1 locus is the same in liver cells and lymphoblasts. Therefore, our results support previous observations of *Drosophila* heat shock loci (42) and the mouse kappa immunoglobulin gene (9), i.e., that the scaffold attachment sites are the same irrespective of transcriptional activity. Furthermore, the promoter region of G6PD was not scaffold associated in halo preparations prepared either in LIS or in 2 M salt; therefore, transcription of this gene and the other X-linked housekeeping genes that we have examined does not require that the promoter region be attached to scaffold or matrix.

SARs and X-chromosome activity. The rationale for carrying out these experiments was to obtain insights into the molecular basis of X-chromosome inactivation. The differences in replication, transcription, and condensation between the homologous X chromosomes makes it highly likely that there are underlying differences in chromatin structure. However, the lack of detectable differences in the patterns of scaffold attachment indicates that the functional differences between the two X chromosomes are not mediated by attachment to the nuclear scaffold. Although the late onset of DNA replication that makes the inactive X chromosome unique among chromosomes might be detectable only if our studies were carried out in synchronized cell populations, we would expect to see differences in the locations of genes on the active and inactive chromosomes if transcription status affects or is affected by the position of the locus on chromatin loops. Instead, our data suggest that transcriptional status is not a primary determinant of scaffold association or lack thereof. In fact, the similarity of attachment sites in homologous genes on homologous chromosomes in the same cell, differing only in transcriptional activity, provides further evidence that the positions of genes in the loops, while perhaps influencing transcriptional potential, are independent of transcriptional activity. In any event, one needs to examine another level of chromosome organization to explain the characteristic differences between active and inactive X chromosomes.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant HD05465 from the National Institutes of Health.

We are grateful to Tracy Thranhardt for her contributions to these studies.

LITERATURE CITED

1. **Alcalay, M., and D. Toniolo.** 1988. CpG islands of the X chromosome are gene associated. Nucleic Acids Res. **16**:9527–9543.
2. **Anson, D. S., K. H. Choo, D. J. G. Rees, F. Giannelli, K. Gould, J. A. Huddleston, and G. G. Brownlee.** 1984. The gene structure of human anti-haemophilic factor IX. EMBO J. **3**:1053–1060.
3. **Barr, M. L., and E. G. Bertram.** 1949. A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. Nature (London) **163**:676–677.
4. **Bird, A. P.** 1986. CpG-rich islands and the function of DNA methylation. Nature (London) **321**:209–213.
5. **Bishop, D. F., R. Kornreich, and R. J. Desnick.** 1988. Structural organization of the human alpha-galactosidase A gene: further evidence for the absence of a 3′ untranslated region. Proc. Natl.

Acad. Sci. USA **85**:3903–3907.
6. **Chimera, J. A., and P. R. Musich.** 1985. The association of the interspersed repetitive KpnI sequences with the nuclear matrix. J. Biol. Chem. **260**:9373–9379.
7. **Ciejek, E. M., J. L. Nordstrom, M.-J. Tsai, and B. W. O'Malley.** 1982. Ribonucleic acid precursors are associated with the chicken oviduct nuclear matrix. Biochemistry **21**:4945–4953.
8. **Ciejek, E. M., M. J. Tsai, and B. W. O'Malley.** 1983. Actively transcribed genes are associated with the nuclear matrix. Nature (London) **306**:607–609.
9. **Cockerill, P. N., and W. T. Garrard.** 1986. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell **44**:273–282.
10. **Cook, P. R., and I. A. Brazell.** 1980. Mapping sequences in loops of nuclear DNA by their progressive detachment from the nuclear cage. Nucleic Acids Res. **8**:2895–2906.
11. **Cook, P. R., J. Lang, A. Hayday, L. Lania, M. Fried, D. J. Chiswell, and J. A. Wyke.** 1982. Active viral genes in transformed cells lie close to the nuclear cage. EMBO J. **1**:447–452.
12. **Cooper, D. N., M. H. Taggart, and A. P. Bird.** 1983. Unmethylated domains in vertebrate DNA. Nucleic Acids Res. **11**: 647–658.
13. **Davies, K. E., P. Briand, G. Ionasescu, R. Williamson, C. Brown, C. Cavard, and L. Cathelineau.** 1985. Gene for OTC: characterization and linkage to Duchenne muscular dystrophy. Nucleic Acids Res. **13**:155–165.
14. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
15. **Gasser, S. M., and U. K. Laemmli.** 1986. The organisation of chromatin loops: characterization of a scaffold attachment site. EMBO J. **5**:511–518.
16. **Gasser, S. M., and U. K. Laemmli.** 1986. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. Cell **46**:521–530.
17. **Gasser, S. M., and U. K. Laemmli.** 1988. A glimpse at chromosomal order. Trends Genet. **3**:16–22.
18. **German, J.** 1962. DNA synthesis in human chromosomes. Trans. N.Y. Acad. Sci. **24**:395–407.
19. **Graves, J. A. M., and S. M. Gartler.** 1986. Mammalian X chromosome inactivation: testing the hypothesis of transcriptional control. Somat. Cell. Mol. Genet. **12**:275–280.
20. **Hata, A., T. Tsuzuki, K. Shimada, M. Takiguchi, M. Mori, and I. Matsuda.** 1986. Isolation and characterization of the human ornithine transcarbamylase gene: structure of the 5'-end region. J. Biochem. **100**:717–725.
21. **Jarman, A. P., and D. R. Higgs.** 1988. Nuclear scaffold attachment sites in the human globin gene complexes. EMBO J. **7**:3337–3344.
22. **Kas, E., and L. A. Chasin.** 1987. Anchorage of the chinese hamster dihydrofolate reductase gene to the nuclear scaffold occurs in an intragenic region. J. Mol. Biol. **198**:677–692.
23. **Kaslow, D. C., and B. R. Migeon.** 1987. DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. Proc. Natl. Acad. Sci. USA **84**: 6210–6214.
24. **Keith, D. H., J. Singer-Sam, and A. D. Riggs.** 1986. Active X chromosome DNA is unmethylated at eight CCGG sites clustered in a guanine-plus-cytosine-rich island at the 5' end of the gene for phosphoglycerate kinase. Mol. Cell. Biol. **6**:4122–4125.
25. **Keppel, F.** 1986. Transcribed human ribosomal RNA genes are attached to the nuclear matrix. J. Mol. Biol. **187**:15–21.
26. **Lyon, M. F.** 1972. X chromosome inactivation and developmental patterns in mammals. Biol. Rev. **42**:1–35.
27. **Martini, G., D. Toniolo, T. Vulliamy, L. Luzzatto, R. Dono, G. Viglietto, G. Paonessa, M. D'Urso, and M. G. Persico.** 1986. Structural analysis of the X-linked gene encoding human glucose 6-phosphate dehydrogenase. EMBO J. **5**:1849–1855.
28. **McCarry, J. R., and K. Thomas.** 1987. Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. Nature (London) **326**:501–505.
29. **McCready, S. J., J. Goodwin, D. W. Mason, I. A. Brazell, and P. R. Cook.** 1980. DNA is replicated at the nuclear cage. J. Cell Sci. **46**:365–386.
30. **Michelson, A. M., C. C. F. Blake, S. T. Evans, and S. H. Orkin.** 1985. Structure of the human phosphoglycerate kinase gene and the intron-mediated evolution and dispersal of the nucleotide-binding domain. Proc. Natl. Acad. Sci. USA **82**:6965–6969.
31. **Michelson, A. M., G. A. P. Bruns, C. C. Morton, and S. H. Orkin.** 1985. The human phosphoglycerate kinase multigene family. J. Biol. Chem. **260**:6982–6992.
32. **Migeon, B. R., J. Axelman, and G. Stetten.** 1988. Clonal evolution in human lymphoblast cultures. Am. J. Hum. Genet. **42**:742–747.
33. **Mirkovitch, J., M. E. Mirault, and U. K. Laemmli.** 1984. Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell **39**:223–232.
34. **Mirkovitch, J., P. Spierer, and U. K. Laemmli.** 1986. Genes and loops in 320,000 base-pairs of the *Drosophila melanogaster* chromosome. J. Mol. Biol. **190**:255–258.
35. **Nguyen, C., P. Pontarotti, D. Birnbaum, G. Chimini, J. A. Rey, J. F. Mattei, and B. R. Jordan.** 1987. Large scale physical mapping in the q27 region of the human X chromosome: the coagulation factor IX gene and the *mcf-2* transforming sequence are separated by at most 270 kilobase pairs and are surrounded by several HTF islands. EMBO J. **6**:3285–3289.
36. **Persico, M. G., G. Viglietto, G. Martini, D. Toniolo, G. Paonessa, C. Moscatelli, R. Dono, T. Vulliamy, L. Luzzatto, and M. D'Urso.** 1986. Isolation of human glucose-6-phosphate dehydrogenase (G6PD) cDNA clones; primary structure of the protein and unusual 5' non-coding region. Nucleic Acids Res. **14**: 2511–2522.
37. **Robinson, S. I., B. D. Nelkin, and B. Vogelstein.** 1982. The ovalbumin gene is associated with the nuclear matrix of chicken oviduct cells. Cell **28**:99–106.
38. **Ruta-Cullen, C., P. Hubberman, D. C. Kaslow, and B. R. Migeon.** 1986. Comparison of factor IX methylation on human active and inactive X chromosomes: implications for X inactivation and transcription of tissue-specific genes. EMBO J. **5**:2223–2229.
39. **Sander, M., and T. S. Hsieh.** 1985. *Drosophila* topoisomerase II double-strand DNA cleavage: analysis of DNA sequence homology at the cleavage site. Nucleic Acids Res. **13**:1057–1072.
40. **Singer-Sam, J., R. L. Simmer, D. H. Keith, L. Shively, M. Teplitz, K. Itakura, S. M. Gartler, and A. D. Riggs.** 1983. Isolation of a cDNA clone for human X-linked 3-phosphoglycerate kinase by use of a mixture of synthetic oligodeoxyribonucleotides as a detection probe. Proc. Natl. Acad. Sci. USA **80**:802–806.
41. **Small, D., B. Nelkin, and B. Vogelstein.** 1982. Nonrandom distribution of repeated DNA sequences with respect to supercoiled loops and the nuclear matrix. Proc. Natl. Acad. Sci. USA **79**:5911–5915.
42. **Small, D., B. Nelkin, and B. Vogelstein.** 1985. The association of transcribed genes with the nuclear matrix of *Drosophila* cells during heat shock. Nucleic Acids Res. **13**:2413–2430.
43. **Small, D., and B. Vogelstein.** 1985. The anatomy of supercoiled loops in the *Drosophila* 7F locus. Nucleic Acids Res. **13**: 7703–7713.
44. **Spitzner, J. R., and M. T. Muller.** 1988. A consensus sequence for cleavage by vertebrate DNA topoisomerase II. Nucleic Acids Res. **16**:5533–5556.
45. **Tani, K., J. Singer-Sam, M. Munns, and A. Yoshida.** 1985. Molecular cloning and structure of an autosomal processed gene for human phosphoglycerate kinase. Gene **35**:11–18.
46. **Toniolo, D., G. Martini, B. R. Migeon, and R. Dono.** 1988. Expression of the G6PD locus on the human X chromosome is associated with demethylation of three CpG islands within 100 kb of DNA. EMBO J. **7**:401–406.
47. **Tsuji, S., B. M. Martin, D. C. Kaslow, B. R. Migeon, P. V. Choudary, B. K. Stubblefield, J. A. Mayor, G. J. Murray, J. A. Barranger, and E. I. Ginns.** 1987. Signal sequence and DNA-mediated expression of human lysosomal alpha-galactosidase

A. Eur. J. Biochem. **165**:275–280.

48. **Vogelstein, B., D. M. Pardoll, and D. S. Coffey.** 1980. Supercoiled loops and eucaryotic DNA replication. Cell **22**:79–85.
49. **Wolf, S. F., S. Dintzis, D. Toniolo, G. Persico, K. D. Lunnen, J. Axelman, and B. R. Migeon.** 1984. Complete concordance between glucose-6-phosphate dehydrogenase activity and hypomethylation of 3′ CpG clusters: implications for X chromosome dosage compensation. Nucleic Acids Res. **12**:9333–9348.
50. **Wolf, S. F., D. J. Jolly, K. D. Lunnen, T. Friedmann, and B. R. Migeon.** 1984. Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation. Proc. Natl. Acad. Sci. USA **81**: 2806–2810.
51. **Wolf, S. F., and B. R. Migeon.** 1985. Clusters of CpG dinucleotides implicated by nuclease hypersensitivity as control elements of housekeeping genes. Nature (London) **314**:467–469.