DNA Sequence Organization in *Drosophila* Heterochromatin

D. BRUTTAG, M. CARLSON, K. FRY, AND T. S. HSI

Department of Biochemistry, Stanford University, Stanford, California 94305

*Drosophila melanogaster* is an ideal organism for the study of the structure and function of heterochromatin because it has only four chromosomes and 25% of its genome is heterochromatic. All of the heterochromatin is constitutive and much of it is located in the sex chromosomes, which have been well characterized genetically and cytogenetically (Cooper 1959; Ashburner and Novitski 1976). The study of *Drosophila* heterochromatin is also attractive at the molecular level since the bulk of the DNA of these regions can be isolated as several discrete satellite DNAs in CsCl equilibrium gradients (Peacock et al. 1974; Endow et al. 1975; Brutlag et al. 1977a). These unusual DNAs consist of short nucleotide sequences (5–378 base pairs [bp]) repeated in tandem arrays over 1,000,000 bp long (Goldring et al. 1975; Brutlag et al. 1977a).

The biological roles of satellite DNA and heterochromatin have long been controversial. The molecular properties of satellite DNA make it unlikely that it is involved in transcription or the normal regulation of genetic expression. The misconception that satellite DNA has no essential function is based on the lack of genes in heterochromatin and on the viability of individuals with large heterochromatic deletions. However, individual *Drosophila* males carrying deletions of X heterochromatin have abnormal meiosis, defective spermatogenesis, and a markedly reduced fertility (Gershenson 1933; Sandler and Braver 1953; Peacock et al. 1975). These germ-line aberrations are directly correlated with the failure of deficient X chromosomes to pair properly during meiosis. Recent evidence shows that females carrying heterochromatic deletions also have a defective meiotic mechanism resulting in a reduced level of recombination (Yamamoto and Miklos 1977). These results indicate a strong selective pressure against chromosomes carrying heterochromatic deletions. They also argue strongly for a role of heterochromatin in the germ line rather than in somatic cells. Heterochromatin, therefore, may be dispensable for the survival of a cell or an individual, but it is essential for the survival of a chromosome in the germ line. A germ-line function is also consistent with the large variations of heterochromatin and satellite DNA between closely related species (Hennig and Walker 1970; Sutton and McCallum 1972; Gall et al. 1974). Indeed, the necessity for proper meiotic pairing of heterochromatin for successful gametogenesis or recombination suggests that variations in satellite DNA could lead to speciation.

To begin an analysis of these meiotic functions at the molecular level, we have studied the organization of the satellite DNA sequences that compose the bulk of the heterochromatin of *D. melanogaster*. We describe here the properties of two simple-sequence satellites (1.705 and 1.672 g/cm³) and of a complex satellite (1.688 g/cm³) which has been located primarily in the regions of the X and Y chromosomes essential for proper meiotic pairing (Gershenson 1940; Peacock et al., this volume).

**Simple-sequence Satellite DNAs of *D. melanogaster***

Figure 1 shows a summary of the highly repeated DNA sequences present in the heterochromatin of *D. melanogaster*. Each of these satellite DNA species constitutes 3–5% of the *Drosophila* genome; altogether, 19% of the haploid genome is highly repeated (Brutlag et al. 1977a). The three simple-sequence satellites (1.672, 1.686, and 1.705 species) have been isolated in a very-high-molecular-weight form, which suggested that chromosomeal regions containing a single satellite sequence may be larger than 1000 kb (Goldring et al. 1975). The physical homogeneity of each of the purified simple-sequence satellites has been demonstrated by neutral CsCl gradients, thermal denaturation, renaturation kinetics, and alkaline CsCl gradients which separate the complementary strands (Brutlag et al. 1977a).

The repeating units of each satellite have been determined by sequence analysis of either pyrimidine tracts or RNA complementary to the satellite DNA (cRNA). The sequence poly(AAGAG)TTCTC has been reported for the 1.705 satellite by Sederoff and Lowenstein (1975) from pyrimidine tract analysis and by Endow et al. (1975) from cRNA. Birnboim and Sederoff (1975) showed that pyrimidine tracts isolated from 1.705 DNA averaged 750 nucleotides in size. Transversion base substitutions must therefore be very rare in this DNA. The major repeating units of the 1.686 satellite were determined primarily from cRNA, and the arrangement of these three sequences with respect to each other is not known.
Figure 1. Sequences of highly repeated heterochromatic DNAs of *D. melanogaster*. The original sources for these sequences are: 1.672, Peacock et al. (1974); 1.672 and 1.686, Brutlag and Peacock (1975); 1.686 and 1.705, Endow et al. (1975); 1.705, Sederoff and Lowenstein (1975); 1.688, Carlson and Brutlag (1977); 1.705 and 1.688, this work.

(Fig. 1) (Brutlag and Peacock 1975; Endow et al. 1975).

Many results suggested a predominantly alternating arrangement of A and T in the 1.672 DNA (Fansler et al. 1970; Peacock et al. 1974; Endow et al. 1975). However, both physical and template properties of the separated complementary strands of the 1.672 DNA indicated that this satellite actually contained two distinct DNA species with the closely related sequences poly(AATAT) and poly(TTATA) (Brutlag and Peacock 1975).

The efficient replication of the 5-bp repeat, poly(AATAT) and poly(TTATA), by DNA polymerase in the complete absence of dGTP and dCTP indicates that very few GC substitutions have occurred within this sequence. Thus the simple-sequence satellites, 1.672 and 1.705, both seem to be very homogeneous in their repeating units.

The physical properties of the two 1.672 DNAs containing the 5-bp or 7-bp repeat are so similar as to make them inseparable by classical physical procedures. Moreover, the finding of two such closely related sequences within a single homogeneous satellite suggested to us that the sequence variation observed in many satellites may result from mixtures of sequences, sequences so similar that they might even cross-hybridize (Sutton and McCallum 1972; Blumenfeld 1974). We therefore developed procedures for cloning individual molecules of satellite DNA in recombinant plasmids to resolve individual sequences and to study sequence variations within a single molecule of satellite DNA.

Cloning Reveals Two DNA Species in the 1.705 Satellite

In our attempts to resolve satellite DNA into individual cloned molecules, we encountered instability of regions containing tandem repeats even in recombination-deficient bacteria (Brutlag et al. 1977b). Regions of satellite DNA in hybrid plasmids are often heterogeneous in length (Fig. 2). Subcloning bacteria containing the hybrid plasmids or transforming other bacteria with a single hybrid plasmid DNA molecule eliminates the heterogeneity, but it is often regenerated upon bacterial growth. These results show that DNA repeated in tandem may subject to rearrangements or recombinational instability within *recA* bacteria. Nevertheless, the cloned satellite DNA in these plasmids still hybridizes with satellite DNA from *Drosophila*, and we show below that it still contains the same repeating sequences.

To facilitate our analysis of the 1.705 sequences in each of these hybrid plasmids, we have utilized the restriction endonuclease *MboII* from *Moraxella bovis* (R. Roberts, unpubl.). Endow (1977) discovered that this enzyme cleaves 1.705 DNA into a series of short oligonucleotides 5 bp, 10 bp, and 15 bp in size. At first we felt that this indicated three repeating sequences in 1.705, one 5 bp, one 10 bp, and one 15 bp in length, each containing a single *MboII* recognition site. The mechanism of *MboII* cleavage, however, indicated that all these fragments could arise from a single repeating sequence. *MboII* recognizes the sequence (GAAGA) but cleaves the DNA eight nucleotides to the right of the terminal AT base pair of the recognition sequence (Sanger et al. 1977) (see legend to Fig. 3). Cleavage of poly(AAGAC) by this enzyme would result in products that were all integral multiples of 5 bp (Fig. 3). Products 20 bp or longer would contain both a recognition site and a cleavage site and would be cleaved further. Products 15 bp or smaller should be resistant to digestion. Cleavage of 1.705 satellite DNA with *MboII* produces the fragments shown in Figure 4. Most of the fragments are integral multiples of 5 bp and form a pattern which extends from
Figure 3. Fragments produced from the 1.705 satellite upon MboII digestion. The recognition site for the MboII restriction nuclease is shown in the box and the cleavage site is shown by the broken arrows. Note that the cleavage site is 10 bp or exactly one turn of the DNA duplex from the center of the recognition site. The enzyme leaves 3' hydroxyl and 5' phosphoryl termini. Cleavage of poly(AAGAG) (TTCTC) by this mechanism releases oligonucleotides which are integral multiples of 5 bp and which begin with 5' GAGAA TCTCTTTC and end with ... CTC. Cleavage near the 3' end of such a fragment results in the release of either a 5-bp or 10-bp or 15-bp fragment, each of which is too short to contain both a recognition and a cleavage site. The shortest oligonucleotide that can be produced by cleavage near the 5' end is a 15-bp fragment. Oligonucleotides 20 bp or longer resulting from cleavage at internal sites can always be cleaved further to yield 5-bp, 10-bp, and 15-bp fragments.

The MboII fragments longer than 15 bp have several possible origins. Control experiments in which SV40 DNA was included or more enzyme was added indicated that digestion was complete. However, due to the proximity of restriction sites in 1.705 DNA, we cannot eliminate the possibility that the binding of one molecule of MboII blocks recognition or cleavage by another molecule at a nearby site. A third alternative is that longer oligonucleotides result from the presence of sequence alterations in the 1.705 DNA which eliminate some MboII recognition sites.

Digestion of 15 of our hybrid plasmids with MboII confirms that they contain the sequence poly(AAGAG) (TTCTC). However, one of the plasmids (crDm705.5; see Fig. 4) contained a different repeating sequence. MboII digestion of this plasmid produced no short DNA fragments which were integral multiples of 5 bp. Instead, oligonucleotides 7, 14, 21, and 28 bp in size were present. Fragments of these sizes can also be seen in digests of purified satellite DNA. The recovery of this cloned fragment shows that the 1.705 satellite, like the 1.672 satellite,
Figure 4. Fragments produced by MboII endonuclease digestion of 1.705 satellite DNA and recombinant plasmid DNA containing 1.705 sequences. Satellite and plasmid DNAs were purified and digested as previously described (Brutlag et al. 1977b). After MboII digestion, samples were heated to 100°C for 2 min, cooled to 37°C, and bacterial alkaline phosphatase added. After 30 min the reaction was diluted with appropriate salts and [γ-32P]ATP to give final concentrations of 70 mM Tris-HCl (pH 8.05), 20 mM NaCl, 6.7 mM K2HPO4, 12.3 mM MgCl2, 16.6 mM 2-mercaptoethanol, 0.33 mM DTT, and 1.6 μM ATP. Polynucleotide kinase (Richardson 1965) was added to 33 units/ml and the reaction incubated at 37°C for 30 min. The MboII fragments were resolved by electrophoresis in a 20% polyacrylamide-urea gel at 15 V/cm (Maxam and Gilbert 1977) and detected by autoradiography. Indicated lengths of fragments were determined by external standards (not shown) and by direct sequence analysis.

consists of two discrete DNAs with different but closely related repeating sequences.

We have determined the sequence of the 7-bp and 14-bp fragments both from the cloned segment and from the total satellite DNA. Figure 6 shows a comparison of the sequences of the 7-bp pyrimidine strand of plasmid and of satellite origin. Unlike the fragments in the 5-bp series, these oligonucleotides contain an unpaired terminal G or C nucleotide. Thus the 7-bp and 14-bp fragments can only be arranged adjacent to each other, forming the repeating sequence poly(AAGAGAG)TTCTCTC (Fig. 7). All the hybrid plasmids containing either 1.705 sequence have satellite regions about 1 kb average size. This implies that poly(AAGAGAG)TTCTCTC must be segregated in chromosomal regions of at least this length. Using these hybrid plasmids as hybridization probes for in situ hybridization, we should be able to test whether each sequence has a distinct chromosomal location.

In addition to the major sequence obtained for the 7-bp fragment from satellite, one or more minor sequences appear to be present in this oligonucleotide (Fig. 6). By comparison, the 7-bp fragment from the cloned satellite shows no minor sequences despite prolonged overexposure of the autoradiogram. The minor sequences in the 7-bp oligonucleotide could arise either from sequence alteration of the 7-bp repeat or from insertion or deletion mutations in the 5-bp repeat, giving rise to unusual restriction-site spacing. Regardless of the origin of the variations in satellite DNA, it appears that we have cloned an extremely homogeneous region of the 7-bp repeated sequence.

The similarities between the two sequences in 1.672 DNA and 1.705 DNA (Fig. 1) are reminiscent of the relationship between repeating sequences observed in Drosophila virilis by Gall and Atherton (1974). Their analysis of three separable satellite DNAs showed each to consist of a repeating heptanucleotide which was related to the others by one or two base substitutions (TA → CG transitions). The 1.672 and 1.705 sequences described here are related by two or three base substitutions (TA → GC transversions). These two satellites of D. melanogaster share several other characteristics. The 5-bp repeat of each satellite contains very few base substitutions as described above. During purification these two species are the most widely resolved in all the gradients utilized, and yet 1.672 DNA is the major contaminant of the 1.705 satellite (Goldring et al. 1975). DNA molecules with 1.705 and 1.672 sequences covalently linked indicated that these species are adjacent in the genome (Brutlag et al. 1977a). This conclusion is supported by the cytological arrangement of the satellite DNAs (Peacock et al., this volume). Using clonally purified sequences of 1.705, we can now test whether the 5- and 7-bp repeats of each satellite are adjacent to each other.

The Complex 1.688 Satellite

The 1.688 g/cm² satellite is much more complex than the three simple-sequence satellites of D. melanogaster.
Figure 5. Sequence analysis of the 10-bp fragment from $MboI$ digestion of 1.705 satellite. These chromatograms are two-dimensional fractionations of snake venom phosphodiesterase partial digests of separated strands of the 10-bp fragment shown in Fig. 4. The mobility changes in the first and second dimensions are characteristic of the nucleotide lost from the 3' end as shown in the line drawings (Galibert et al. 1974; Jay et al. 1974). The sequences determined are arranged to show complementarity. After elution from the polyacrylamide gel, the strands were separated by homochromatography in one dimension. The strand composed of pyrimidines migrates faster than the one of purines. Transfer to the second dimension was carried out as described previously (Southern 1974).

5' GAGAAGAGAA
TCTCTCTCT
5'

7 mer FROM 1.705 SATELLITE DNA

Figure 6. Sequence comparison between the 7-bp fragments of 1.705 satellite and cloned 1.705 DNA. These two-dimensional fractionations resolve snake venom phosphodiesterase partial digests of the pyrimidine strands of the 7-bp fragments from either satellite DNA ($A$) or from cloned 1.705 DNA in crDm705.5 ($B$). Methods are as described in Fig. 5. The two sequences determined are identical; however, note the minor sequences present in satellite and absent in cloned DNA. The sequence shown is at the bottom of the figure along with the complementary sequence determined independently.

5' AGAAGAG
CTCTCTCT
5'
nogaster, as demonstrated both by renaturation studies and by restriction-site mapping (Carlson and Brutlag 1977). Several workers have shown that digestion of 1.688 DNA with either HaeIII or Hinf endonuclease generates a series of DNA fragments which are integral multiples of a 365-bp monomer (Fig. 8) (Manteuil et al. 1975; Shen et al. 1976; Carlson and Brutlag 1977). It was proposed that the arrangement of restriction sites in this satellite DNA evolved by random alteration of sites in an originally homogeneous array of 365-bp tandem repeats. Random alterations which varied the sequence of 25% of the HaeIII sites or 10% of the Hinf sites would quantitatively produce the pattern of fragments shown in Figure 8.

We have shown that the restriction sites are not altered in a random fashion by cloning segments of 1.688 satellite DNA. One hybrid plasmid (pDm688.26; see Carlson and Brutlag 1977) contained a 5.8-kb region of 1.688 satellite DNA with no altered HaeIII or Hinf sites in 16 tandem 365-bp repeats. On the other hand, we also showed that long regions of 1.688 DNA contained no HaeIII sites or no Hinf sites. In the HaeIII and the Hinf digests of Figure 8, we detected oligomers as long as 15 monomer units by hybridization with cloned monomer DNA. Moreover, the long DNA fragments near the top of the gel are resistant to these enzymes and also hybridize with cloned monomer DNA. These data suggest that there are long regions of 1.688 DNA missing every HaeIII site or every Hinf site and other regions which contain every site intact.

By cloning segments of 1.688 DNA we have demonstrated that the altered HaeIII sites are variant sequences rather than restriction sites with modified bases. Any bases specifically modified in Drosophila would not be similarly modified after propagation in bacteria. We have recovered several randomly cloned segments of 1.688 DNA which yield oligomer-length fragments upon HaeIII digestion. One cloned region yielded only monomers upon Hinf digestion but gave two large DNA fragments, in addition to a few monomer fragments, after digestion with HaeIII (pDm688.66 in Carlson and Brutlag 1977). Thus some of the repeats in this plasmid clearly had variant sequences at the expected positions for HaeIII sites. More direct proof that the missing HaeIII sites are due to sequence alterations comes from cloning the dimer fragments produced by HaeIII digestion of 1.688 satellite DNA. Seventeen independent plasmids contained HaeIII dimers.
which were still resistant to \textit{Hae}III after propagation in \textit{Escherichia coli}. We also cloned the long \textit{Hae}III-resistant DNA at the top of the gel in Figure 8. Six independent hybrid plasmids contained long \textit{Hae}III-resistant 1.688 regions. These data show that sequence alterations are primarily responsible for the pattern of restriction fragments shown in Figure 8. Furthermore, they confirm that there are long regions of 1.688 satellite DNA which contain no \textit{Hae}III sites.

The Repeating Sequence of 1.688 Satellite DNA

We have examined the sequence of the 1.688 repeat unit and variations within it by direct nucleotide sequence analysis of 15 tandem monomers in the plasmid pDm688.23. In addition to the \textit{Hin}f and \textit{Hae}III sequences, two \textit{Alu} sites facilitate sequence determination (Fig. 1). Using the technique of Maxam and Gilbert (1977), we have determined a preliminary sequence around both the \textit{Hae}III and \textit{Hin}f sites and in one direction from each \textit{Alu} site. Some of the data for the sequence around the \textit{Hin}f site are shown in Figure 9 and the preliminary sequence is presented in Figure 10.

The 15 tandem repeats present in pDm688.23 do not all have the same sequence. The procedure of Maxam and Gilbert (1977) usually results in a single possible nucleotide at any given distance from the restriction site. The sequence in Figure 10 is ambiguous, with at least two nucleotides possible at many positions. A variant sequence must be rather common among the 15 monomer units of pDm688.23 to be detected by this sequencing method. Also, the procedure of Maxam and Gilbert (1977) can only detect sequence variants that are related by transversion, i.e., only purine-pyrimidine mixtures can be detected. These technical limitations mean that there may be more variation among these 15 repeats than is shown in Figure 10. There also appear to be rather long regions which are free of sequence variation, in particular, around the \textit{Hae}III and \textit{Hin}f sites. This may explain the conservation of each of the 16 \textit{Hae}III and \textit{Hin}f sites in this cloned region of 1.688 DNA. The fact that we can obtain relatively unique sequences over long distances suggests that insertion and deletion mutations are rare. The primary mechanism of sequence variation appears to be base substitution. The presence of sequence variations among tandem repeats argues against hypothetical rectification mechanisms in the maintenance of 1.688 satellite (Thomas 1974).

This partial sequence shows that the basic repeating unit of the 1.688 satellite is 378 ± 5 bp and is not composed of a number of shorter repeats or inverted subunits. There is no relationship between the sequence of the 1.688 satellite and those of the three simple-sequence satellites of \textit{D. melanogaster}. Instead of the alternating AT or AG sequences found in 1.672 and 1.705 satellites, the 1.688 satellite contains runs of (dA) · (dT) or (dG) · (dC) up to 9 bp long.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure9.png}
\caption{Sequence analysis of cloned 1.688 satellite near the \textit{Hin}f site. DNA of the plasmid pDm688.23 was cleaved with \textit{Hin}f endonuclease and the 15 copies of the 1.688 monomer were purified by polyacrylamide gel electrophoresis. The fragments were then labeled at their 5' ends by a modification of the method of Maxam and Gilbert (1977) and further cleaved with \textit{Hae}III endonuclease. The 147-bp and 210-bp fragments, each labeled only at their 5' \textit{Hin}f ends, were again separated by polyacrylamide electrophoresis. Each fragment was then partially cleaved as described by Maxam and Gilbert (1977) specifically at the T and C residues, C residues, preferentially at G residues, or preferentially at A residues. All of the partial digests were resolved by several polyacrylamide gels over ranges extending from 2 bp to over 100 bp. Sequence ambiguity is seen most clearly in the 147-bp fragment beginning at nucleotide 17 at the bottom and reading up: T, T, T, C, C or G, G, etc. Because of the sequence ambiguity in these 15 monomer units, we did not analyze any sequences more than 70 nucleotides from any restriction site. The lengths of the 147-bp and 210-bp fragments, originally determined by polyacrylamide gel electrophoresis (Carlson and Bruttag 1977), are more precisely 154 ± 5 bp and 223 ± 5 bp, based on the partial sequence in Fig. 10.}
\end{figure}
Figure 10. Partial nucleotide sequence of the cloned 1.688 monomer repeating unit. This figure shows the sequence obtained as described in Fig. 9. The boxes locate the occurrences of (TTTCC) and single-base changes of that sequence. When a particular nucleotide was ambiguous, one base-pair was placed above and below the central sequence. No attempt could be made to quantitate which nucleotide predominated due to the strong sequence effects on the rates of partial cleavage. The sequence about the Hinf site was determined in Fig. 9. The sequence about the HaeIII site was determined similarly, except the plasmid DNA (pDm688.23) was cleaved first with HaeIII and the monomer was purified, labeled at the 5' end, and then cleaved with Hinf. The sequence from the Alul sites was determined by cleaving pDm688.23 DNA with Alul, purifying the resulting 196-bp and 181-bp fragments on polyacrylamide gels, and then labeling their 5' ends. The smaller fragment was cleaved with HaeIII, resulting in 56-bp and 125-bp fragments whose sequences provided overlap with the sequence about the HaeIII site. The 196-bp fragment is cleaved by Hinf into two fragments of similar size which we have not yet resolved. Therefore, there are about 29 bp missing in the center of the sequence and about 39 bp missing at the end. The first 60 nucleotides have had both strands analyzed (one strand from the HaeIII site and the other from an Alul site). The rest of the sequence, although determined by multiple gel analysis, must remain tentative. This is especially true of nucleotides within 7 bp of either Alul site.

Despite the lack of an internal repeating structure, there is some evidence for short-range periodicity.

The sequences (TTTCC) and (AAATTT) and single-base changes of them are very frequent components of this complex sequence. The sequence (AAATTT) is present eight times and single-base variations of it occur an additional eight times. There are nine exact repeats and 18 single-base changes of (TTTCC), which compose 43% of all the nucleotides in the repeat (Fig. 10). The precise frequencies with which these sequences are present in the 15 tandem repeats of pDm688.23 are difficult to determine because of the sequence variations observed.

The presence of short repeats within a more complex satellite DNA has often been taken as evidence for an ancestral simple-sequence satellite from which the more complex repeats have derived (Sutton and McCallum 1972; Cooke 1975; Biro et al. 1975; Southern 1975). However, the evolution of the 378-bp repeat of 1.688 from poly(TTTCC) does not
seem very likely. Since the sequence 5' (TTTCC) 3' occurs on both strands of 1.688 DNA, inversions would be necessary in addition to single-base changes in any hypothetical evolutionary scheme.

Moreover, the distribution of (TTTCC) (AAAGG) sequences indicates that either multiple overlapping inversions or insertions and deletions occurred as well. All such inversions, insertions, and deletions must have occurred prior to any amplification event that established the regular arrangement of restriction sites. Subsequent to the formation of this regular pattern, only single-base changes could be permitted since inversions, insertions, and deletions would alter the regular pattern. Although it does not rule out such an unusual series of events, this partial sequence suggests that the complex 1.688 satellite did not evolve from a simple-sequence satellite such as the 1.672 or 1.705 species. The presence of 5' (TTTCC) 3' on both strands also argues against the 1.688 satellite being an intermediate in the evolution of a simple-sequence satellite by unequal sister-chromatid exchange as proposed by Smith (1976).

An alternative proposal for the origin of short repeats within the 378-bp monomer is that sequences such as (TTTCC) (AAAGG) or (AAAAATTT) (TTTTAA) may have a functional significance and be subject to selection. This complex satellite has been located adjacent and proximal to the nucleolus organizers on both the X and Y chromosomes (Peacock et al., this volume). These are precisely regions that have been implicated as essential for proper meiotic pairing, chromosome segregation, and spermatogenesis (Gershenson 1940; Sandler and Braver 1953; Cooper 1964; Peacock et al. 1975). Were this satellite to be involved in recognition between the predominantly nonhomologous X and Y chromosomes, then the shorter repeats might be essential for sequence-specific DNA–protein interactions. Sequence-specific proteins could readily utilize sequences on both strands. Moreover, such interactions might be modulated by sequence variation, allowing different regions of satellite DNA to react differently, just as they do with the sequence-specific restriction endonucleases. We are currently using our knowledge of DNA sequence organization in heterochromatin to isolate chromatin fractions containing exclusively a single satellite sequence. By developing as stringent criteria for the homogeneity of chromatin fractions as we described here for DNA fractionation, we hope to demonstrate the presence or absence of sequence-specific proteins in heterochromatin.

Acknowledgments

This work was supported by grants from the National Foundation March of Dimes (Basil O'Connor Starter Grant) and from the National Institute of General Medical Science. M. C. is a National Insti-


