RESEARCH ARTICLE

Crystallographic Structure of the Octameric Histone Core of the Nucleosome at a Resolution of 3.3 Å

Rufus W. Burlingame, Warner E. Love, Bi-Chen Wang
Ron Hamlin, Nguyen-Huu Xuong, Evangelos N. Moudrianakis

Abstract. *The structure of the (H2A-H2B-H3-H4)$_2$ histone octamer has been determined by means of x-ray crystallographic techniques at a resolution of 3.3 angstroms. The octamer is a prolate ellipsoid 110 angstroms long and 65 to 70 angstroms in diameter, and its general shape is that of a rugby ball. The size and shape are radically different from those determined in earlier studies. The most striking feature of the histone octamer is its tripartite organization, that is, a central (H3-H4)$_2$ tetramer flanked by two H2A-H2B dimers. The DNA helix, placed around the octamer in a path suggested by the features on the surface of the protein, appears like a spring holding the H2A-H2B dimers at either end of the (H3-H4)$_2$ tetramer.*

The eukaryotic chromosome undergoes many structural changes as it condenses during mitosis and decondenses during the dynamic functional processes in the cell cycle. Almost all of the DNA in a chromosome is primarily complexed with histones. The histones and DNA are organized in repeating units called nucleosomes (*1*). The four core histones are complexed in an (H2A-H2B-H3-H4)$_2$ octamer (*2*) around which about 165 base pairs of DNA (*3*) are wrapped in two superhelical turns (*4–6*). Histone H1 appears to be bound to the outside of this particle and to interact with the spacer DNA which leads to the next nucleosome (*1*).

The histone octamer dissociates from the DNA in 2*M* NaCl at neutral *p*H as an octameric complex (*7*). In solutions of high ionic strength, this complex yields the same chemical cross-linking pattern (*2*) and spectral properties (*8*) as do the histones in chromatin since the equilibrium between the octamer and its subunits is shifted far toward association. At low ionic strength, or at nonneutral *p*H, the histone octamer dissociates into two H2A-H2B dimers and one (H3-H4)$_2$ tetramer (*7*). When these physiological subunits reassociate, they first form an (H2A-H2B)(H3-H4)$_2$ hexamer intermediate to which the second H2A-H2B dimer binds with positive cooperativity (*9*), forming the octamer. These and other studies (*10, 11*) on the mode and regulation of the assembly of the histone octamer in solution have demonstrated that the octamer is a dynamic tripartite entity.

Over the past several years, the low-resolution structure of the DNA-histone complex has been deduced by combining information from various sources, including x-ray (*6, 12, 13*) and neutron (*4, 5, 14*) diffraction, electron microscopy (*15, 16*), nuclease digestion of chromatin (*17, 18*), protein-protein (*2, 19, 20*) and protein-DNA (*21, 22*) cross-linking, and measurements of histone associations in solution (*7, 9*). Because of the central role of the histone octamer in the organization of DNA in chromatin, it is expected that the determination of its structure at high resolution should provide insight into several of the molecular processes that occur in the chromosome during the cell cycle. In this article, we report the structure of the unproteolized chicken erythrocyte (H2A-H2B-H3-H4)$_2$ histone octamer at 3.3-Å resolution determined by single crystal x-ray diffraction techniques. The histone octamer is a tripartite ellipsoid with a length of 110 Å and a diameter of 65 by 70 Å. Its general shape is that of a rugby ball. Furthermore, the resolution of the electron density maps at hand allows the visualization of polypeptide domains, from which we have thus far traced the backbone of the two chains in the H2A-H2B dimer.

Structure determination. Crystals of the histone octamer from chicken erythrocytes were grown as described (*23*). The space group is $P3_221$; $a = b = 118.7$ Å, and $c = 102.9$ Å, and the crystals diffract to 3.1-Å resolution. There is one half octamer per asymmetric unit, showing that the histone octamer has perfect twofold symmetry at least to this resolution. The solvent content of the crystal is 65 percent.

Heavy atom derivatives. In an extensive search for a heavy atom derivative to phase the diffracted x-rays, the most useful compound tested was tetrakis-(acetoxymercuri)methane (TAMM). Biochemical experiments with the TAMM derivative yielded information about the identity of the amino acid to which it bound and movement of that amino acid in the protein. Crystals soaked for a month in 5 m*M* TAMM in crystallization buffer without 2-mercaptoethanol showed no change in the pattern of a screened precession photograph. However, when a crystal was first soaked in 2 m*M* dithiothreitol (DTT), rinsed, and then soaked in 1 m*M* TAMM for 2 days, there was a change in the diffraction pattern. The chicken histone octamer contains only two cysteines, at position 110 on each H3, and it is known that these can be oxidized to form a disulfide bridge between the two H3 molecules (*24*). The fact that an effective mercurial derivative was obtained only after the protein in the crystal was reduced offers strong evidence that the single cysteine in H3 was specifically labeled. Furthermore, crystals of the TAMM derivative diffracted as well as native crystals and were used in the structure determination described below.

Data collection. Preliminary data from the native and derivative crystals were collected on a diffractometer (Syntex $P2_1$). By scaling together a series of native data sets, it was found that the crystals were isomorphous to 7-Å resolution, but there was some nonisomor-

Rufus W. Burlingame is a postdoctoral fellow in the Department of Biology, Johns Hopkins University, Baltimore, Maryland 21218. Warner E. Love is a professor in the Thomas Jenkins Department of Biophysics, Johns Hopkins University. Bi-Chen Wang is a member of the Biocrystallography Laboratory, V.A. Medical Center, Pittsburgh, Pennsylvania 15240, and a professor in the Department of Crystallography, University of Pittsburgh, Pittsburgh, Pennsylvania 15260. Ron Hamlin is an associate research physicist and Nguyen-Huu Xuong is a professor in the Department of Physics, University of California at San Diego, La Jolla 92093. Evangelos N. Moudrianakis is a professor in the Department of Biology, Johns Hopkins University, Baltimore, Maryland 21218.

phism at higher resolution, even though the lattice constants of the different crystals were virtually identical. In order to overcome this lack of isomorphism among crystals at high resolution, we utilized the Mark II multiwire area detector facility in San Diego (*25, 26*). The area detector, coupled with a rotating anode x-ray generator, collects data 100 times faster than a standard diffractometer, and with a tenfold improvement in the signal-to-noise ratio. On that diffractometer it was possible to collect a complete 3.1-Å data set (including Bijvoet pairs) with 3.5-fold redundancy from a single crystal. To circumvent any possible lack of isomorphism between the different crystals used to collect the native and derivative data, a large crystal (about 2.0 by 2.0 mm) was gently broken into a few fragments. One fragment was used for the native data set (after it was soaked in DTT) and one was treated with DTT and then with TAMM. The diffraction pattern of the fragments showed that no appreciable disorder, cracking, or mosaic spread was introduced when the parent crystal was cut. Complete data sets of the native protein (3.1 Å) and the TAMM derivative (3.3 Å) were collected from these crystal fragments. The ratio of intensity to background for the reflections at 3.5-Å resolution for both the native and TAMM derivatives is 10, and the R value for all equivalent reflections is about 0.05. The data were reduced in San Diego with the use of programs that have been described (*26*). All the data used in this structure determination thus came from a single crystal.

Data analysis. The heavy atom position in the TAMM derivative was determined by inspection of an isomorphous difference Patterson map. Least-squares refinement in space group $P3_221$ of centric reflections to 3.5-Å resolution yielded the position 0.322, 0.329, 0.992, and an R value of 0.49. The single mercury atom (see below) in the asymmetric unit is thus only 1.5 Å from the position with fractional coordinates of 1/3, 1/3, 0. If the mercury atom were exactly on the position defined by these coordinates, it would contribute to the phase of only one-third of the reflections, that is, those with indices such that $-h+k+l = 3n$. However, because the mercury atom does not lie exactly on this position, its position caused no significant problems in the determination of the structure reported here.

The initial phases were calculated from the single isomorphous replacement (SIR) and anomalous scattering (AS) differences to 3.9-Å resolution, and the SIR differences from 3.9- to 3.3-Å resolution, since the statistics for the AS data showed that the observed scattering differences became less reliable beyond 3.9-Å resolution. At Pittsburgh, we then used the iterated single isomorphous replacement (ISIR) method (*27*) to improve the 3.9-Å resolution phases and to break the phase ambiguity for reflections between 3.9- and 3.3-Å resolution. In the ISIR method, the initial electron density map is first transformed into a map that is related to the probability of finding molecules instead of atoms. The protein-solvent boundary is located from this map and the boundary information is used to construct a filter for removing noise from the map. An inverse Fourier transform is calculated and the phases from it are used either to improve the accuracy or to remove the phase ambiguity of the initial phases. During our analysis, we upgraded the noise filter one time after four cycles of phase improvement. The overall process took eight cycles. The final figure of merit for the approximately 13,000 reflections in the 3.3-Å data set was 0.65, and the average phase change from initial values was 44 degrees.

The handedness of the anomalous scatterer was determined by comparing statistics in the ISIR program generated from each of the two possibilities. The problem reduced to our choosing either space group $P3_121$ and a heavy atom position near 2/3, 2/3, 0, or space group $P3_221$ and a position near 1/3, 1/3, 0. The statistics favored the latter, and the choice was confirmed by the appearance of clear right-handed alpha helices in the electron density maps (Fig. 1). Cross-peak analysis of difference Patterson maps, refinement of the data with an additional heavy atom site, and three-dimensional difference Fourier maps all showed that there were no minor heavy atom sites. We know that we started with authentic TAMM because x-ray diffraction measurements of TAMM crystals made in this laboratory are identical with those published (*28*), and from the above results we know that there is just one mercury atom per asymmetric unit. Furthermore, two single mercurials, thimerosal and mercuric fulminate, gave data sets virtually identical to that of the TAMM derivative, albeit of lower statistical quality. We therefore conclude that the TAMM dissociated in the crystallization buffer into a species containing a single mercury atom. A difference Patterson map was calculated with only the anomalous scattering differences, and it was easily interpreted by inspection.

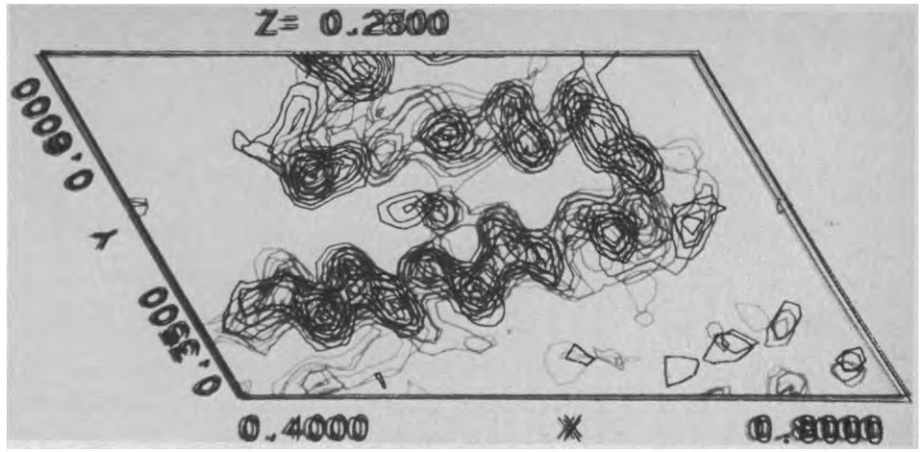


Fig. 1. A region of the electron density map that contains two alpha helices that have been assigned to H2A. The alpha helix at the bottom of the map is 30 Å long and the one at the top is 24 Å long.

Analysis of the electron density map. A number of points concerning the overall structure are immediately apparent upon observing the electron-density map. Because there is one half octamer per asymmetric unit, a molecular twofold axis must be coincident with one of the crystallographic twofold axes. The axis that relates one half octamer to the other half octamer in the same particle is clearly identified by the position of the mercury atom. The other class of twofold axis relates one complete octamer to a different octamer within the unit cell. The boundaries of the octamer are clearly visible because each octamer is surrounded by solvent channels 20 to 25 Å wide within the crystal lattice, and because the positions of the two classes of twofold axes put limitations on the boundary of the protein. The octamer is a prolate ellipsoid about 110 Å long by 65 Å high by 70 Å wide, and the long axis of the ellipsoid is perpendicular to the twofold symmetry axis. The overall organization of the electron density is clearly tripartite. The dimensions and shape of the octamer reported here, as well as the

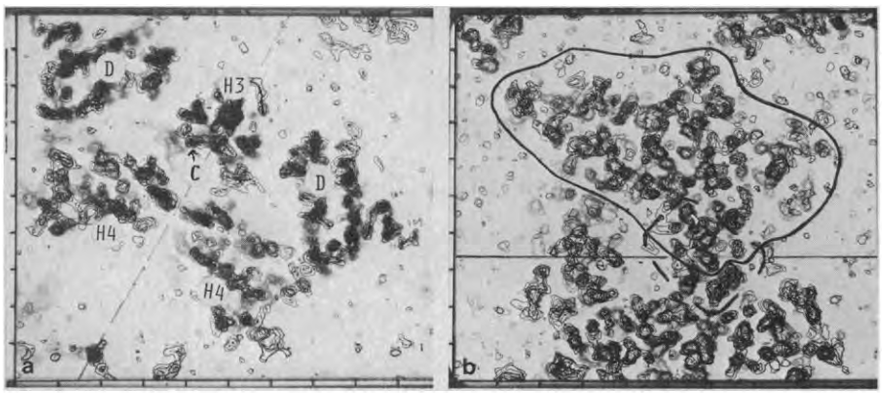


Fig. 2. The electron density maps, viewed down the crystallographic *c*-axis, have been orthogonalized for convenience. The vertical axis is 103 Å long and the horizontal axis is 119 Å long. (a) A portion of the electron density map 5 Å above and 5 Å below the molecular (and crystallographic) twofold axis, marked by the dashed diagonal line. Part of an H3 molecule straddles this axis, and the arrow coming from the "C" on the map marks the cysteine of that H3. Parts of two H4 molecules ("H4") lie closer to the bottom of the figure and are close to the twofold axis. The two H2A-H2B dimers (D) are separated from H3 and H4 by solvent channels on this level. There is also solvent between H3 and H4. Much larger solvent regions surround the electron density comprising a single octamer in all directions, allowing unambiguous assignment of density to a single octamer at this level. (b) The most complex region of electron density in the asymmetric unit. A section of the map 3 Å below and 7 Å above the twofold axis (the horizontal line) that relates two different octamers to each other. The density at this level of the map that is assigned to a single octamer is surrounded by the solid black line. The separation from two of the neighboring octamers (upper corners of the picture) is clear. The region that has not yet been completely resolved is surrounded by a dashed line and represents less than 5 percent of the total volume of the octamer.

orientation of its long axis, are radically different from those found in the diffraction studies of others (*14, 16*) (see below).

Assignment of subunits. The electron density was calculated for a unit cell, which contains exactly three histone octamers. The borders between most of the sites of contact among different octamers within the crystal lattice are clear (Fig. 2, a and b), and all but one have been established thus far (Fig. 2b). The exact placement of this one border affects the placement of only a small amount of electron density and does not affect any of our conclusions.

The electron density comprising one octamer is organized into three separate regions, a central "V" 30 Å wide at its tip and 75 Å wide at its base, flanked by two flattened balls, which are each roughly 40 Å in diameter. The mass in the central "V" region is loosely packed, containing solvent channels 5 to 15 Å wide and 20 Å long (Fig. 2a). Its interior is therefore accessible to solvent. Inspection of the density within this "V" reveals four contiguous regions which correspond to four polypeptide chains. The mass of each of the approximately 40-Å diameter flattened balls flanking the "V" is tightly packed and contains long stretches of alpha helix. We have been able to trace two polypeptide chains almost from one end to the other within each flattened ball. It was shown earlier that the $(H3\text{-}H4)_2$ tetramer and H2A-H2B dimer are stable entities in solution (*7*). Thus the four polypeptide chains in the "V" have been assigned to H3 and H4, and the two chains in each approximately 40-Å diameter ball have been assigned to H2A and H2B. Knowledge of the subunit organization of the octamer and the amino acid sequence of each polypeptide, combined with the salient features of the electron density map, has allowed us to directly identify the individual polypeptide chains.

Figures 1 and 2 are sections of the electron density map, and are used to illustrate the assignment of the subunits. One of the highest peaks on the map is 2 Å away from the mercury position (Fig. 2a). We interpret this peak as the sulfur of cysteine at position 110 of H3. The cysteines of the two H3's are 7 Å apart from each other, in good agreement with fluorescence studies (*29*), and are separated by the molecular and crystallographic twofold axis. At this point in the map, the cysteine is closer to the twofold axis than the bulk of the backbone of H3. When the cysteines were oxidized, they must have been only about 2 Å apart. Thus, the cysteine and at least some of the backbone of each H3 moved 2.5 Å after being reduced, for a combined movement of 5 Å. Defining the cysteines also defines the electron densities that belong to each of the H3 polypeptides, and these make up more than half of the "V." The backbone of H3 has not yet been unambiguously traced from end to end, but there are regions of electron density that are continuous with the cysteine, and these have been assigned to H3. For example, there is a thick rod of density 48 Å long (Fig. 2b) that runs from a point nearly in contact with H2A on one side to a point in contact with H4 on the other side. On close inspection it is seen to be made up of a 36-Å long alpha-helical stretch coming from near H2A. This abuts a separate stretch that continues toward H4 and then turns toward the center of the molecule. This region constitutes what we refer to below as the "propeller."

There are two domains within each foot of the "V" that are separate from H3 but that have continuity with each other. Each of these paired domains has been assigned to one H4, which is thus bilobal. One lobe of each H4 is wedged firmly between H3 on one side and the density assigned to H2B on the other side. The other lobe makes up part of the "foot" of the "V" and contacts an H4 from a different octamer within the crystal (Fig. 2b).

A clear channel 4 to 14 Å wide and up to 40 Å long separates each dimer from the centrally located tetramer (Fig. 2a), except for a few points of limited contact. The H2B chain was identified by two methods. There is a sequence of six amino acids in H2B where there are three tyrosines (*30*), and one or more of them can be cross-linked to H2A (*31*). There is also a sequence of four amino acids in H2B where there are two methionines. Electron density was found that could be interpreted as these regions. It is also known from solution studies that removal of the COOH-terminus of H2B reduces the affinity of the H2A-H2B dimer for the $(H3\text{-}H4)_2$ tetramer (*32*). In the electron density map, an arm of density wraps around from H2B to the tetramer, and if it is the COOH-terminus of H2B, this would explain the solution studies. This arm is part of the continuous chain that contains the density interpreted as tyrosine and methionine residues, and this chain is therefore considered to be H2B. There are large stretches of alpha helix within this polypeptide chain. The other chain within the dimer is therefore H2A. It has a striking set of three alpha helices, two of which can be seen in Fig. 1. These helices are 30, 24, and 37 Å long. Although each one may have a small nonhelical section in it, their continuity is unambiguous.

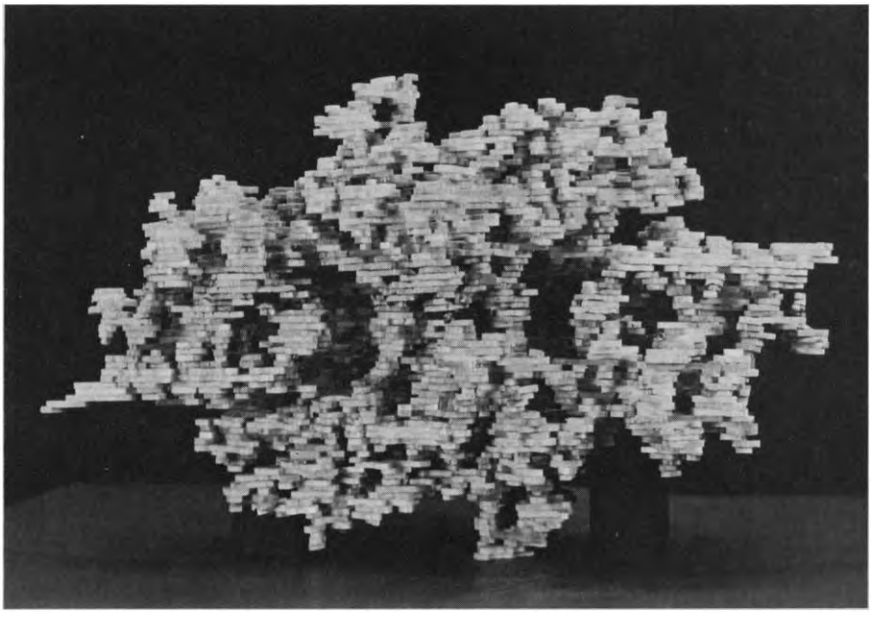
From a preliminary tracing of the backbones of all the polypeptide chains, we calculated that about 90 percent of the expected protein mass is accounted for. The remaining 10 percent probably

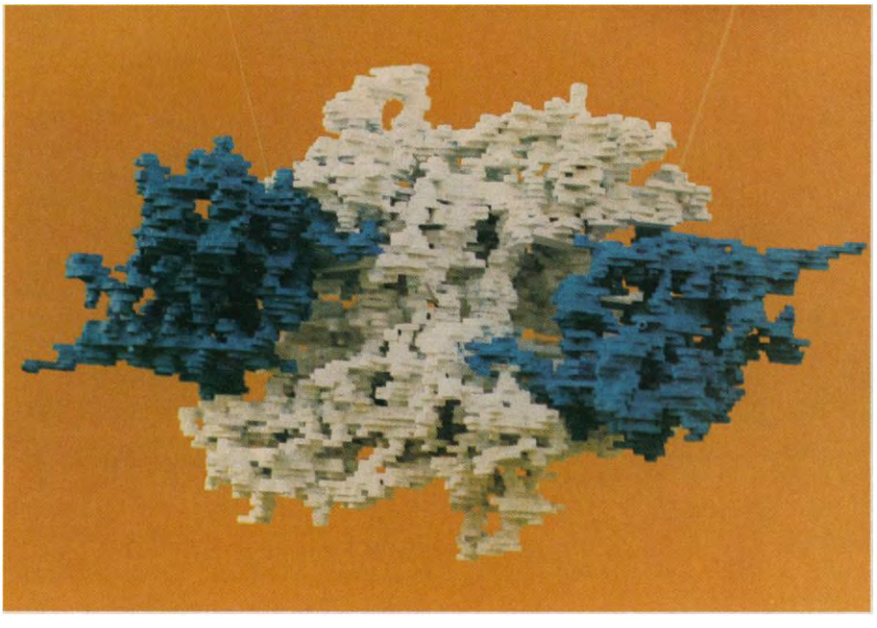
belongs to flexible termini of the histone chains, and would therefore be invisible to x-rays. The absolute assignment of the positions of the polypeptide chains must wait until the backbones for all the chains have been traced and the individual amino acids have been fitted to the electron density.

Analyzing the balsa wood model. We built a balsa wood model using the 3.3-Å data, scaled at 1 cm per 3.0 Å. It was built in two halves and virtually all of the electron density in an asymmetric unit was used for each half. The two halves were joined around the twofold axis of symmetry within the octamer. The lowest contour level is one-eighth that of the highest peak on the map. At this contour level, there were only three minor connections between the dimer and tetramer, all at the site of H2B-H4 interaction. The chains in the dimer are contiguous from one end to the other, as is all of the density within the tetramer.

Different views of the balsa wood model are shown in Figs. 3 to 8. The view along the crystallographic *a*-axis, which is the molecular and crystallographic twofold axis is seen in Figs. 3, 4, and 5. We refer to this part of the octamer as the "front," namely, the face where twofold axis intersects the centrally located tetramer at its narrowest point. The leftmost and rightmost edges on the long axis of the ellipsoid are the "ends." Figure 6 is a view along the crystallographic *c*-axis, showing the "top" of the octamer. The plane that is perpendicular to the *c*-axis and contains the *a*-axis is the "equator." Furthermore, the word "dimer" is used to mean the H2A-H2B dimer, and the word "tetramer" means the $(H3\text{-}H4)_2$ tetramer. The surface grooves of the structure that are pointed out define probable DNA binding regions.

Figure 3 is a view directly down the twofold axis at the front of the octamer. The solvent channels that separate the two dimers from the tetramer are apparent and emphasize the tripartite organization of the octamer. From this vantage point, the front of the tetramer is laterally biconcave. The central mass resembles an elongated left-handed propeller. The narrowest point of the tetramer (the vertical section of the propeller) is the H3-H3 contact through the twofold axis. This could be the site of H3-H3 chemical cross-linking (*2, 20*). The long helical





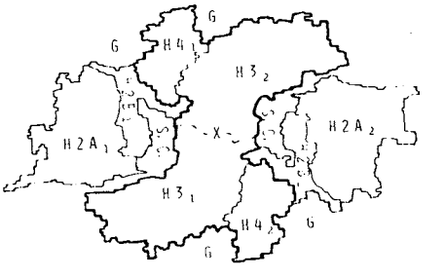


Fig. 3 (upper left). The front of the balsa wood model viewed directly down the twofold axis. The tripartite organization of the octamer is clearly visible even in the unpainted model, since large solvent channels (S. C. in Fig. 4) separate each dimer from the tetramer. There are grooves on the surface of the protein (G in Fig. 4) at the bottom of the dimer-tetramer interface on the right side and at the top of the interface on the left side, where we propose that the DNA interacts with the dimer. Other grooves can be seen at the top and bottom of the tetramer. Along the 110-Å axis, the histone tetramer has characteristics like a left-handed propeller. Fig. 4 (upper right). A tracing of Fig. 3 in which the polypeptide domains are delineated; this serves as a guide to understanding Figs. 3, 5, and 7. The numerical subscripts given to the histones identify their order along the DNA supercoil and do not correspond to any internal organization within the tetramer, since $H4_1$ binds most tightly with $H3_2$. The "X" marks the twofold axis, which is perpendicular to the plane of the page. Fig. 5 (lower left). Interpreted model of the histone octamer which has been painted to emphasize the organization of the subunits. The H2A-H2B dimers appear in blue, while the $(H3\text{-}H4)_2$ tetramer appears white. At the front, the COOH-terminal arm of H2B extends as a bridge to the tetramer.

region of H3 (the horizontal blade of the propeller) interacts at its end with a turn of H2B and part of a helix of H2A. This could be the site of chemical cross-linking between H2B-H3 and H2A-H3 (*2, 20*). The front lobe of H4 can be seen at each right-angle turn of the propeller. This is most probably the site of H3-H4 cross-links (*2, 20*). At the top and bottom surfaces of the tetramer, shallow troughs about 25 Å wide and 15 Å deep are apparent.

The two dimers nestle into the cavities at the front and sides of the tetramer. The COOH-terminal arms of H2B can be seen wrapping around the front of the tetramer about 8 Å above and below the equator, and interacting with H4 (Fig. 5). The end of one H2B arm is about 20 Å from the arm of the H2B from the other dimer. The lysines at the COOH-terminus of each H2B could span this approximately 20-Å gap, and could be the site of H2B-H2B cross-links found with formaldehyde (*20*). Both the COOH-terminal arm of H2B and its globular region make contact with H4, and these regions could be the sites of the H2B-H4 cross-links (*2, 19, 20*).

There are extensive areas of contact between H2B and H2A, as expected from solution studies (*7*) and cross-linking studies (*2, 20, 31*). The two H2A molecules occupy each end of the ellipsoid, and no close contacts are possible with another H2A molecule within the same octamer. However, in nuclei in 10 m*M* Mg^{2+} [known to induce chromatin condensation (*1*)] there are large amounts of H2A-H2A cross-links induced by formaldehyde (*20*). It seems that these must be contacts between the H2A's of two neighboring nucleosomes. The H2B-H2B cross-link found in the same study might also arise from H2B's in different octamers. This new interpretation of the cross-linking data puts certain restrictions on the packing of the octamer into higher order structures in chromosomes (*33*).

The view of the model down the crystallographic *c*-axis (Fig. 6a) shows that on either side, and closer to the front of the "V"-shaped tetramer are the two dimers. The tip of the tetramer does not protrude past them, but about 25 Å of each "foot" of the tetramer extends behind them. Grooves about 20 Å wide on the surface of the structure can be seen at the dimer-tetramer interface at the front of the model. The straight line through the middle of the model going from the front to the back is a piece of wire placed there to represent the molecular twofold axis. Figure 6b shows the model from the end. The orientation of the polypeptides within the dimer is clear from this angle.

Placement of DNA around the histone octamer. Neutron diffraction studies established that in nucleosomes the DNA is wrapped around the outside of the histone core (*4, 5*). Several additional types of evidence have indicated that nucleosome core particles have overall twofold symmetry (*18, 21*). This means that in nucleosomes the pseudo-twofold axis of symmetry of the DNA supercoil (if the base sequence in the DNA were palindromic, the axis would be perfectly twofold) must be coincident with the twofold axis of symmetry of the octamer. This puts a severe constraint on the possible path of the DNA around the octamer.

Guided by this information, we examined the surface of the balsa wood model of the octamer for morphological clues to DNA-binding sites. Simple inspection revealed that the dominant features of the surface of the model are well-defined grooves and ridges traversing it in a discontinuous left-handed spiral path 20 to 25 Å wide (Figs. 3, 4, and 6a). Tubing, with a diameter correctly scaled to represent DNA, was placed around the model so that it passed through the twofold axis and followed the path dictated by these grooves and ridges, seen in Figs. 3, 4, 7, and 8. Starting by the equator at the left front end of the model (Fig. 3), the tubing (DNA) first fits into the 20-Å-wide groove at the dimer-tetramer interface, then the 25-Å-wide troughs at the bottom and top of the tetramer, and finally into the right front groove. Two full superhelical turns of DNA in this orientation saturate the probable DNA-binding surface of the protein and yield a structure with a length of 110 Å and a diameter of 100 to 110 Å, containing about 168 base pairs of DNA (16 turns, with 10.4 base pairs per turn). These dimensions are the same as those predicted by neutron diffraction of nucleosomes (*5*), and x-ray diffraction (*12*) and electron microscopy (*15*) of chromatin (*34*), but are significantly different from diffraction studies of core particles and precipitated histone tubules (*4, 6, 13, 14, 16*). The number of

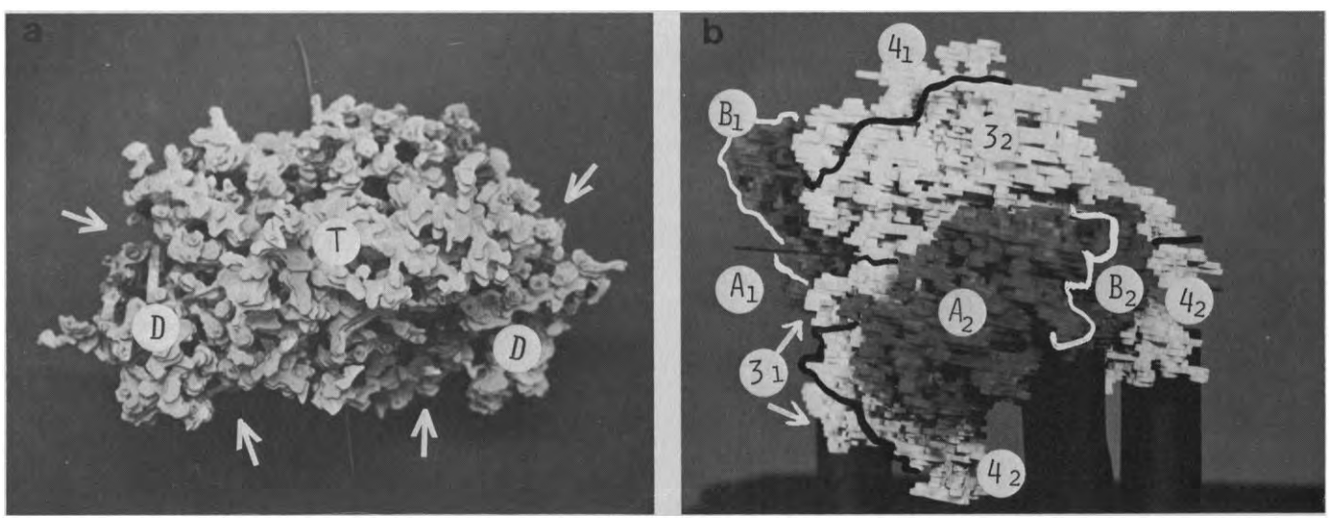


Fig. 6. (a) The balsa wood model viewed down the crystallographic *c*-axis. The tetramer (T) makes up the central portion of the structure, with the dimers (D) on either side. From this vantage point the separation of the dimers from the tetramer is not visible, because the solvent channels are obscured by ledges from the tetramer extending over them. The entrance and exit points of the channels are marked by arrows. The 30-Å-long alpha helix from the ledge of H2A (Fig. 1) can be seen in the dimer on the right, just above the D. The metal rod through the middle of the balsa wood model represents the twofold axis of the molecule. (b) The balsa wood model, which has been painted to accentuate the subunits, viewed from the right end. The dimers are dark and the tetramer is light. The domains of the individual polypeptides are delineated by thick lines and identified by the following code: H2A, A; H2B, B; H3, 3; and H4, 4.

base pairs of DNA in the model is consistent with the number of base pairs determined by nuclease digestion studies (*3*). The DNA appears like a spring around the tripartite core, with the dimers interacting with the ends of the spring (Fig. 7). The tetramer defines one and one-half turns of the DNA, and the dimers define the last one-quarter turn on each end.

Figure 7 shows the model from the same view as Fig. 3, but with the DNA wrapped around it in the orientation described above. In the upper left-hand and lower right-hand corners, part of H4 protrudes from the bulk of the protein and separates the turns of DNA at the regions where they are closest to each other. The flat ledge of H2A does not bind to the DNA, nor would it bind to the linker DNA if the path of the DNA were continued.

Figure 8 shows the model down the twofold axis from the back. Only from this angle does the particle give the illusion of being bipartite to the viewer, because the bulk of the tetramer and DNA obscure the dimers. The back lobe of each H4 is visible in this view. There is very little contact across the open region between each H4, which may explain why there is very little H4-H4 homo-dimer formed in the previously mentioned cross-linking experiments.

DNA-protein cross-linking data (*21, 22*) were not used to assign the identity of individual polypeptide chains within the octamer or to define the path of the DNA around the octamer. However, the above arrangement of DNA around the core histone octamer agrees extremely well with the results of the DNA-protein cross-linking experiments of Belyavsky *et al.* (*22*), lending support to the correctness of both studies (*22*). The cross-linking of H2A to the DNA near the front of the twofold axis, which would not be expected according to our model, only occurs in studies of the nucleosome core particle (146 base pairs of DNA and the octamer) (*21*), not in chromatosomes (I-nucleosomes, 165 or 175 base pairs of DNA, the octamer, and H1) or in chromatin (*22*). This observation suggests altered histone-DNA interactions among these structures (*22*).

We attempted to place the DNA around the balsa wood model in other ways, with the aim of generating a particle with a length closer to 55 Å, the length found by the others (*6, 13, 14*). Rotating the DNA supercoil 90 degrees from the position we have chosen yields a particle 70 Å long, but the diameter is increased to 140 Å, and the known histone-DNA contact points (*22*) are not generated. Furthermore, in this orientation there are no obvious surface features on the protein to dictate the DNA path.

Comparison of nucleosome structures derived by diffraction methods. Over the past 7 years, a number of low-resolution diffraction studies have yielded data that, when combined with the results of biochemical studies from several laboratories, allowed the modeling of the histone octamer and the nucleosome particle. In the models, both the path of the DNA (*6*) and the arrangement of the histones were proposed (*13, 14, 16*). However, the diffraction data contained information that directly defined only the size and overall shape of the particle. At 22-Å resolution, using image reconstruction of electron micrographs of precipitated histone tubules, the octamer was described as a wedge-shaped disk 55 Å long at its thickest point, with a diameter of approximately 70 Å (*16*). At 15-Å resolution, obtained from solvent contrast neutron diffraction of core particle crystals, the same general size and shape for the octamer were found, and its internal organization was described as tetrapartite (*14*).

At 25-Å resolution, the nucleosome core particle was first described as a bipartite wedge-shaped disk with flattened ends containing one and three-quarters turns of tightly packed DNA wrapped around the histone core. Its length was reported to be 57 Å at its thickest point and its diameter 110 Å (*6*). The 7-Å-resolution crystallographic study of core particle crystals gives refined information concerning the path and conformation of the DNA around the histone core (*13*), while other features of the model are similar or identical to the previous work (*6, 14, 16*). The length of the core particle in the 7-Å-resolution study can be calculated from the information given in the legend to figure 6 in that study (*13*) and is about 60 Å. Thus, the general size and shape of the particle, as well as the organization of the proteins and DNA, are internally

Fig. 7. This view, directly down the molecular twofold axis from the front, is the same as that of Fig. 3, but DNA has been placed around the model in a path suggested by features on the surface of the protein. The identification of the polypeptide chains is the same as in Fig. 4. On a scale relative to the balsa wood model, the DNA used for model building has a diameter of 20 Å, and a repeat distance for one turn of the double helix of 34 Å, making the DNA correctly proportioned to the protein. The DNA looks like a spring wrapped around a tripartite core, with the H2A-H2B dimers located on either end of the protein core, and protruding past the DNA on each end. Because the DNA-binding grooves at the dimer-tetramer interface of the model are wide, the double helix can roll closer to or farther away from the center of the structure. In the arrangement shown, the center-to-center distance between the DNA ends is 75 Å.

consistent within the above-mentioned series of studies.

In contrast to the above results (*6, 13, 14, 16*), we find the shape of the histone octamer to be roughly that of a rugby ball with a tripartite organization. It has a length of 110 Å and a diameter of approximately 67 Å (65 Å high by 70 Å wide), and its volume is about twice that previously assigned to it. The shape and volume of just one subunit of the octamer—the $(H3\text{-}H4)_2$ tetramer—are similar to that of the entire octamer of Klug *et al.* (*16*) and Bentley *et al.* (*14*). In addition, the model-built nucleosome (lacking H1) that we report here is a loosely packed structure 110 Å long and approximately 105 Å in diameter. There is at least 15 Å between each turn of the DNA, which is wrapped around a tripartite core, and parts of each dimer protrude past the DNA on each end, forming ledges.

The structures for the histone octamer and model-built nucleosome that we describe here are radically different from the previously reported structures (*6, 13, 14, 16*) and these differences cannot simply be attributed to the higher resolution of our study. Assuming that the diffraction data were correctly analyzed in all studies, we think the differences between the structures must have resulted from differences in the biological materials studied, either as a result of the preparation procedures or the state of the material when the data were collected. Thus the question arises, what is the structure of the histone octamer-DNA complex in vivo? While it is probable that nucleosomes exist in several states of compaction within the chromosome, it is also possible that the core particle represents an altered compaction state induced during its generation via the endonucleotic cleavage of chromatin. As new DNA ends are generated by the nuclease at the distal points of dimer-DNA contacts, "end effects" may cause a new state of equilibrium within the core particle, thereby yielding the reorganization of its components. Indications for such rearrangements exist in the literature. An additional protein-DNA cross-link is found in core particles which is absent in chromatin and chromatosomes (*22*). The size of the nucleosome in solution is greater than that of the core particle in solution (*4, 5*). Furthermore, the low salt transition of core particles is different from the low salt transition of the chromatosome lacking H1 (*35*).

It is also possible that the polypeptides in the histone octamer change their conformation and gain solvent when they dissociate from the DNA. Although this may be true at low ionic strength, it seems unlikely at high ionic strengths [greater than 2*M* NaCl equivalence, such as the crystallization buffer (*23*)] where the ionic shielding of the histones may be similar to what they experience when they are bound to DNA. This premise is supported by the reports that DNA-free histone octamers in high salt exhibit chemical cross-linking (*2*) and spectral properties (*8*) identical to those found for DNA-bound histones in chromatin. Furthermore, the results of x-ray diffraction (*12*) and electron microscopic studies of chromatin (*15, 36*) are consistent with a 110-Å repeating length within the chromatin fiber. The probable explanation for the large differences between the previous structures and our structures requires extensive analysis of the earlier chromatin literature (*33*).

Fig. 8. The back side of the painted model, viewed down the twofold axis. The dimers appear dark and the tetramer light. The path of the DNA follows the left-handed screw pitch of the protein surface. A large solvent channel parallel to the DNA runs through the back of the tetramer. The split in the back of the molecule is only about 15 Å deep. It does not separate the tetramer into two subunits since at the front of the molecule there is tight contact between H3 and H4, and between the two H3's.

Relationship of structure to function. A large body of evidence accumulated over the past 25 years points toward a dual role for the histones in the physiology of chromosomes: one direct and structural, the other indirect and regulatory. First, by associating with the DNA, the histones cause the right-handed double helix to condense about sixfold and form a left-handed supercoil. This condensed structure is further compacted via H1, probably in a second-order supercoil (or solenoid) (*36*), and then folded into even higher order structures. In its compact form, the DNA is kept organized and untangled and thus can be separated as individual chromosomes at mitosis. Second, by participating in the conformational transitions of chromatin, the histones can contribute to polymorphism of the DNA structure and thus influence the mode of interaction of the double helix with primary genetic regulators.

Our view for the functioning of the histone octamer in the compaction-decompaction cycle of chromatin is based on the structure described above, the physical-chemical properties of the histones bound to DNA (*19, 37, 38*) and free in solution (*7, 9, 10, 11*), and one set of biological experiments (*39*). The same low ionic strength that causes swelling in chromatin (*37*) also causes the loss of the ultraviolet light–induced cross-link between H2B and H4, while other cross-links remain (*19*). The cross-links between the histones and the DNA at this low ionic strength also remain (*40*). These data suggest that the histone dimer can lose its contact with the tetramer and that there is concomitant chromosome decondensation while both subunits remain bound to the DNA. That is, protein-protein interactions are minimized while protein-DNA interactions remain maximized. According to this, the H2A-H2B dimer and the $(H3\text{-}H4)_2$ tetramer are the physiological (or functional) subunits within the octamer in chromatin, as proposed earlier (*7, 11*). In accord with the above view is the proposal that at least part of the well-documented increase in nuclease sensitivity of the DNA in active genes (*1*) is the result of a relaxed DNA structure caused by the loss of contact of the histone dimer with the tetramer within the nucleosomes of these genes, while the histone subunits remain bound to the DNA.

Studies with newly replicated chromatin (*39*) have shown that the $(H3\text{-}H4)_2$ tetramer is the first subunit to bind to newly replicated DNA, and that about 10 minutes later the H2A-H2B dimers associate with the tetramer-DNA complex. We have interpreted the above experiment, taking into account the tripartite structure of the histones (the work reported in this article), and the finding of cooperative association of the histones in solution (*9*) (also indicative of a tripartite structure). This interpretation yields a model for chromatin decondensation that proceeds along the reverse path of its assembly. The separation of the first dimer subunit within the octamer allows the DNA around it to start to decondense, and the cooperativity within the histone complex then facilitates the sep-

aration of the second dimer from the tetramer. Conversely, in chromatin condensation, the binding of one dimer to the tetramer can stabilize the tetramer from conformational plasticity, and thus facilitate the binding of the second dimer. Since the COOH-termini of H2A (*10*), H2B, and H4 (*19*) are all involved in the dimer-tetramer interaction, it is likely that any molecular complex that causes decompaction within the chromosome will interact at this region of the dimer-tetramer interface, causing the opening up of the nucleosome. Once the two dimers have broken contact with the tetramer, further conformational changes can occur.

From the mercury binding studies during our structure determination, we know that the cysteines and the stretch of amino acids connected to them moved 5 Å within the middle of the tetramer. This is a significant distance for two polypeptide chains to move within the interior of a protein (*41*), and is indicative that conformational changes are available to this subunit (*38*). There is one large solvent channel at the back of the $(H3\text{-}H4)_2$ tetramer (Fig. 8), and a number of smaller channels within the interior of the tetramer, mostly near the equator (Fig. 2a). When this porous subunit associates with the H2A-H2B dimers, two additional long solvent channels are generated, one along each dimer-tetramer interface (Figs. 2a and 3). In our model, all these channels are still accessible to the solvent when the histone octamer is complexed with DNA to form a nucleosome (Figs. 7 and 8). We suggest that small molecules that can penetrate these solvent channels may modulate protein-protein and protein-DNA interactions in the nucleosome, resulting in coupled conformational changes in the protein-DNA complex, as well as opening and closing of the channels themselves. These conformational changes may ultimately be manifested as functional transitions in the chromosome.

References and Notes

1. J. D. McGhee and G. Felsenfeld, *Annu. Rev. Biochem.* **49**, 1115 (1980).
2. R. D. Kornberg and J. O. Thomas, *Science* **184**, 865 (1974); J. O. Thomas and R. D. Kornberg, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2626 (1975).
3. R. T. Simpson, *Biochemistry* **17**, 5524 (1978).
4. J. F. Pardon, D. L. Worcester, J. C. Wooley, K. Tatchell, K. E. Van Holde, B. M. Richards, *Nucleic Acids Res.* **2**, 2164 (1975).
5. R. P. Hjelm, G. G. Kneale, F. Suau, J. P. Baldwin, E. M. Bradbury, K. Ibel, *Cell* **10**, 139 (1977).
6. J. T. Finch, L. C. Lutter, D. Rhodes, R. S. Brown, B. Rushton, M. Levitt, A. Klug, *Nature (London)* **269**, 29 (1977).
7. T. H. Eickbush and E. N. Moudrianakis, *Biochemistry* **17**, 4955 (1978).
8. G. J. Thomas, Jr., B. Prescott, D. E. Olins, *Science* **197**, 385 (1977); R. I. Cotter and D. M. J. Lilley, *FEBS Lett.* **82**, 63 (1977).
9. J. E. Godfrey, T. H. Eickbush, E. N. Moudrianakis, *Biochemistry* **19**, 1339 (1980); R. C. Bennedict, E. N. Moudrianakis, G. K. Ackers, *ibid.* **23**, 1214 (1984).
10. T. H. Eickbush, D. K. Watson, E. N. Moudrianakis, *Cell* **9**, 785 (1976).
11. C. L. Hatch, W. M. Bonner, E. N. Moudrianakis, *Biochemistry* **22**, 3016 (1983).
12. J. F. Pardon, M. H. F. Wilkins, B. M. Richards, *Nature (London)* **215**, 508 (1967); B. M. Richards and J. F. Pardon, *Exp. Cell Res.* **62**, 184 (1970).
13. T. J. Richmond, J. T. Finch, D. Rhodes, A. Klug, *Nature (London)* **311**, 532 (1984).
14. G. A. Bentley, A. Lewit-Bentley, J. T. Finch, A. D. Podjarny, M. Roth, *J. Mol. Biol.* **176**, 55 (1984).
15. A. L. Olins and D. E. Olins, *Science* **183**, 330 (1974); P. L. Anderson and E. N. Moudrianakis, *Biophys. J. Soc. Abstr.* **9**, A-54 (1969); P. Oudet, M. GrossBellard, P. Chambon, *Cell* **4**, 281 (1975).
16. A. Klug, D. Rhodes, J. Smith, J. T. Finch, J. O. Thomas, *Nature (London)* **287**, 509 (1980).
17. D. R. Hewish and L. A. Burgoyne, *Biochem. Biophys. Res. Commun.* **52**, 504 (1973); M. Noll, *Nature (London)* **251**, 249 (1974).
18. R. D. Camerini-Otero, B. Sollner-Webb, G. Felsenfeld, *Cell* **8**, 333 (1976); B. Sollner-Webb, W. Melchior, Jr., G. Felsenfeld, *ibid.* **14**, 611 (1978).
19. H. G. Martinson, R. J. True, J. B. E. Burch, *Biochemistry* **18**, 1082 (1979).
20. V. Jackson, *Cell* **15**, 945 (1978).
21. A. D. Mirzabekov, V. V. Shick, A. V. Belyavsky, S. G. Bavykin, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4184 (1978).
22. A. V. Belyavsky, S. G. Bavykin, E. G. Goguadze, A. D. Mirzabekov, *J. Mol. Biol.* **139**, 519 (1980).
23. R. W. Burlingame, W. E. Love, E. N. Moudrianakis, *Science* **223**, 413 (1984).
24. R. D. Camerini-Otero and G. Felsenfeld, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5519 (1977).
25. R. Hamlin, *Trans. Amer. Cryst. Assoc.* **18**, 95 (1982).
26. Ng.-H. Xuong, S. Freer, R. Hamlin, C. Nielsen, W. Vernon, *Acta Cryst.* **A34**, 289 (1978).
27. B. C. Wang, *ibid.* **A40**, C12 (1984); B. C. Wang, in *Diffraction Methods for Biological Molecules*, volume of *Methods of Enzymology*, H. Wycoff, Ed. (Academic Press, New York, in press). An early objective of the ISIR method was to resolve the phase ambiguity problem in an SIR without AS information or in a single wavelength AS set.
28. D. Grdenic, B. Kamenar, B. Korpar-Colig, M. Sikirica, G. Jovanovski, *Chem. Commun.* (1974), p. 646.
29. M. Zama, P. N. Harrington, A. L. Olins, D. E. Olins, *Cold Spring Harbor Symp. Quant. Biol.* **42**, 31 (1977).
30. D. K. Grandy, J. D. Engle, J. B. Dodgson, *J. Biol. Chem.* **257**, 8577 (1982).
31. R. J. DeLange, L. C. Williams, H. G. Martinson, *Biochemistry* **18**, 1942 (1979); H. G. Martinson, R. True, C. K. Lau, M. Mehrabian, *ibid.*, p. 1075.
32. E. N. Moudrianakis, unpublished observations.
33. _____ and R. W. Burlingame, in preparation.
34. R. D. Kornberg, *Annu. Rev. Biochem.* **46**, 931 (1977).
35. A. E. Dieterich, H. Eshaghpour, D. M. Crothers, C. R. Cantor, *Nucleic Acids Res.* **11**, 2475 (1980); V. C. Gordon, V. N. Schumaker, D. E. Olins, C. M. Knobler, J. Horwitz, *ibid.* **6**, 3845 (1979).
36. F. Thoma, Th. Koller, A. Klug, *J. Cell Biol.* **83**, 403 (1979).
37. R. D. Carlson and D. E. Olins, *Nucleic Acids Res.* **3**, 89 (1976).
38. J. B. E. Burch and H. G. Martinson, *ibid.* **9**, 4367 (1981); A. E. Dieterich, R. Axel, C. R. Cantor, *J. Mol. Biol.* **129**, 589 (1979).
39. A. Worcel, S. Han, M. L. Wong, *Cell* **15**, 969 (1978); V. Jackson and R. Chalkley, *ibid.* **23**, 121 (1981).
40. V. W. Zayetz, S. G. Bavykin, V. L. Karpov, A. D. Mirzabekov, *Nucleic Acids Res.* **9**, 1053 (1981).
41. C. Chothia, A. M. Lesk, G. G. Dodson, D. C. Hodgkin, *Nature (London)* **302**, 500 (1983).

42. This work was supported in part by a grant from the Cell Biology Program of the National Science Foundation (E.N.M.) and a National Institutes of Health grant (W.E.L.). We thank Dr. Gary K. Ackers for numerous discussions during the course of this work and substantive critique of this manuscript. We wish to acknowledge the far-reaching contributions of professors James Bonner and Ru Chih Huang who pioneered the modern era of chromatin biochemistry.

4 February 1985; accepted 21 March 1985