PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Introductory Remarks

A. Klug

Phil. Trans. R. Soc. Lond. B 1978 **283**, 233-239
doi: 10.1098/rstb.1978.0020

Phil. Trans. R. Soc. Lond. B. **283**, 233–239 (1978) [233]
Printed in Great Britain

I. STRUCTURE OF CHROMATIN

Introductory remarks

By A. Klug, F.R.S.
M.R.C. Laboratory of Molecular Biology, Cambridge, U.K.

The DNA of a eukaryotic chromosome is probably one single molecule, amounting to several centimetres in length if laid out straight, and it must be highly folded to make the compact structure one can see in a chromosome. At the same time it is organized into separate genetic or functional units, and the manner in which this folding is achieved, genes organized and their expression controlled is the subject of this Discussion Meeting. The subject is a wide ranging one but there has been such rapid progress in the various branches that the organizers felt it timely to review the current position and to see what connections could be made between structural organization at various levels and functional control.

The first session is devoted to the topic of chromatin, the name given to the chromosomal material when extracted. It consists mainly of DNA, tightly associated with an equal weight of a small set of rather basic proteins called histones. Other proteins and RNA are present in lesser amounts, and these non-histone proteins include polymerases which transcribe the DNA, other enzymes and putative regulatory proteins. For many years the histones themselves appeared rather complex, but by about 1970 (see, for example, DeLange & Smith 1973) it had become clear that there were only five types, the apparent proliferation of types or species arising from various post-synthetic chemical modifications such as acetylation or phosphorylation, which take place during various phases of the cell cycle.

Nucleosomes

The large amounts in which the histones occur suggest that their rôle is structural and it is known that the four histones H2A, H2B, H3 and H4 are responsible for the first level of structural organization in chromosomes (reviewed by Elgin & Weintraub 1975; Kornberg 1977). They fold successive segments of the DNA, about 200 base pairs long, into compact bodies of about 10 nm diameter called nucleosomes. A string of nucleosomes or repeating units is thus created, and when these are closely packed they form a fibre or filament about 10 nm in diameter. The rôle of the fifth histone, H1, is not clear. It is much more variable in sequence than the other four, being not only species-specific but very probably tissue-specific. It is likely that it is concerned with the folding of the nucleofilament into the next highest level of organization.

This is not the place to tell in detail how this clear picture of the basic organization of chromatin has emerged, but the idea of the nucleosome arose from the convergence of several different lines of work. The first indication for a regular structure came from X-ray diffraction studies on chromatin, or nucleohistone as it was then called (Wilkins, Zubay & Wilson 1959; Luzzati & Nicolaieff 1959), which showed that there must be some sort of repeating unit, albeit not very well ordered, on the scale of about 10 nm. The first biochemical evidence for

regularity came from the work of Hewish & Burgoyne (1973) who showed that an endogenous nuclease in rat liver cut the DNA into multiples of a unit size, which they later showed to be about 180–230 base pairs. The work of Noll (1974*a*), using a different enzyme, micrococcal nuclease, gave a value of 205 ± 15. The fact that the nuclease cuts the DNA of chromatin at regularly spaced sites, quite unlike its action on free DNA, is attributed to the fact that the DNA is folded in such a way as to make only these sites available to the enzyme. The third piece of evidence which led to the idea of a nucleosome was the observation that the two 'arginine-rich', highly conserved histones, H3 and H4, existed in solution as a specific oligomer, the tetramer $(H3)_2(H4)_2$, which behaved rather like a common multisubunit globular protein (Kornberg & Thomas 1974). On the basis of these studies, Kornberg (1974) proposed a definite model for the basic unit of chromatin as a 'bead' of about 10 nm diameter, containing a stretch of DNA 200 base pairs long condensed around a protein core made out of eight histone molecules, namely the $(H3)_2(H4)_2$ tetramer and two each of H2A and H2B. The fifth histone, H1, was somehow associated with the outside of each nucleosome. The placing of the DNA on the outside of the bead, perhaps the most surprising feature of the model, was suggested by the presumed globular nature of the tetramer of H3 and H4. The external location of the DNA has been supported by neutron scattering experiments on isolated nucleosomes in solution (Pardon *et al.* 1975; Hjelm *et al.* 1977) which shows that the radius of gyration of the DNA is greater than that of the protein. A summary of this work will be given later by Dr Richards and Dr Bradbury. Evidence that the repeat unit of chromatin contains eight histones comes from chemical cross-linking with bifunctional amino group reagents and analysis of the products by gel electrophoresis (Thomas & Kornberg 1975).

In Kornberg's model, the basic chromatin fibre consists of a flexibly jointed chain of nucleosomes with very little internucleosomal DNA. However, the first electron micrographs to be obtained which showed the particulate nature of chromatin (Olins & Olins 1974; Woodcock 1973) showed linear arrangements of beads, named ν-bodies, separated by long open stretches of DNA and this led some workers to hold the idea of a 'beads-on-a-string' model, in which nucleosomes, containing 120–160 base pairs, were spaced by internucleosomal segments of free DNA about 40–80 base pairs long. It now appears that such electron micrographs represent chromatin in a partly extended form, resulting from shearing forces present during the preparation. When chromatin is prepared by brief micrococcal digestion of nuclei, and which is native in the sense that further nuclease digestion gives the same discrete pattern as does direct digestion of nuclei (Noll, Thomas & Kornberg 1975), this native chromatin shows more or less continuous 10 nm diameter filaments in which the beads are in close contact (Finch, Noll & Kornberg 1975; Finch & Klug 1976). Evidence that this native chromatin has a compact structure with beads in contact also comes from cross-linking experiments (Thomas & Kornberg 1975) and from the measurements of the mass per unit length of a nucleofilament by low angle X-ray scattering (Sperling & Tardieu 1976).

Perhaps the most striking, though not the first, demonstration of the identity between the physical subunit, as determined by sedimentation analysis and electron microscopy, and the biochemical subunit, as determined by chemical composition and nuclease digestion, is given by the so-called 'one, two, three, four' experiment of Finch *et al.* (1975). Noll (1974*a*) had earlier shown that fragments of nuclease-digested chromatin give a series of discrete peaks when sedimented in a sucrose gradient, which corresponded to one, two, three or more repeating units. Finch *et al.* (1975) showed that the material which sediments in, say, the

3-unit position and which contains 600 base pairs of DNA, appears in the electron microscope as three beads (in close contact), and so on. These experiments do not, of course, tell whether all nucleosomes are identical, but there is a strong indication that this is so from antibody binding experiments which show that all nucleosomes contain H2B (Bustin, Goldblatt & Sperling 1976); similar experiments are in progress to test for the other histones.

The idea that the biochemical unit corresponds to the nucleosome was in fact first demonstrated by the electron microscopy of complexes formed *in vitro* between histones and the DNA of adenovirus (Oudet, Gross-Bellard & Chambon 1975) and of Simian virus 40 (Germond *et al.* 1975). These complexes appear as particles or beads on the DNA molecule, separated by stretches of naked DNA. However, the quantity of DNA associated with each particle was estimated to be close to 190 base pairs, in good agreement with the results of nuclease digestion. The SV40 system is a particularly attractive one for studying the basic structure and assembly of chromatin since it has been shown by a number of workers that SV40 DNA and the four histones H2A, H2B, H3 and H4 are associated in infected cells and in virions in a chromatin-like structure called a minichromosome. In particular, viral DNA has been used in reconstitution experiments *in vitro* to demonstrate that all nucleosomes contain the four histones H2A, H2B, H3 and H4 in equimolar amounts (Oudet *et al.* 1975). If one out of the four histones is absent from the reconstitution mixture, no particles are seen in the electron microscope: these reconstitution experiments, and extensions of them which illuminate the rôle of the various histones, will be discussed later by Oudet *et al.* Other reconstitution experiments which lead not to the complete nucleosome but to the 'core particle', will be described by Steinmetz, Streeck & Zachau.

Structure of nucleosomes

There is now a large body of evidence to indicate that nucleosomes exist in all eukaryotes including yeast and fungi, but what has emerged is that the DNA content of the repeating unit is variable, ranging from about 160 to 240 base pairs (reviewed by Kornberg 1977). But the micrococcal nuclease digestion experiments which have demonstrated this variation have also shown that all chromatins investigated have a common feature which bids to become universal. While the action of micrococcal nuclease on chromatin is first to cleave between nucleosomes, it subsequently acts as an exonuclease on the excised nucleosomes, shortening the DNA first to about 160 base pairs, where there is a brief pause in the reaction (Noll & Kornberg 1977; Varshavsky *et al.*, this volume), and then to about 140 base pairs, where there is a clear plateau in the course of digestion, before more degradation occurs. During this last stage the histone H1 is released, leaving a major metastable intermediate in digestion, containing 140 base pairs of DNA complexed with the set of 8 histone molecules. This is called the 'core particle' (Shaw *et al.* 1976) and the remaining DNA, which connects one core particle to the next and which varies in length over cell types from about 20 to 100 base pairs, is called the 'linker DNA'. Since H1 is not present in the core particle there have been suggestions that the variation in the length of the linker DNA is associated with the variation in H1 between cell types, but the situation is clearly more complicated since the experiments on viral nucleosomes by Chambon's group show that all the DNA in the repeat unit is associated with a nucleosome which contains no H1.

Nucleosome core particles may be obtained sufficiently pure and homogeneous enough to

be crystallized. This has been achieved by the Russian group (Bakayev, Melnickov, Osicka & Varshavsky 1975) who have, however, made no crystallographic analysis, and by our research group in Cambridge (L. Lutter, D. Rhodes, J. T. Finch & A. Klug, unpublished). At the present stage our crystals diffract to a resolution of about 2 nm, but the unit cell, which is orthorhombic and of sides $11.0 \times 19.2 \times 34.0$ nm, is formidably large for X-ray analysis. However, we can conclude that the particles pack in hexagonal layers of dimension 11.0 nm, and this has been confirmed by electron microscopy. This value agrees well with the 'hydrodynamic' diameter estimated from the particle weight and sedimentation coefficient (Van Holde *et al.* 1975). Two dimensions of the nucleosome are therefore known, but the third remains to be determined. However, the striking thing one can say already is that the first fold of DNA in chromatin is regular enough to lead to particles homogeneous enough to be crystallized.

It is to be hoped that the X-ray analysis will establish just how the DNA is folded. We already know that at the level of the double helix the DNA cannot depart much from the normal B, or wet, form, since digestion of chromatin by the enzyme DNase I produces a set of single stranded DNA fragments differing in size by 10 bases (Noll 1974*b*). This can only mean that the DNA is folded in some very regular way and moreover that it is very probably on the outside of the bead. The simplest explanation of the limited cleavage by DNase I is that it is recognizing some structural feature which repeats with the periodicity of the DNA, for example, the point on each strand furthest from the bead surface, or one of the grooves. Another possible explanation is that there are sharp bends or kinks (Crick & Klug 1975; Sobell *et al.* 1976) at intervals of ten base pairs in the DNA, and the nucleases cleave preferentially at or between kinks. Kinks were postulated to explain how the DNA of the nucleosome of length 68 nm could be folded without undue strain into a compact structure of maximum diameter about 10 nm. Sobell, Tsai, Jain & Sakore will later describe a particularly attractive type of kink, but it remains to be seen whether the folding of the DNA on the nucleosome is indeed accomplished in this way or by smooth bending.

Higher levels of organization

So far, I have discussed only the first level of folding, the basic fibre or nucleofilament composed of a chain of nucleosomes. When these are closely packed along the filament, as we believe to be the case in native interphase chromatin, the DNA is compacted by about a factor of 7, a value very small with the factor believed to apply to interphase chromosomes, let alone metaphase chromosomes for which the factor may approach several thousand. I would like briefly to mention some experiments we have carried out in Cambridge (Finch & Klug 1976) which suggest that the next level of organization after the nucleofilament may be achieved by coiling this into a helix, or what would be called a supercoil if the word had not already been used in a different context. Because the turns of the superhelix are spaced fairly closely together, we call it a solenoid.

It had been known for some time that the thickness of fibres observed in electron microscopical studies of 'whole mount' chromosome specimens varied from about 10–25 nm in diameter, depending on whether chelating agents had been used or not in the preparation (Ris & Kubai 1970). Taking this as a clue, Finch and I carried out some experiments *in vitro* on short pieces of chromatin, prepared by brief micrococcal digestion of nuclei and therefore native

in the sense described earlier. In the presence of chelating agents this native chromatin appears as fairly uniform filaments of 10.0 nm diameter, but when Mg^{2+} ions are added, these coil up into thicker fibres of about 30 nm diameter, which are transversely striated at intervals of about 12–15 nm, corresponding apparently to the turns of a solenoidal structure. The thick fibres are rather knobbly so that the solenoids are distorted, but not in an obviously regular manner.

Moreover, we found that when the same experiments were carried out on H1 depleted chromatin, the thicker fibres failed to form, suggesting that the fifth histone, H1, is needed for the formation or stabilization of the solenoidal structure. H1 is indeed a good candidate for involvement in higher order folding since it is somehow on the 'outside' of the nucleosome (Kornberg 1974), it is the most variable in its amino acid composition from one source to another, and there is some evidence that it is involved in the 'condensation' of chromatin (see Egin & Weintraub 1975).

These experiments *in vitro* in solution, carried out without control of the non-histone protein complement of the chromatin, do not demonstrate that solenoidal structures actually exist in chromosomes, but can only be regarded as suggestive. However, solenoidal structures might well correspond to the 'superunit threads' seen by Davies and his colleagues (summarized by Davies & Haynes 1975) in condensed interphase chromosomes (chromatin bodies) from several tissues. If solenoids and superunit threads can be equated and if the distance between nucleosomes along a nucleofilament is about 10 nm, then there would be about six nucleosomes per turn in the solenoid. This gives a compaction ratio for the DNA of about 40:1. It is, however, very possible that in ordinary interphase chromatin there might be a family of structures with different numbers of nucleosomes per turn and that chromatin might be a mixture of these. One might speculate that the next order of folding may involve solenoids folded into loops, and here one begins to approach the scale of gene organization (figure 1). Taking a typical genetic unit to contain about $50–100 \times 10^3$ base pairs, i.e. 250–500 nucleosomes, this could be accommodated in a solenoid about 500–1000 nm long, so that if the solenoids are curved, looping out from the chromosome axis, the thickness of the chromosome would be about 0.2–0.5 μm, a very reasonable value.

Independent, folding domains consisting of 'supercoiled DNA loops' were first described in the *Escherichia coli* chromosome (Worcel & Burgi 1972) and have now been shown to exist in the DNAs of *Drosophila* (Benyajati & Worcel 1976) and of cultured human cells (Cook & Brazell 1975). This is an important step towards understanding DNA packing in chromosomes and Worcel & Benyajati will tell us about it in another session. The loops in *Drosophila* contain on the average about 400 nucleosomes, or 80 000 base pairs, but the nature of the loop stabilizers, whether RNA or protein or both is not known. Similarly it is not known what relation these DNA domains bear to cytological units or to the units of genetic expression, but they are of the same order of magnitude as a band of a *Drosophila* polytene chromosome, or strictly to one chromoneme of that band. These correspondences are tantalizing and Crick (1976) has been moved to coin the slogan 'one band–one loop–one solenoid'. Solenoids would be present in inactive genes, but would be opened up into filaments in active genes (figure 1). Indeed Crick has suggested that a solenoid might be activated by transcribing it, starting from a promoter at its base. It is not going to be easy to test ideas like these, but there are more immediate questions which it might be possible to answer soon, for example whether the nucleosomal repeats are maintained in regions of chromatin being actively transcribed. These matters will be discussed

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

by Gottesfeld and Franke & Scheer in other sessions. That a question like this relating structural to functional aspects can be discussed at all is a measure of the progress I referred to at the beginning, and illustrates the hope that the organizers have had for this meeting.

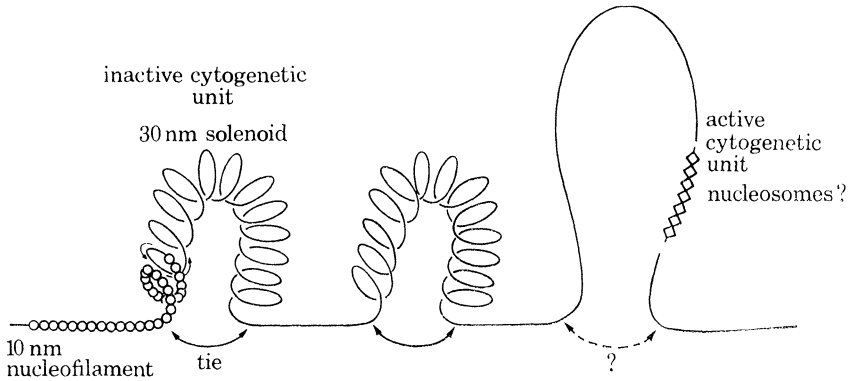

FIGURE 1. Highly schematic diagram of possible levels of organization in chromatin. The representation of the transcriptionally active loop is based upon studies of lampbrush chromosomes.

Note added in proof (17 January 1978)

The crystals of the 140 base-pair nucleosome cores described above have now been analysed by X-ray diffraction and electron microscopy (Finch, J. T., Lutter, L. C., Rhodes, D., Brown, R. S., Rushton, B., Levitt, M. & Klug, A. 1977 *Nature, Lond.* **269**, 29–36). The core is a flat particle of dimensions about $11 \times 11 \times 5.5$ nm, somewhat wedge-shaped and strongly divided into two layers, consistent with a model in which the DNA is wound into something less than two turns of a flat superhelix of pitch about 2.8 nm. There are about 80 bases in each turn of the superhelix, so that the nucleosome core contains $1\frac{3}{4}$ turns. This model for the organization of DNA in a nucleosome core provides an explanation for the results of certain enzyme digestion studies on nucleosomes. As mentioned above, the action of DNase I on chromatin or nucleosomes is to produce a set of single stranded DNA fragments differing in size by 10 bases. The enzyme does not, however, cut equally frequently every 10 bases, and further work in a number of laboratories, including our own, has revealed the distribution of frequency of cutting among the susceptible sites. The pattern is complicated, and its interpretation was not clear until considered in relation to the superstructure of DNA determined by the X-ray analysis. Thus it was noted that the least susceptible sites are 80 bases apart. Since points along the double helix 80 bases apart will be brought close together after one turn of the DNA superhelix, such points will have similar environments and therefore be expected to have similar susceptibilities to the nuclease, as is observed.

Another recent advance, which bears on the question touched on above of the nature of the folding domains of DNA in chromosomes, has been made by Laemmli and his colleagues at Princeton (Adolph, K. W., Cheng, S. M. & Laemmli, U. K. 1977 *Cell* **12**, 805–816; Paulson, J. R. & Laemmli, U. K. 1977 *Cell* **12**, 817–828). They have studied histone-depleted metaphase chromosomes and show that, after removal of the histones, the DNA retains a highly folded configuration organized by a certain set of non-histone proteins, which they call 'scaffolding proteins'. As seen in the electron microscope, the structure consists of a central fibrous scaffold, which, apart from an increase in length, retains the characteristic morphology of metaphase chromosomes, and this scaffold is surrounded by a widespread 'halo' of DNA released laterally from the chromatid axes. The DNA is clearly attached to the scaffold, and in a few favourable

electron micrographs it is possible to trace loops which leave the scaffold and return to an adjacent point, indeed rather like the schematic diagram in figure 1. The loops are at least 10–30 nm long, corresponding to $30–90 \times 10^3$ bases, that is, in the same range which had been inferred from the physiocochemical studies of eukaryotic interphase nuclei referred to above. If the two sets of values are indeed as close in size as this, this would be a good indication that the folding domains of DNA present in interphase chromosomes might be maintained, though more condensed overall, during metaphase.

References (Klug)

Bakayev, V. V., Melnickov, A. A., Osicka, V. D. & Varshavsky, A. J. 1975 *Nucl. Acids Res.* **2**, 1401–1419.
Benyajati, C. & Worcel, A. 1976 *Cell* **9**, 393–407.
Bustin, M., Goldblatt, D. & Sperling, R. 1976 *Cell* **7**, 297–304.
Cook, P. R. & Brazell, I. A. 1975 *J. Cell Sci.* **19**, 261–279.
Crick, F. H. C. 1976 In *Organisation and expression of chromosomes* (eds V. G. Allfrey, E. K. F. Bautz, B. J. McCarthy R. T. Schimke & A. Tissières), p. 71. Berlin: Dahlem Konferenzen.
Crick, F. H. C. & Klug, A. 1975 *Nature, Lond.* **255**, 530–533.
Davies, H. G. & Haynes, M. E. 1975 *J. Cell Sci.* **17**, 263–285.
DeLange, R. J. & Smith, E. L. 1973 *Acc. Chem. Res.* **5**, 368–373.
Elgin, S. C. R. & Weintraub, H. 1975 *A. Rev. Biochem.* **44**, 725–774.
Finch, J. T., Noll, M. & Kornberg, R. D. 1975 *Proc. natn. Acad. Sci. U.S.A.* **72**, 3320–3322.
Finch, J. T. & Klug A. 1976 *Proc. natn. Acad. Sci. U.S.A.* **73**, 1897–1901.
Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. 1975 *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843–1847.
Hewish, D. R. & Burgoyne, L. A. 1973 *Biochem. biophys. Res. Commun.* **52**, 504–510.
Hjelm, R. P., Kneale, G. G., Suau, P., Baldwin, J. P. & Bradbury, E. M. 1977 *Cell* **10**, 139–151.
Kornberg, R. D. 1974 *Science, N.Y.* **184**, 868–871.
Kornberg, R. D. 1977 *A. Rev. Biochem.* **46**, 931–954.
Kornberg, R. D. & Thomas, J. O. 1974 *Science, N.Y.* **184**, 865–868.
Luzzati, V. & Nicolaieff, A. 1959 *J. molec. Biol.* **1**, 127–133.
Noll, M. 1974*a* *Nature, Lond.* **251**, 249–251.
Noll, M. 1974*b* *Nucl. Acids Res.* **1**, 1573–1478.
Noll, M., Thomas, J. O. & Kornberg, R. D. 1975 *Science, N.Y.* **187**, 1203–1206.
Noll, M. & Kornberg, R. D. 1977 *J. molec. Biol.* **109**, 393–404.
Olins, A. L. & Olins, D. E. 1974 *Science, N.Y.* **183**, 330–332.
Oudet, P., Gross-Bellard, M. & Chambon, P. 1975 *Cell* **4**, 281–300.
Pardon, J. F., Worcester, D. L., Wooley, J. C., Tatchell, K., Van Holde, K. E. & Richards, B. M. 1975 *Nucl. Acids Res.* **2**, 2163–2176.
Ris, H. & Kubai, D. F. 1970 *A. Rev. Genet.* **4**, 263–294.
Shaw, B., Herman, T., Kovacic, R., Beaudreau, G. & Van Holde, K. E. 1976 *Proc. naln. Acad. Sci. U.S.A.* **73**, 505–509.
Sobell, H., Tsai, C., Gilbert, S., Jain, S. & Sakore, T. 1976 *Proc. natn. Acad. Sci. U.S.A.* **73**, 3068–3072.
Sperling, L. & Tardieu, A. 1976 *FEBS Lett.* **64**, 89–91.
Thomas, J. O. & Kornberg, R. D. 1975 *Proc. natn. Acad. Sci. U.S.A.* **72**, 2626–2630.
Van Holde, K. E., Shaw, B. R., Lohr, D., Herman, T. M. & Kovacic, R. T. 1975 *FEBS Symp.* **38**, 57–72.
Wilkins, M. H. F., Zubay, G. & Wilson, H. R. 1959 *J. molec. Biol.* **1**, 179–185.
Woodcock, C. L. F. 1973 *J. Cell Biol.* **59**, 368a.
Worcel, A. & Burgi, E. 1972 *J. molec. Biol.* **71**, 127–147.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES