PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The Bending of DNA in Nucleosomes and Its Wider Implications

A. A. Travers and A. Klug

Phil. Trans. R. Soc. Lond. B 1987 **317**, 537-561
doi: 10.1098/rstb.1987.0080

Phil. Trans. R. Soc. Lond. B **317**, 537–561 (1987)
Printed in Great Britain

The bending of DNA in nucleosomes and its wider implications

By A. A. Travers and A. Klug, F.R.S.
Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

The DNA of a nucleosome core particle is wrapped tightly around a histone octamer with approximately 80 base pairs per superhelical turn. Studies of both naturally occurring and reconstituted systems have shown that DNA sequences very often adopt well-defined locations with respect to the octamer. Recent work in this laboratory has provided a structural explanation for this sequence-dependent positioning in terms of the differential flexibility of different sequences and of departures from smooth bending.

The 'rules' that are emerging for DNA bendability and, from the results of other workers, on intrinsically bent DNA, are likely to be useful in considering looping and bending of DNA in other processes in which it is thought to be wrapped around a protein core.

Introduction

When DNA is packaged, whether in a phage head or a eukaryotic chromosome, it is often tightly bent. In chromatin the elementary unit of the structure is the nucleosome, in which the DNA wraps about two times around an octamer of the histone proteins H3, H4, H2A and H2B as a left-handed superhelix with a diameter of 86Å† (Finch *et al.* 1977; Richmond *et al.* 1984). The path of the DNA between histone octamers is not known, but it appears that variable lengths of 'linker' DNA separate individual nucleosomes (Prunell & Kornberg 1982; Widom & Klug 1985). A single copy of the fifth histone, H1, is associated with the linker in a way that is not fully understood. When the linker DNA is digested away by nucleases, the H1 drops off leaving the so-called nucleosome core particle, consisting of about 145 base pairs (b.p.) wrapped in approximately 1.8 superhelical turns around the histone octamer. The two copies of each histone in the octamer are related by a unique axis of twofold symmetry, or dyad, which at one end passes through the midpoint of DNA. Although in the chromosome the histone octamers are associated with a great variety of DNA sequences, studies of nucleosome positioning in both reconstituted and naturally occurring systems have shown that these proteins can adopt well-defined, even precise, locations with respect to the primary DNA sequence (Simpson & Stafford 1983; Edwards & Firtel 1984; Palen & Cech 1984; Ramsay *et al.* 1984; Rhodes 1985; Drew & Travers 1985*a*; Thoma & Simpson 1985; Ramsay 1986).

What are the molecular interactions that determine this precise positioning? The association of histone octamers with an immense variety of DNA sequences suggests that sequence-specific recognition, as classically exemplified by the interaction of the *lac* and λ repressors with their respective operators, is not a dominant determinant of nucleosome positioning. However, it has recently become apparent that the structural and mechanical properties of the DNA double helix vary in a sequence-dependent manner and could contribute significantly to the specific

† 1 Å = 10^{-10} m = 10^{-1} nm.

interaction of many proteins with DNA (Lomonossoff *et al.* 1981; Drew & Travers 1984). In this respect DNA is not merely a passive ligand in binding to a protein but its physical properties can actively influence the structure and stability of the complexes.

The property of DNA that is relevant to complexes in which it does not follow a straight path is the bendability of the molecule. DNA does not behave as an isotropic rod (Trifonov & Sussman 1980; Widom 1985; Calladine & Drew 1986); it may bend more easily in one plane rather than another, i.e. it has anisotropic flexibility. In a long molecule this will depend on the overall effect of the distribution of short sequences within it that are differentially flexible in different directions. In many large DNA–protein complexes, including those involved in the enzymatic manipulation of DNA in replication and recombination (Echols 1986), the DNA molecule is wrapped in a nucleosome-like manner tightly around a central core of protein molecules. In such structures all the grooves (both major and minor) on the inside of the curve must narrow somewhat because of the compression associated with bending, whereas those on the outside of the curve become correspondingly wider. It is therefore clear that in such a DNA–protein complex the structure of the DNA must be able to accommodate this deformation. There are also sequences that, even in the absence of external forces imposed by interaction with proteins, impart a preferred direction of curvature on a DNA molecule, i.e. give an 'intrinsic bend' (Marini *et al.* 1982; Hagerman 1984; Wu & Crothers 1984). We shall discuss both these kinds of bending, and try to draw a relation between them.

For the subsequent discussion we need to state the terminology we shall use to describe the positioning of DNA on a nucleosome or similar complex. Two parameters must be considered: a translation, marking where the histone octamer is placed along the DNA; and a rotation, which defines the local orientation (azimuth) of the DNA relative to the direction of curvature. The DNA sequences that determine nucleosome positioning must also reflect these parameters. In a DNA molecule that is wrapped uniformly around a protein any DNA sequences that are selected to correlate with particular rotational orientations should occur regularly with a periodicity equal to the local helical twist (Drew & Travers 1985*a*). By contrast, translational sequence markers would be expected to occur non-periodically and be found either at unique locations within the nucleosome, such as the dyad, or at positions symmetrically related about the dyad.

Use of DNase I as a probe for DNA organization

When a DNA molecule is constrained either by binding to a surface or by circularization into a small circle, its three-dimensional organization may restrict the accessibility of a chemical reagent or enzyme relative to that in the unconstrained linear state. A carefully selected reagent can thus act as a probe for the configuration of the DNA molecule. The utility of such a probe depends on both its chemical specificity and the dependence of reactivity on the accessibility of the DNA substrate. One reagent that has been widely used in previous studies is the enzyme DNase I (Lutter 1978; Rhodes & Klug 1980). This enzyme has a broad chemical specificity for the cleavage of phosphodiester bonds in DNA (Drew 1984; Drew & Travers 1985*a*) and thus can be used to investigate gross structural features without the complication of cleavage being restricted by high sequence specificity. In addition, the enzyme protein is large relative to the DNA double helix (Suck *et al.* 1984; Suck & Oefner 1986) and consequently when a DNA molecule is lying on a surface access of the enzyme to one side of the DNA helix is

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

hindered. In this situation only the most exposed phosphodiester bonds will be accessible to cleavage (figure 1). Because DNase I cleaves the phosphate backbone independently on either side of the minor grooves (Lutter 1978; Drew & Travers 1984; Suck & Oefner 1986), the pattern of cuts reveals the orientation of the DNA molecule relative to the surface on which it lies.

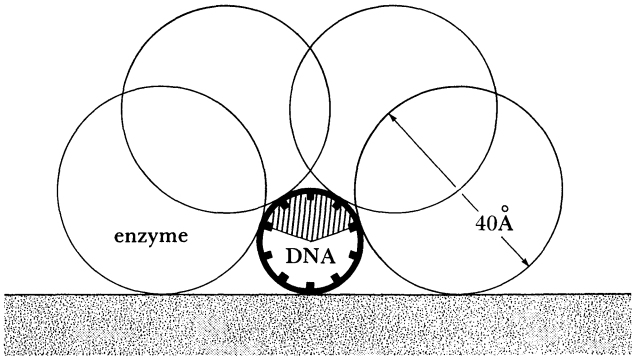

FIGURE 1. Schematic drawing of the accessibility of an enzyme to DNA lying on a flat surface. The surface restricts a spherical enzyme with a diameter of 40 Å to a maximum of four accessible phosphodiester bonds (redrawn from Rhodes & Klug (1980)).

Although not directly relevant to the theme of this paper, it should be noted that DNase I cleavage patterns can also be used to detect local variations in helical structure as well as the orientation of the double helix. This follows because in DNase I cleavage profiles there is a stagger between the cuts on the two strands. For classical B-form DNA this stagger is on average about two to three nucleotides in the 3′ direction. However, the stagger is dependent on the distance between the two DNA strands and hence on the width of the minor groove. This means that A-form DNA, which has a wider minor groove, a high tilt and a low average twist angle, would have a stagger of 0–1 b.p. (Rhodes & Klug 1986) whereas a DNA with a greater-than-average twist angle and hence a narrow minor groove (Drew & Travers 1984) would have a stagger of 3–4 b.p.

THE BENDING PREFERENCE OF DNA IS A DETERMINANT OF NUCLEOSOME POSITIONING

The tight wrapping of DNA about the histone octamer strongly suggests that the anisotropic bendability of DNA may be a major determinant of nucleosome positioning. To test this hypothesis experimentally we need to establish that the bending preference of a particular DNA molecule is the same when reconstituted into a nucleosome core particle as it is when constrained in the absence of any bound proteins. To approach this question a 169 b.p. DNA fragment of bacterial origin (which had thus undergone no natural selection for binding to a histone octamer) was covalently closed into a small, relaxed circle (Drew & Travers 1985*a*). Such a molecule, if bent uniformly, would have an average diameter of *ca.* 170Å, that is, about twice that of the superhelical turns on the nucleosome core. The problem is thus to determine by experiment whether the DNA in such a circle assumes a preferred direction of curvature, and if it does, to establish the rotational orientation of a particular DNA sequence relative to the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

inside and outside of the circle. The probe used (see previous paragraphs) was the nuclease DNase I, which has an effective average diameter of 40Å (Suck *et al.* 1984), chosen on the assumption that enzymic access to the inside of the circle would be kinetically impeded relative to that on the outside. The experimental result, which shows this assumption to be correct, was that cuts on the circle follow a sinusoidal periodicity of 10.56 b.p. on both DNA strands, in contrast to the irregular cleavage pattern on the corresponding linear molecule. This shows immediately that in a small circle the DNA assumes a highly preferred configuration. The average cutting periodicity found has the value expected from the helical periodicity of mixed sequence DNA, because the overall length of the molecule (169 b.p.) was carefully chosen so as to complete a precisely integral number of double helical turns (16×10.56), thereby producing efficient closure during ligation (Shore & Baldwin 1983; Horowitz & Wang 1984).

Because DNase I cleaves DNA where the minor groove is accessible to the enzyme (Lutter 1978; Drew 1984; Drew & Travers 1985*a*; Suck & Oefner 1986) we can also deduce the angular orientation of the DNA sequence in the small circle. The general result is that (A + T)-rich sequences are cleaved at a substantially lower rate in the circle than in the linear form, whereas cleavage at (G + C)-rich sequences is not significantly reduced by circularization. This means that, on average, (A + T)-rich minor grooves face in towards the centre of the circle whereas (G + C)-rich minor grooves face out. Of course, in any given piece of mixed-sequence DNA, it is unlikely that all the bending preferences in that sequence can be simultaneously satisfied by the configuration assumed by the whole molecule, and the overall setting will be determined by the balance of local preferences. Consequently, there will always be a few helix segments whose rotational position is imposed not by local constraints but by the preferred configuration of the whole molecule. At some sites where the DNA is unfavourably positioned in this way, the rate of DNase I cleavage in the circle is greater than that on the linear molecule, suggesting a local deformation of the DNA structure. This effect is particularly apparent where (A + T)-rich minor grooves face outwards rather than inwards.

This experiment demonstrates that a short DNA molecule, when constrained in the absence of protein, adopts a highly preferred rotational orientation. Is this preference maintained when the DNA fragment, now in a linear form, is placed on a histone octamer? Analysis of such a reconstituted nucleosome core particle showed that the angular setting of the DNA remained largely conserved in going from circle to nucleosome. The few slight differences that are observed may be attributed to a reduction in the average cutting periodicity from 10.56 b.p. in the circle to 10.32 b.p. upon nucleosome formation. As argued previously (Klug & Lutter 1981) this latter periodicity sets an upper bound to the average local twist of the DNA in the core particle.

Sequence determinants of DNA bending: statistical sequencing of nucleosome core DNA

If the rotational positioning of DNA on a nucleosome is influenced by particular sequences, we would expect that, in a population of nucleosome core particles, the occurrence of such sequences would exhibit a periodic modulation that would reflect the structural periodicity of the DNA molecule lying on the surface of the histone octamer. To determine the general nature of such DNA sequences, a population of DNA molecules was extracted from a nucleosome core preparation isolated from chicken erythrocytes. This population had essentially the same

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

dinucleotide composition as that previously determined for total chicken erythrocyte DNA (Swartz *et al.* 1962) and by this criterion was therefore representative of chicken DNA as a whole. The analysis of sequence content was done by a technique we have termed 'statistical sequencing' (Drew & Travers 1985*a*). In this method (figure 2) the predominant locations of particular DNA sequences in a mixed population of DNA molecules are detected by binding

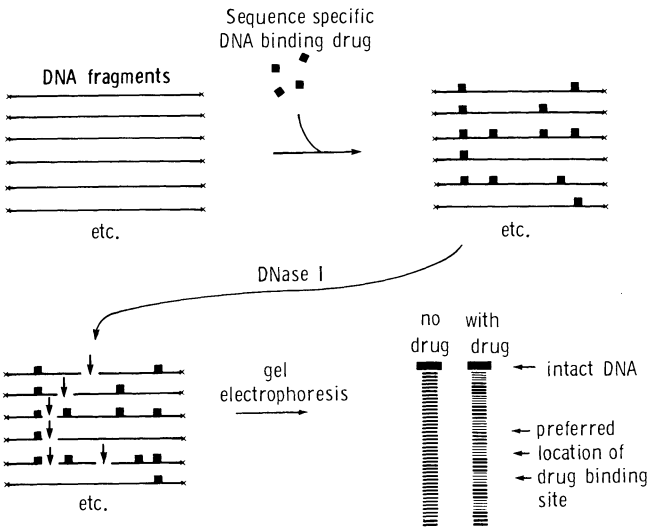

FIGURE 2. Flow diagram for statistical sequencing. The figure shows the method applied to an aligned population of short DNA molecules of uniform length.

antibiotic drugs of known sequence specificity to the DNA and then treating the drug–DNA complex with DNase I. At sites where the drug binds, the rate of DNase I cleavage is substantially reduced and can be measured. By this method of quantitative footprinting, any preferred location of such sites along the length of DNA can be determined.

Because it had been shown that the angular orientation of a particular DNA sequence (the *Escherichia coli tyrT* promoter) correlated with the disposition of (A+T)-rich and (G+C)-rich stretches, the antibiotic drugs chosen for the statistical sequencing of core DNA isolated from chicken erythrocytes were distamycin and chromomycin. The former binds selectively to runs of four or more (A, T) base pairs (Van Dyke *et al.* 1982; Fox & Waring 1984) whereas the latter has the reverse specificity, selecting runs of four or more (G, C) base pairs (Van Dyke & Dervan 1983; Fox & Howarth 1985). This analysis showed that on the core DNA the occurrence of short runs of both (A+T) and (G+C) are periodically modulated with an average period of 10.17 ± 0.05 b.p. Also, these two modulations are in opposite phases, a maximum in the occurrence of (A+T) runs coinciding with a minimum in the occurrence of (G+C) runs, and vice versa. This periodic fluctuation in sequence content is thus not primarily in single-nucleotide composition but instead relates to the orientation of short runs of (A+T) or of (G+C) relative to the histone proteins. Because it had been shown that the minor groove points out of the particle dyad (Lutter 1978; Richmond *et al.* 1984), a position that corresponds to the midpoint of the population of DNA molecules analysed, the angular disposition of such runs could be deduced. As on nucleosome core particles reconstituted with defined DNA sequences, (A+T) runs are preferentially placed where the minor groove faces approximately

inward towards the histone octamer whereas (G + C) runs prefer to occupy positions where the minor groove points outward. This kind of sequence variation argues strongly that the rotational orientation of DNA within the nucleosome is determined principally by certain directional bending preferences of the DNA, rather than by any sequence-specified protein–DNA contacts.

Sequence determinants of DNA bending: direct sequencing of nucleosome core DNA

Statistical sequencing allows a description of the most dominant sequence features of nucleosome core DNA, but it cannot describe the occurrence of all particular sequence combinations. Nor can it assess the detailed nature of helix curvature in regions such as those close to the protein dyad where the path of the DNA deviates markedly from a uniform superhelix (Richmond *et al.* 1984). To investigate these aspects of nucleosome core structure, 177 individual DNA molecules from the same DNA sample that was used for statistical sequencing were cloned and sequenced (Satchwell *et al.* 1986). These sequences were aligned about their midpoint and analysed for the occurrence of dinucleotides and trinucleotide. The most striking result from this analysis is a well-defined periodicity in the distribution of the dinucleotide ApA/TpT between base pair steps 1 and 56, where the dyad is defined as step 72.5 at the midpoint of the observed average DNA length of 145 bases (figure 3). These maxima of ApA/TpT occurrence are on average sited where the crystal structure shows the minor groove to point in towards the histone octamer (Richmond *et al.* 1984). In this region

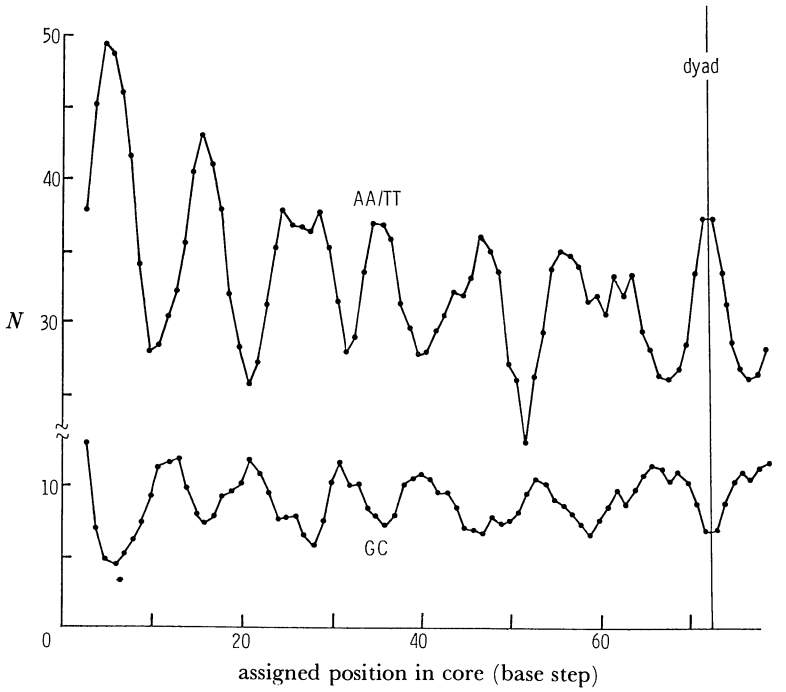

Figure 3. Variations in the occurrence of the dinucleotides AA/TT and GC versus position in the nucleosome core DNA sequence. *N* is the running 3-bond average of the frequency of occurrence of a dinucleotide also averaged about the dyad at base-step 72.5. Data are taken from Satchwell *et al.* (1986).

there are five marked peaks at an average spacing of 10.1 bases. This regular periodic pattern changes between position 56 and the dyad, such that a maximum occurs at the dyad instead of the minimum which would have been expected had the previous pattern simply continued. This reversal of phase corresponds to a region of the nucleosome where the path of the DNA departs from a uniform superhelix with a 'jog' (figure 4*a*) between the two adjacent turns of the supercoil (Richmond *et al.* 1984).

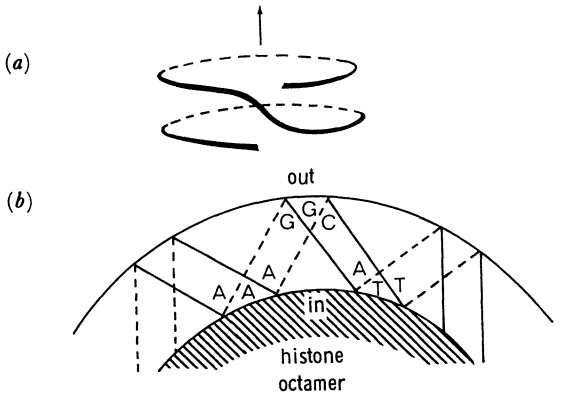

FIGURE 4.(*a*) The path of the DNA in the nucleosome core particle. The path is represented as a left-handed 'key ring'; for clarity the jog at the dyad is exaggerated in comparison with the actual structure (Richmond *et al.* 1984). (*b*) Preferred rotational orientation of trinucleotide sequences in the nucleosome core. The schematic drawing shows the relation of the sequences with respect to the minor groove of DNA.

To obtain a quantitative measure of amplitude and phase for each of the ten possible dinucleotide steps, their distributions as a function of position were analysed by Fourier transformation. This analysis confirmed the rotational orientation of the dinucleotide ApA/TpT and showed that its fractional variation in amplitude is $\pm 20\,\%$ about the mean. Of the other dinucleotides with significant periodic modulations, GpC is the strongest and is found in the opposite phase to ApA/TpT with a fractional variation in amplitude of $\pm 27\,\%$. Similarly the dinucleotides GpC/CpC and TpG/CpA are in phase with GpC but their preferences for outward facing minor groove are smaller, having variations in amplitude of $\pm 12\,\%$ and $\pm 8\,\%$ respectively (figure 3).

It is clear from the crystal structures of short DNA fragments that the conformation at a particular base step depends not only on the base pairs on either side of it but also on the sequence context in which it is located. We should thus ideally consider at least the flanking base steps on either side of a particular base step, that is, tetranucleotide sequences. However, the number of core DNA sequences available was insufficient to allow a reliable statistical evaluation of tetranucleotide occurrences and therefore only the occurrences of the 32 possible trinucleotides were analysed by Fourier transformation (taking complementary trinucleotides to be equivalent). This analysis showed clearly that the positioning of some sequences on the nucleosome core was determined by a trinucleotide or longer sequence, rather than by the dinucleotide components acting independently. Thus of the seven trinucleotides containing the dinucleotide ApA/TpT, only three, ApApA/TpTpT, ApApT/ApTpT and TpApA/ApTpT, show significant relative preferences for an inward-facing minor groove with respective amplitudes of $\pm 36\,\%$, $30\,\%$ and $20\,\%$ about their mean values. Indeed, the dinucleotide ApA/

TpT in isolation (that is, the sequence ApA not flanked by A on either side and the sequence TpT not flanked by T) does not exhibit any significant periodic modulation. Again the conclusion is that for this class of sequences, rotational positioning is determined by at least a trinucleotide sequence. By contrast, the periodic modulations of the GpC containing trinucleotides are all the same phase but differ substantially in amplitude.

The periodicity of nucleosome core DNA and the linking-number paradox

When the results of the X-ray crystal analysis on the organization of DNA in the nucleosome core were first obtained and compared with physicochemical data a paradox emerged that has come to be known as the linking-number problem (Finch *et al.* 1977). The X-ray analysis showed that the double helix of DNA in a nucleosome is wound into about two turns of a shallow superhelix, whereas measurements on closed circular DNA extracted from SV40 chromatin gave a linking number nearer to one per nucleosome (Germond *et al.* 1975), rather than the value of two which would be expected if the helical screw of the DNA remained the same. This subject is discussed by Finch *et al.* (1977) and Klug & Lutter (1981) who proposed that the explanation of the paradox lay in a difference in the screw of the DNA when in the nucleosome and free in solution.

The new results on the periodic modulation of sequence in nucleosome core DNA have important implications with regard both to the change in linking number on nucleosome formation and also to the packing of the two adjacent superhelical turns of DNA on the histone octamer. The main fact is that the periodicity in sequence of nucleosome DNA is different from helical periodicity of DNA in solution (Drew & Travers 1985*a*; Satchwell *et al.* 1986). By the method of statistical sequencing, Drew & Travers (1985*a*) obtained a value of 10.17 ($\pm$0.05) b.p. for the modulation of sequence content within a mixed population of chicken core DNA molecules. A similar value of 10.21 b.p. was obtained by direct sequence analysis of a small sample of this same population (Satchwell *et al.* 1986). In both cases the value for the periodicity is an average calculated across the 120–140 b.p. centred on the nucleosome dyad. To relate the sequence periodicity of core DNA to the problem of the linking number in the path of the DNA about the histone octamer we must first consider how the sequence periodicity is related to the helical (i.e. structural) periodicity of the DNA molecule. For the discussion that follows, we define the local twist as the instantaneous torsion about a moving axis tangential to the space-curve of the molecule (Crick 1953, 1976). In supercoil this is not the same function as the angle between two successive base pairs in a laboratory (i.e. external) frame of reference.

If a particular repeated sequence is to be used over and over again for preferred bending at equivalent positions on a uniform superhelical path, then the structural periodicity in the local frame of reference, i.e. the local twist, must match the sequence periodicity. This conclusion remains true for a non-uniform superhelical path, as on the nucleosome, where the average sequence periodicity will correspond to the average local twist, provided that between positions of identical sequence phase at each extremity of the periodically modulated sequence there is no net change in the direction of curvature relative to the average superhelical axis. (In this context the direction of maximum curvature is given by the normal in the osculating plane, defined as containing in the limit two successive tangents to the curve, or three successive points

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

(see, for example, Salmon & Rogers (1914).)) If we assume that this condition is satisfied, then the average local twist on the nucleosome core would be 10.17 b.p. per turn. Note that in the laboratory frame of reference this corresponds to a structural periodicity of 10.31 b.p. per turn, a number that is not relevant to the calculation shown here.

The consequences for the change in linking number when DNA passes from solution, where its helical periodicity is 10.6 b.p. per turn (Wang 1979; Rhodes & Klug 1980; Peck & Wang 1981) are as follows. The change in linkage number is equal to the number of superhelical turns if there is no change in the screw, as expressed in the local frame of reference (Finch *et al.* 1977). If the screw changes this gives an extra contribution because the change of twist is not accounted for by the superhelical path. These contributions are additive provided the calculation is done in a local frame of reference (Fuller 1978). Then, in this case, for one superhelical turn, which contains 7.6 turns of the double helix (Richmond *et al.* 1984) and thus 77.3 b.p. ($= 7.6 \times 10.17$),

$$\Delta L = -1 + [(77.3/10.17) - (77.3/10.6)],$$

hence

$$\Delta L = -1 + 0.31 = -0.69,$$

where ΔL is the change in linking number, -1 is due to the superhelical path and 0.31 is due to the change in local twist.

For 1.8 superhelical turns on the nucleosome core,

$$\Delta L = -0.69 \times 1.8 = -1.24.$$

This result for ΔL can be derived in a different way by an argument due to J. T. Finch. When the superhelical path is pulled into an almost straight line the writhe is reduced to almost zero and all the linkage, which is conserved, is converted into twist. In this operation the number of base pairs per turn in the local frame of reference is also conserved. Hence for one superhelical turn on the nucleosome the actual linking number is

$$L_{\text{nucleosome}} = -1 + (77.3/10.17).$$

Now the linking number of DNA in solution is simply calculated from the twist. Hence for 77.3 b.p.,

$$L_{\text{solution}} = (77.3/10.6).$$

Hence

$$\Delta L = [-1 + (77.3/10.17)] - (77.3/10.6),$$

which is the same result as above. In this method one is simply calculating the change in linkage between two, as it were, straight pieces of DNA so only changes in twist are involved.

Both methods of calculation demonstrate the essential point that the change in helical periodicity makes a substantial contribution to the change in linkage number, as previously predicted (Finch *et al.* 1977; Klug & Lutter 1981). The value of ΔL of -1.24 compares with the experimentally determined value of *ca.* -1 per nucleosome observed for nucleosome arrays containing four or more nucleosome core particles (Germond *et al.* 1975; Simpson *et al.* 1985). The change in linkage number consequent upon formation of a single nucleosome is undetermined so we do no know whether this small discrepancy is due to a contribution from linker DNA or to a more complex topology for the path of nucleosome core DNA than we have assumed.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

It is possible that the average value of 10.17 b.p. per turn for the local twist of DNA in the nucleosome core particle obscures local variations in helical periodicity. Indeed when the periodically modulated sequences in the region between 3 and 58 b.p. and 88 and 143 b.p. from the dyad are analysed independently, the average periodicity of the sequence modulations of the trinucleotides with the greatest periodic modulation is 10.02 b.p. (B. F. Luisi, S. C. Satchwell & A. A. Travers, unpublished observations). This means that, to maintain an overall average periodicity of 10.17 b.p., the average local twist of the central two double helical turns nearest the dyad must be close to 10.7 b.p., i.e. to that of DNA in solution or even slightly underwound. Two lines of experimental evidence support this conclusion. First the low cross strand stagger of DNase I cleavage at and close to the dyad observed both for a population of chicken erythrocyte core particles (Lutter 1979) and for a nucleosome reconstituted with a defined single DNA sequence (Rhodes 1985; Drew & Travers 1985*a*) is consistent with a low twist angle (Rhodes & Klug 1986). Secondly, the periodicity of cleavage by both DNase I and DNase II within the core increases in the vicinity of the dyad (Lutter 1979, 1981; Cockell *et al.* 1983). Although the precise value of this cutting periodicity could be influenced by steric effects (for a full discussion see Klug & Lutter (1981)) the qualitative pattern of DNase I cleavage is closely paralleled by the sequence periodicities in core DNA (figure 5).

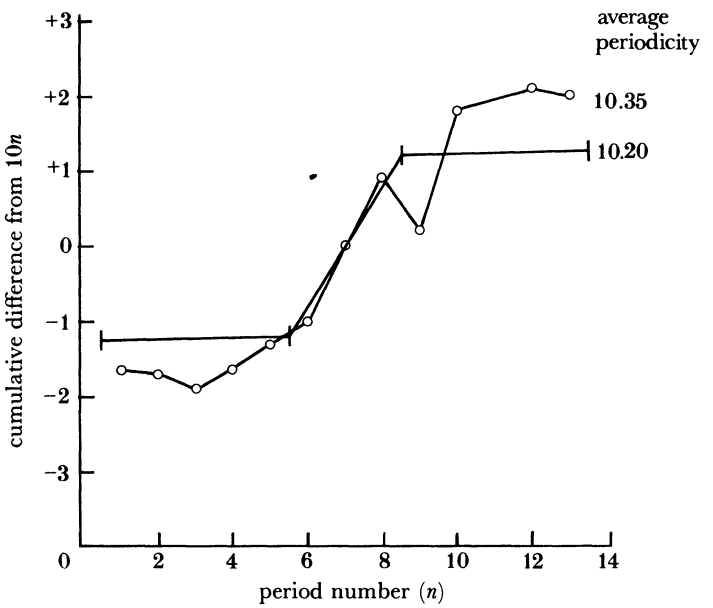

FIGURE 5. Comparison of average sequence periodicity with periodicity of cleavage by various nucleases in nucleosome core DNA. The average sequence periodicity (|——|) for the outer arms of core DNA was calculated by Fourier analysis of the trinucleotide sequences with the strongest periodic modulation (GGC/GCC, AAA/TTT, AAT/ATT) separately over the two outer windows shown. These separate values were then averaged and the periodicity in the region of the dyad was calculated from the difference of that number and the average sequence periodicity of 10.2 b.p. across the whole extent of nucleosome core DNA (Satchwell *et al.* 1986). The cleavage periodicities (○——○) are the average of the DNase I and DNase II cutting patterns of Lutter (1978, 1981). The dyad is taken as the origin and thus any period in excess of 10 b.p. is scored as positive in the rightward direction and negative in the leftward direction. The difference between the two curves is a measure of the difference in angle of attack of the nucleases from perpendicularity to the superhelical axis produced by steric hindrance between neighbouring superhelical turns (compare figure 6 in Prunell *et al.* (1979) with figure 4 in Klug & Lutter (1981)).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A local twist of 10.02 b.p. in the outer arms of nucleosome core DNA has a second important implication for nucleosome structure. As pointed out previously (Finch *et al.* 1977; Klug & Lutter 1981) an integral number of base pairs per turn of the double helix would mean that the phosphate groups of the two adjacent superhelical turns would keep in phase. A periodicity of 10.0 b.p. per turn thus allows the same stabilizing interactions to occur repeatedly along the chain just as it does between adjacent molecules in fibres of B DNA (Dover 1977). In addition, calculations using empirical energy functions suggest that smoothly bent DNA with a 45 Å radius of curvature is most stable with this same helical screw of 10.0 ± 0.1 b.p. per turn rather than 10.6 b.p. per turn (Levitt 1978). Hence the observed sequence periodicity of core DNA strongly suggests that the structure of the nucleosome core particle optimises the packing of adjacent superhelical turns of DNA, except in the neighbourhood of the dyad. On the whole nucleosome, the exit and entry point of the DNA are also in the dyad region which thus contains segments of three DNA helices whose precise relative configuration is, as yet, unknown.

We note finally that the above discussion of sequence periodicity relates to the 'average' nucleosome core particle. However, it is clear that the distance between DNase I cuts within reconstituted core particles is not constant but can vary within a narrow range from 10.2 to 10.5 b.p. (Drew & Travers 1985*a*; Rhodes 1985; Drew & Calladine 1987). One must not therefore equate the 'average' helical periodicity with that of a particular core particle.

Sequence markers for translational positioning on the nucleosome

The periodically modulated sequences in nucleosome core DNA can be directly related to the periodic rotation of the DNA double helix as it wraps around the histone octamer and thus defines the angular orientation of the sequence relative to the direction of curvature of the helical axis. Note that the same average rotational setting can be achieved by simply translating the DNA along its axis relative to the histone octamer by a multiple of the helical periodicity. What determines this translational positioning? We might expect that local variations in both the magnitude and direction of bending, as well as variations in other physical parameters such as torsion, could select for the preferential location of sequences in a non-periodic manner. The most striking example of such preferential location revealed by sequence analysis is the avoidance of the central superhelical turn of core DNA by runs of $(dA) \cdot (dT)$ greater than 8 b.p. in extent. Within core DNA the occurrence of certain trinucleotides is also strongly favoured or disfavoured at particular positions. Such markers include a preference for TpTpG/CpApA at the dyad itself (S. C. Satchwell, H. R. Drew and A. A. Travers, unpublished observations), and for YpGpG/CpCpR in the regions immediately flanking the outer bounds of the nucleosome (Drew & Calladine 1987).

Three experimental studies have analysed translational positioning by introducing local alterations in longer sequences that position precisely on a single histone octamer (Fitzgerald & Simpson 1985; Ramsay 1986; Neubauer *et al.* 1986). These alterations involve either successive deletions of the DNA sequence from the outer limits of the nucleosome core or, alternatively, internal insertions in the vicinity of the dyad. The latter alter the frames of part of the DNA sequence relative to its previous rotational orientation and to its contact points with specific histones. The general conclusion from all these studies is that the particular translational setting must arise through interaction with a relatively large segment of DNA.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A secondary conclusion, consistent with the sequence analysis and deduced largely from deletion studies, is that those sequences which lie within two to three double helical turns of the dyad exert a stronger influence on translational positioning than sequences distant from the midpoint of the DNA.

It also remains possible that sequences in the internucleosomal DNA or on the edge of the nucleosome core are equally important or more important than the core DNA in determining translational positioning. Indeed, studies of the chromatin structure in the vicinity of the yeast *trp1* gene show that the neighbouring origin of DNA replication is normally nucleosome-free *in vivo* (Thoma 1986). This region contains several short stretches of (dA)·(dT) and is intrinsically bent (Snyder *et al.* 1986). This raises the possibility that some structures of this type may be incompatible with the path of the DNA around the histone octamer and thus would exclude nucleosomes. Such sequences may not be the only type that have this property. Further sequence analysis is required to clarify this issue.

With the available data, Drew & Calladine (1987) have constructed an algorithm based solely on the correlation between particular sequences and their location in nucleosome core DNA. This algorithm assumes that the rotational sequence preference for successive 10.2 b.p. periods is constant, except in the region of the dyad where a change in phase is necessary to accommodate changes in the direction of curvature. In addition, the translational position is determined by the relative frequency of occurrence of the trinucleotides ApApA/TpTpT and YpGpG/CpCpR over the same region. Even with this simplified picture of sequence distribution in core DNA the algorithm accounts well for the position of most reconstituted single nucleosomes, as well as for that of three out of five histone octamers reconstituted on a piece of DNA 860 b.p. long (Drew & Calladine 1987).

Relation of position-dependent sequence preferences to DNA structure

To what extent are the rotational sequence preferences consistent with the known sequence-dependent polymorphism of DNA? The striking result of the sequencing of nucleosome DNA described above is that all such periodically repeated sequence elements fall into two classes whose settings on the nucleosome core differ in phase by approximately 180°. This means that the preferred base pair steps of these two classes occupy positions separated by half a double-helical turn. By reference to the crystal structure analysis (Richmond *et al.* 1984) and DNase I probing (Lutter 1978) we can deduce that these two classes correspond to positions where the minor groove points either in or out in relation to the histone octamer. In other words, the base pairs lie in an orientation that allows the major and minor grooves of the double helix to open and close smoothly as the DNA winds round the protein. This bending is consistent with the 'roll' deformation in which the short axes of adjacent base pair planes are inclined relative to each other (Dickerson & Drew 1981; Fratini *et al.* 1982; Calladine 1982; Dickerson 1983; Calladine & Drew 1984, 1986).

At present there are insufficient data to relate the observed sequence periodicities to the several crystal structures of DNA that have been solved in the past few years. However, certain correlations are evident. For example, the sequence GpGpC in the crystal structure of d(GGGGCCCC) has a large total roll of +20° that opens the minor groove (McCall *et al.* (1985); see also Wang *et al.* (1982)). The sequence ApApTpT in the crystal structure of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

d(CGCGAATTCGCG) has a slightly negative roll and a large 'propeller twist' that closes the minor groove (Dickerson & Drew 1981). For the sequences ApApA and RpTpG no detailed structural information is yet available.

Despite the paucity of X-ray data, Calladine & Drew (1986) have recently developed a second algorithm that accounts for the preferred positions of the histone octamer in terms of a set of preferred values of the roll angle at each of the ten types of dinucleotide step. The algorithm measures the closeness of fit between an ideal distribution of roll angle values required to establish the given configuration of the DNA and the permissible values corresponding to a particular base sequence. The set of permissible values used is based partly on crystal structure data, and partly chosen to fit the statistical preferences found in the observed data from nucleosome core DNA sequences. To obtain a good fit, most dinucleotide steps are allowed two possible values of roll angle in accordance with the bistability deduced from crystal data (Calladine & Drew 1984), and in order to introduce a degree of context dependence the values for certain dinucleotide steps are chosen according to the preceding base pair.

The algorithm accounts successfully for the rotational setting observed for several DNA sequences that have been reconstituted with core histones and studied in solution (Simpson & Stafford 1983; Ramsay *et al.* 1984; Drew & Travers 1985*a*; Rhodes 1985; Drew & Calladine 1987). To account for the translational setting as well, Calladine and Drew (1986) found it necessary, as in the first algorithm based directly on statistical occurrences, to introduce a change of phase in the vicinity of the dyad, such as had been observed by Satchwell *et al.* (1986) for certain dinucleotide periodicities (e.g. ApA, figure 3), corresponding apparently to a jog in the path of the DNA, as described previously.

This last point emphasizes the fact that sequence-dependent anisotropic flexibility is not the sole determinant of translational positioning on the nucleosome. Indeed, there are examples of strong sequence selection at particular positions that do not correlate with those for rotational orientation. One example already referred to is the nucleosome dyad, the point at which the two copies of histone H3 come into contact. Although the minor groove points outwards at this position, the observed local sequence preferences show no correlation with the rotational sequence preferences expected for such an orientation on the nucleosome core (Satchwell *et al.* 1986). Another example is the preferred occurrence of both TpGpG and CpGpG in locations other than in the region of the nucleosome dyad (Drew & Calladine 1987). In this case, it has been suggested that the torsional requirement for a high average twist-angle in core DNA would result in the exclusion of these sequences from the central region, because in the crystal structures of TpGpG, and also of CpGpG, the YpG steps are notable for an unusually low twist-angle of 16° (Wang *et al.* 1982; McCall *et al.* 1986). Another possible determinant of nucleosome positioning would be specific chemical interactions between the DNA and the histone proteins of positions of close contact. At present there is insufficient information to evaluate the contribution, if any, of such interactions.

The T–A base step

Not only can we correlate known DNA structures with certain observed sequence preferences in nucleosome core DNA but also we can ask what these preferences tell us about the flexibility or rigidity of certain DNA sequences. For instance, simple alternating sequences that can wrap around the histone octamer must possess axial flexibility. There are two

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

examples of such sequences: poly(dAT)·(dAT), which can be reconstituted into nucleosomes (Rhodes 1979), and sequences from the collection of nucleosome core DNA clones (Satchwell *et al.* 1986) that consist almost entirely of mixed purines on one strand and mixed pyrimidines on the opposite strand. These latter sequences thus contrast with poly(dA)·(dT) and poly(dG)·(dC), which cannot be reconstituted into nucleosomes (Rhodes (1979); Simpson & Kunzler (1979); see also Kunkel & Martinson (1981); Prunell (1982)) and are discussed in more detail below.

What are the distinguishing features of these axially flexible DNA sequences? Poly-(dAT)·(dAT) contains two base steps that differ substantially in thermal stability, the dinucleotide step TpA melting at a temperature 20 °C below the step ApT (Gotoh & Tagashira 1981). This difference in stability is presumably a consequence of a greater stacking overlap for the ApT step compared with the TpA step (Klug *et al.* 1979; Calladine & Drew 1984). Several lines of evidence confirm the relative instability of the TpA step. First, in the crystal structure of pATAT the helix is unstacked at its central TpA step giving two separate ApT dinucleotide units (Viswamitra *et al.* 1978). Secondly, Patel *et al.* (1982), on determining the rate of exchange of the thymine imino proton by nuclear magnetic resonance spectroscopy as a measure of transient base pair opening, found that such exchange was 3-fold faster for the sequence GTATAC than GAATTC and even faster for the sequence TATAAT. Thirdly, there are no kinetic barriers to the formation of cruciform structures by short poly(dAT)·(dAT) blocks in a negatively supercoiled plasmid, in contrast to other palindromic sequences that undergo this structural transition (Greaves *et al.* 1985; Haniford & Pulleyblank 1985). Finally, in both relaxed and supercoiled DNA the TpA step is preferentially sensitive to cleavage by micrococcal nuclease and S1 nuclease (Dingwall *et al.* 1980; Hörz & Altenburger 1981; Drew *et al.* 1985; Flick *et al.* 1986). Because both of these nucleases require an exposed single strand as a substrate (Drew 1984), these results suggest that some, but not necessarily all, sequences of the type NTAN are often, at least transiently, unwound (Drew *et al.* 1985). Taken together, these results show that the TpA dinucleotide step is less stable than other dinucleotide steps. This characteristic accounts for the ubiquitous utilization of the 'TATA' sequence in processes such as transcription, site-specific recombination and the initiation of DNA replication, all of which involve DNA strand separation (Drew *et al.* 1985).

There is also reason to believe that the second example of simple sequence found in nucleosome core particles is easily unwound. The simplest asymmetric poly(dR)·(dY) sequence of this type is poly(dAG)·(dCT). Again, blocks of this alternating copolymer are preferentially cleaved by S1 nuclease both in linear and in negatively supercoiled DNA (Hentschel 1982; Mace *et al.* 1983; Htun *et al.* 1984; Pulleyblank *et al.* 1985; Evans & Efstratiadis 1986). It may therefore be more than a coincidence that both of these simple repeating sequences can wrap around the histone octamer.

Intrinsically bent DNA

We have so far been discussing the bendability of DNA with respect to its reaction to applied external forces imposed by interaction with proteins or by closure into a small circle. But there are now clearly established cases of permanently bent-form DNA without external constraints (Marini *et al.* 1982; Hagerman 1984). This phenomenon was first recognized as a property of small circular DNA molecules found in the kinetoplast of certain flagellate Protozoa

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

(Simpson 1979; Marini *et al.* 1982). It seems likely that the organized packing of these mini-circles into a catenated network is due to such an inherent structural feature of the DNA (Silver *et al.* 1986). The topic of intrinsic bending, which is now under intensive study by many groups, is not the main subject of this paper, but some of the results on the bendability of DNA merge into the subject of permanently bent DNA. Indeed, we can relate some of the features found in the disposition of sequences on the nucleosome directly to the structure of bent DNA.

It has been demonstrated, both for the kinetoplast DNA and for artificially designed sequences, that any short runs of A or T nucleotides longer than 3 b.p. (such as AAAA or TTTTT), when periodically repeated at intervals of 10–11 b.p., can confer a detectable amount of curvature on an isolated DNA molecule (Marini *et al.* 1982; Wu & Crothers 1984; Hagerman 1985, 1986; Ulanovsky *et al.* 1986). This intrinsic bending is detectable by the anomalous migration in polyacrylamide gel electrophoresis of DNA fragment containing such sequences (Simpson 1979). From such measurements of mobility it remains unclear by what amount the DNA bends and in what direction. However, direct measurements of birefringence decay show that a DNA molecule which is anomalously retarded on gel electrophoresis behaves in solution as though the average distance between the ends of the molecule is less than would be expected for a linear DNA molecule of average persistence length (Hagerman 1984; Levene *et al.* 1986). These observations strongly suggest that DNA molecules with anomalous gel mobility possess a time-averaged net curvature. This conclusion is supported by two further observations: first, the proposed curvature can be directly visualized by electron microscopy (Griffith *et al.* 1986; Kitchin *et al.* 1986) and secondly, an otherwise normal DNA molecule may be induced to migrate abnormally by a rigid structure in the form of a DNA cruciform or a bound protein located near its centre (Gough & Lilley 1985; Wu & Crothers 1984; Shuey & Parker 1986).

How do these results relate to the bending of DNA on the nucleosome? We know that on the nucleosome core there is a strong tendency for the centres of short runs of (dA)·(dT) such as AAA or AAAA to lie with their minor grooves along the inside of the DNA supercoil whereas runs of somewhat longer length such as AAAAA and AAAAAA tend to be on the upper and lower surfaces of the supercoil (figure 6*a*). In other words, the junctions of these longer runs of (dA)·(dT) with their flanking sequences coincide (on average) on one side with a maximum in the distribution of the centres of the trinucleotide ApApA/TpTpT, and on the other side with GpGpC/GpGpC (figure 4*b*). From the arguments advanced above this means that these junctions are sited where the minor groove either closes or opens out whereas the centre of a homopolymer run lies in a region of minimum curvature. This is consistent with a structure in which the homopolymer run itself has little or no curvature whereas at the junctions there is a significant 'roll' angle responsible for a change in the direction of the helix axis. A structural discontinuity at certain such junctions is indeed apparent from digestion studies with DNase I (Drew & Travers 1985*a*).

If we assume that the structure of longer (dA)·(dT) runs is the same in solution as it is on the nucleosome core particle we can attempt to predict the rotational orientation configuration of intrinsically bent DNA. The simplest sequence that forms such a structure is $(T_5A_5)_n$ (Hagerman 1985). This sequence has two junctions: an ApT step that can adopt a negative roll angle (Fratini *et al.* 1982) and a TpA step that in two crystals has a positive roll angle (Wang *et al.* 1982; Shakked *et al.* 1983). Because runs of $(dA)_n$, where $n \geqslant 4$, are sufficient for the formation of intrinsic bends (Koo *et al.* 1986), the sequence $(T_5A_5)_n$ should adopt a curved

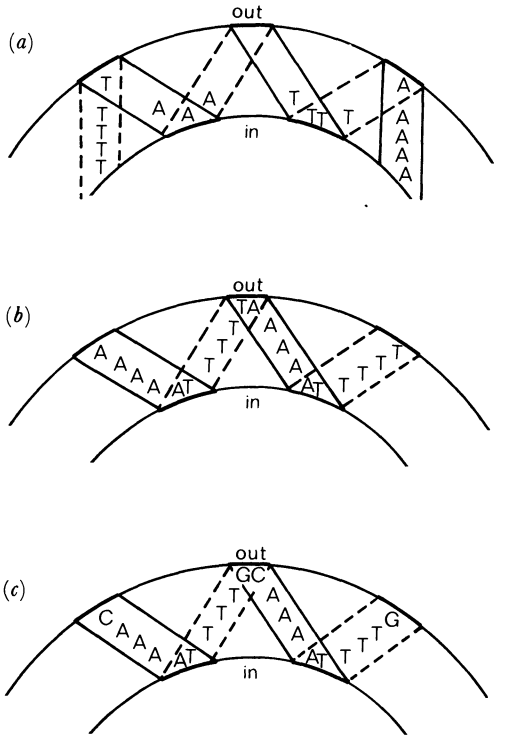

FIGURE 6. (*a*) Preferred orientations of A_3, T_4, T_5, A_5 sequences in nucleosome core DNA. Schematic details are the same as those in figure 4*b*. (*b*) Conjectured rotational orientation of the intrinsically bent DNA sequence $(A_5T_5)_n$ (Hagerman 1985). (*c*) Conjectured rotational orientation of the intrinsically bent DNA sequence $(CA_4T_4G)_n$ (Hagerman 1986).

configuration with the ApT and TpA steps positioned where the minor groove is respectively on the inside and outside of the curve (figure 6*b*). Similarly for the sequences $(GA_4T_4C)_n$ and $(CA_4T_4G)_n$, the ApT step would occupy an inward-facing minor groove whereas the CpG and GpC step would occupy a corresponding outward-facing position as they do on the nucleosome. These sequences indeed migrate unusually slowly on polyacrylamide gel electrophoresis. However, the related sequences $(GT_4A_4C)_n$ and $(CT_4A_4G)_n$, in which a TpA step replaces the ApT step, migrate normally (Hagerman 1986). This alteration could either be due to the instability of the TpA step or to a preference for this step, with a positive roll angle, to be positioned where the minor groove is on the outside of the curve. This latter possibility would cancel the outward facing preference of the CpG or GpC steps found on the nucleosome and result in a 'jogged' structure.

For the sequences discussed above, an alternation of positive and negative roll angles spaced on average at half turn intervals could account for a curvature of the helical axis. However, such an alternation is not sufficient by itself (Diekmann 1986). The role of the oligo (dA)·(dT) stretch might therefore be, by virtue of its conformational rigidity, to stabilize the base stacking interaction with flanking base pairs.

There are other indications that there is something special about oligo(dA)·(dT) stretches. It has long been apparent that poly(dA)·(dT) is structurally distinct from random sequence DNA, adopting a helical periodicity of 10 b.p. per turn in contrast to the average 10.5–10.6 b.p. per turn (Rhodes & Klug 1981; Peck & Wang 1981). Some of its other anomalous

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

properties are discussed by Rhodes & Klug (1981). The reluctance of this homopolymer to reconstitute into nucleosome cores may arise because of a lack of conformational flexibility, a property that may be intrinsic to its unusual helical structure. In mixed-sequence DNA, (dA)·(dT) runs of 5 b.p., but not 3 b.p., in extent are relatively resistant to digestion by DNase I (Drew & Travers 1984). Thus a critical length of 4–5 b.p. is required for runs of (dA)·(dT) to assume a DNase I resistant structure which we assume to be closely related or identical to that of poly (dA)·(dT). This is the same minimum length that is necessary for the formation of intrinsic bends (Koo *et al.* 1986; Diekmann 1986).

Models for DNA bending

A number of largely theoretical, specific structural models have been proposed to account for the bendability of DNA, particularly in the context of the nucleosome core particle, and also for the properties of intrinsically bent DNA. Trifonov & Sussman (1980) were the first to put forward the idea that there would be a relation between a DNA sequence and its ability to be bent in a preferred direction. However, their specific proposals on this relation have not been sustained. In addition, many detailed predictions made by subsequent models are also not consistent with preferred positioning of particular sequences in nucleosome DNA.

The specific model proposed by Trifonov and his colleagues is the 'wedge' structure for a DNA base pair step (Trifonov 1980, 1985; Mengeritsky & Trifonov 1983). They found by an iterative analysis of various eukaryotic DNA sequences that purine–purine steps in general, and ApA in particular, exhibit weak modulations with a 10.5 b.p. periodicity. Similarly, the occurrence of pyrimidine–pyrimidine steps appeared to be modulated with the same periodicity but in the opposite phase. They equated these weak periodic occurrences with the bending of DNA around the nucleosome core particle and suggested that the rotational orientation of these sequences was a consequence of a hypothetical 'wedge' structure assumed by these dinucleotides (Trifonov & Sussman 1980). Neither the proposed periodicity nor the 'wedge' structure is in accord with the experimental data. If, on the nucleosome core particle, ApA and TpT were to occupy positions *ca.* 5 b.p. apart then the occurrences of ApA and TpT combined should show a periodic fluctuation of 5 b.p. Such a periodic modulation is notably absent (see figure 2 of Satchwell *et al.* (1986)). Instead, the occurrences of both ApA and TpT separately show an average modulation of 10.2 b.p. in identical phases. The wedge model thus does not account for the observed sequence modulations in nucleosome core DNA and hence such a proposed structure cannot be a major determinant of DNA bendability.

An alternative model was proposed by Zhurkin (1983, 1985) on the basis of the limited set of crystallographic data then available for DNA structures. He suggested that, when DNA bends, pyrimidine–purine steps prefer to adopt a position with the minor groove facing out away from the direction of curvature whereas purine–pyrimidine steps adopt the opposite orientation. Although such a relation appears to hold for particular dinucleotide steps such as CpG, this suggestion is not generally true. Thus both in crystals of G_4C_4 and in nucleosome core DNA the purine–pyrimidine step GpC opens to the minor groove (McCall *et al.* 1985; Satchwell *et al.* 1986).

Specific models for the structure of intrinsically bent DNA associated with runs of (dA)·(dT) fall into two classes: first, those that propose that the run of (dA)·(dT) is essentially straight and therefore that the overall curvature results from a particular angled structure at the junctions

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

between the homopolymer sequence and its flanking base pairs (Hagerman 1984; Wu & Crothers 1984; Widom 1985; Diekmann 1986; Koo *et al.* 1986), and secondly, an extension of the 'wedge' model which proposes that the homopolymer runs are themselves curved (Trifonov 1985). This latter model is inconsistent with recent experimental results (Diekmann 1986; Hagerman 1986; Koo *et al.* 1986). The 'junction' models again can be separated into two classes. The first (Diekmann 1986; Koo *et al.* 1986) ascribes the curvature to a change in the absolute angle (ϕ_r) of base-pair tilt at the junction while maintaining parallel stacking of base pairs, as originally envisaged by Selsing *et al.* (1979). The second (Hagerman 1984; Widom 1985) proposes that the change in direction of the helical axis is a consequence of base-roll at the junctions (Fratini *et al.* 1982; Dickerson 1983; Calladine & Drew 1986). If similar internal 'forces' are at work in the induced bending of DNA, as in determining permanent bends, these two types of model can be compared with the preferred sequence arrangement of nucleosome core DNA. In particular, the two models have different predictions for the preferred placement of runs of (dA)·(dT) five to six base pairs in extent on the nucleosome core particle. The 'tilt' model, if applied to the nucleosome, would predict that the junctions of such runs with neighbouring base pairs should be sited where the minor groove points 'up' or 'down' relative to the histone octamer. The 'roll' model predicts that such junctions should be where the minor groove points 'in' or 'out'. The actual placement on the nucleosome of many of the (dA)·(dT) runs of this length (Satchwell *et al.* 1986) is consistent only with the 'roll' model. Nevertheless we note that this model does not easily explain why the intrinsically bent sequences $(A_6CG_2C)_n$ and $(CG_3T_6G_3CA_6)_n$ have approximately the same anomalous mobility (Koo *et al.* 1986).

The biological implications of DNA bendability

We have argued that DNA bendability is a major determinant of the specific positioning in the interaction between a long piece of DNA and the histone octamer. The precise placement of nucleosomes in chromatin can thus be directly related to a particular structural property of DNA. Positioning of this nature has important implications for the regulation of eukaryotic genes transcribed by both RNA polymerase II and RNA polymerase III. Several cases have now been described where nucleosomes are positioned to overlap or occlude transcriptional control regions (Rhodes 1985; Almer & Hörz 1986; Cordingley *et al.* 1987). Specific activation of such genes can require removal of these positioned nucleosomes. One such example is the yeast *PHO5* gene where activation results in the removal of four nucleosomes upstream of the transcription startpoint with the consequent unmasking of additional regulatory sequences (Almer *et al.* 1986). A similar situation occurs in the long terminal repeat of mouse mammary-tumour virus where hormone-mediated promoter activation requires the interaction of the glucocorticoid receptor with DNA exposed on the surface of a positioned nucleosome (Cordingley *et al.* 1987). This interaction alters the chromatin structure in this region to expose a binding site for a transcriptional activator, nuclear factor I.

In other situations it may be functionally advantageous for a particular DNA sequence to remain nucleosome free. Such a role has been suggested for runs of (dA)·(dT) acting as upstream activating elements for constitutive promoters in yeast (Struhl 1985) and also for the presence of intrinsically bent DNA in a yeast DNA replication origin (Snyder *et al.* 1986).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The influence of DNA bendability on nucleosome positioning also has implications for protein–DNA interactions in general. In one other studied case, that of the *E. coli* catabolite activator protein (CAP) bound to the *lac* control region, the DNA is bent by 90°–180° (Liu-Johnson *et al.* 1986). From the dimensions of the CAP dimer (McKay & Steitz 1981) we can deduce that the magnitude of DNA curvature in this complex must be similar to that in the nucleosome core particle. We note that there is a striking identity between the strongest rotational sequence preferences on the nucleosome core and sequences in the corresponding orientations in the CAP–DNA complex (figure 7). In other large DNA–protein complexes,

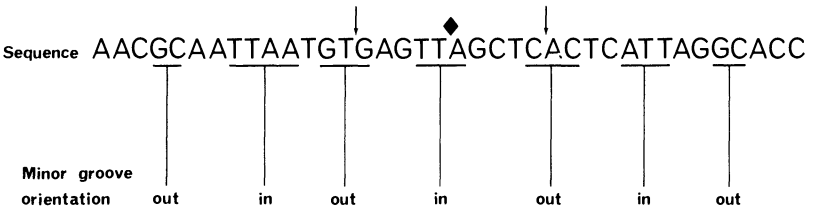

FIGURE 7. Sequence of the CAP binding site in the *E. coli lac* promoter region showing the relation of the DNA sequence to the rotational orientation of the minor groove. A structural periodicity of 10 b.p. is assumed between the DNase I accessible regions (arrows) in accord with the observed cutting periodicity (data from Spassky *et al.* (1984)).

although the precise path of the DNA is not known, electron microscopy of three such complexes, involving the *E. coli* DNA gyrase and DnaA protein and phage λ integrase (Better *et al.* 1982; Fuller *et al.* 1984; Kirchausen *et al.* 1985), shows that, just as in the nucleosome, a long stretch of DNA is tightly wrapped around an aggregate of protein molecules. One consequence of this arrangement is to bring together parts of the DNA molecule that are far apart in the linear sequence and that are involved in a 'DNA transaction' (Echols 1986). By inspection of the patterns of DNase I digestion of such complexes it is possible to deduce the general direction of curvature of the DNA with respect to the protein. In all cases so far examined† it appears the sequence-dependent preferences that are characteristic of the nucleosome core are conserved in the interactions of other proteins with DNA (A. A. Travers, unpublished observations). The sequence conservation implies that the DNA curvature in protein–DNA complexes is positive; that is, the curvature is directed towards the protein. Were this not the case no correlation would have been observed between minor grooves directed outwards from the protein (and therefore accessible to DNase I) and sequences in corresponding positions on the core associated with a known direction of curvature.

These examples imply that DNA bending is of general occurrence and complements sequence-specific recognition, although clearly the relative contribution of bending to complex formation must depend on both the magnitude of curvature and the extent of the curved DNA segment. One would imagine that the energy required to deform DNA mildly in a non-preferred direction is a lot less than that which might be available from direct sequence recognition (by H bonding, etc.) so that, in general, the influence of bendability is a second-

† The particular proteins surveyed were: *E. coli* DNA gyrase (Morrison & Cozzarelli 1981; Kirkegaard & Wang 1981), DnaA protein (Fuller *et al.* 1984), Tn3 resolvase (Grindley *et al.* 1982; Sherratt *et al.* 1984), λ integrase (Ross *et al.* 1979), λ O protein (Zahn & Blattner 1985) and *E. coli* RNA polymerase holoenzyme (Siebenlist *et al.* 1980; Spassky *et al.* 1985).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

order effect. But there can be instances where sequence-dependent bendability might be used to play a significant role as in the nucleosome and the cases just mentioned.

DNA bending must also occur when the double helix is constrained in a loop by the cooperative binding of proteins to separated sites. Such a structure has been proposed to occur naturally in the regulatory region of the *E. coli ara*BAD operon whose repression depends on two binding sites for the AraC protein separated by *ca.* 225 b.p. (Dunn *et al.* 1984) and also for the *lac*, *gal* and *deo* operons (Mossing & Record 1986; Irani *et al.* 1983; Valentin-Hansen *et al.* 1986). In another striking, but perhaps artificial example, the λcI repressor can bind cooperatively to two operator sites separated by five or six double helical turns but not by five and a half turns (Hochshild & Ptashne 1986). However, in such cooperative binding the sequence preferences affecting the rotational orientation of DNA would only be expected to be of significance where the magnitude of curvature is substantial, and the binding energy stabilizing protein–protein contacts is insufficient to override adverse bending preferences. However, the formation of looped complexes does require the interacting proteins to be in the correct angular orientation relative to each other. This means that the distance between protein binding sites can only vary by approximately integral multiples of double helical turns, because the torsional flexibility of DNA over distances of less than 500 b.p. is insufficient to stabilize binding between sites that are not in angular register (Shore & Baldwin 1983). We note that in this situation a stretch of relatively rigid DNA between the binding sites would have the potential to either facilitate or impede cooperative binding by influencing the spatial proximity of the binding sites.

In summary, it would appear that the sequence-dependent preferences for bending a DNA double helix are of wide occurrence and utility. These sequence dependent features thus constitute a further kind of information – a structural code – present in DNA.

Endnote

Two recent results, obtained since this paper was submitted, bear on the question of the origin of the curvature in intrinsically bent DNA. In the crystal of $CGCA_6GCG$, the structure of which was recently determined in this laboratory (Nelson *et al.* 1987), the (dA)·(dT) tract is essentially straight. This feature shows that intrinsic curvature is not a consequence of bending within the homopolymer run and thus must arise from a change in direction of the helical axis at or near the junctions of the (dA)·(dT) tract. This change in direction could in fact be more widely distributed through the intervening sequences. Complementary to this conclusion is the observation that in small circles of *Leishmania* kinetoplast DNA runs of (dA)·(dT) 5–6 base pairs long are placed so that their centres are approximately midway between an outward facing and an inward facing minor groove (A. A. Travers, unpublished data). This rotational orientation, which is the same as that adopted on the nucleosome, is (as discussed earlier) consistent with models that propose that curvature is principally a consequence of a change in roll angle at or near the junctions but is inconsistent with those models that invoke a change in tilt as the major determinant of curvature.

We thank Dr Michael Levitt for stimulating discussions and helpful comments on the manuscript.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

References

Almer, A. & Hörz, W. 1986 Nuclease hypersensitive regions with adjacent positioned nucleosomes mark the gene boundaries of the PHO5/PHO3 locus in yeast. *EMBO J.* **5**, 2681–2687.

Almer, A., Rudolph, H., Hinnen, A. & Hörz, W. 1986 Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. *EMBO J.* **5**, 2689–2696.

Better, M., Lu, C., Williams, R. C. & Echols, H. 1982 Site-specific DNA condensation and pairing mediated by the int protein of bacteriophage λ. *Proc natn. Acad. Sci. U.S.A.* **79**, 5837–5841.

Calladine, C. R. 1982 Mechanics of the sequence-dependent stacking of bases in B-DNA. *J. molec. Biol.* **161**, 343–352.

Calladine, C. R. & Drew, H. R. 1984 A base-centred explanation of the B-to-A transition in DNA. *J. molec. Biol.* **178**, 773–782.

Calladine, C. R. & Drew, H. R. 1986 The principles of sequence-dependent flexure of DNA. *J. molec. Biol.* **192**, 907–918.

Cockell, M., Rhodes, D. & Klug, A. 1983 Location of the primary sites of micrococcal nuclease cleavage on the nucleosome core. *J. molec. Biol.* **170**, 423–446.

Cordingley, M. G., Riegel, A. T. & Hager, G. L. 1987 Steroid-dependent interaction of transcription factors with the inducible promoter of mouse mammary tumor virus *in vivo*. *Cell* **48**, 261–270.

Crick, F. H. C. 1953 The Fourier transform of a coiled-coil. *Acta crystallogr.* **A6**, 685–689.

Crick, F. H. C. 1976 Linking numbers and nucleosomes. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2639–2643.

Dickerson, R. E. 1983 Base sequence and helix structure variation in B- and A-DNA. *J. molec. Biol.* **166**, 419–441.

Dickerson, R. E. & Drew, H. R. 1981 Structure of a B-DNA dodecamer: influence of base sequence on helix structure. *J. molec. Biol.* **149**, 761–786.

Diekmann, S. 1986 Sequence specificity of curved DNA. *FEBS Lett.* **195**, 53–56.

Dingwall, C., Lomonossoff, G. P. & Laskey, R. A. 1981 High sequence specificity of micrococcal nuclease. *Nucl. Acids Res.* **9**, 2659–2673.

Dover, S. D. 1977 Symmetry and packing in B-DNA. *J. molec. Biol.* **110**, 699–700.

Drew, H. R. 1984 Structural specificities of five commonly used DNA nucleases. *J. molec. Biol.* **176**, 535–557.

Drew, H. R. & Calladine C. R. 1987 Sequence specific positioning of core histones on an 860 bp DNA: experiment and theory. *J. molec. Biol.* **195**, 143–173.

Drew, H. R. & Travers, A. A. 1984 DNA structural variations in the *E. coli tyrT* promoter. *Cell* **37**, 491–502.

Drew, H. R. & Travers, A. A. 1985*a* DNA bending and its relation to nucleosome positioning. *J. molec. Biol.* **186**, 773–790.

Drew, H. R. & Travers, A. A. 1985*b* Structural junctions in DNA: the influence of flanking sequence on nuclease digestion specificities. *Nucl. Acids Res.* **13**, 4445–4467.

Drew, H. R., Weeks, J. R. & Travers, A. A. 1985 Negative supercoiling induces spontaneous unwinding of a bacterial promoter. *EMBO J.* **4**, 1025–1032.

Dunn, T. M., Hahn, S., Ogden, S. & Schleif, R. F. 1984 An operator at −280 base pairs that is required for repression of *ara*BAD operon promoter: addition of DNA helical turn between the operator and promoter cyclically hinders repression. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5017–5020.

Echols, H. 1986 Multiple DNA-protein interactions governing high precision DNA transactions. *Science, Wash.* **233**, 1050–1056.

Edwards, C. A. & Firtel, R. A. 1984 Site-specific phasing in the chromatin of the rDNA in *Dictyostelium discoideum*. *J. molec. Biol.* **180**, 73–90.

Evans, T. & Efstratiadis, A. 1986 Sequence dependent S1 nuclease hypersensitivity of a heteronomous DNA duplex. *J. biol. Chem.* **261**, 14771–14780.

Finch, J. T., Lutter, L. C., Rhodes, D., Brown, R. S., Rushton, B., Levitt, M. & Klug, A. 1977 Structure of nucleosome core particles of chromatin. *Nature, Lond.* **269**, 29–36.

Fitzgerald, P. C. & Simpson, R. T. 1985 Effects of sequence alterations in a DNA segment containing the 5S RNA gene from *Lytechinus variegatus* on positioning of a nucleosome core particle *in vitro*. *J. biol. Chem.* **260**, 15318–15324.

Flick, J. T., Eissenberg, J. C. & Elgin, S. C. R. 1986 Micrococcal nuclease as a DNA structural probe: its recognition sequences, their genomic distribution and correlation with DNA structure determinants. *J. molec. Biol.* **190**, 619–633.

Fox, K. R. & Howarth, N. R. 1985 Investigations into the sequence selective binding of mithramycin and related ligands to DNA. *Nucl. Acids Res.* **13**, 8695–8714.

Fox, K. R. & Waring, M. J. 1984 DNA structural variations produced by actinomycin and distamycin as revealed by DNAase I footprinting. *Nucl. Acids Res.* **12**, 9271–9285.

Fratini, A. V., Kopka, M. L., Drew, H. R. & Dickerson, R. E. 1982 Reversible bending and helix geometry in a B-DNA dodecamer CGCGAATTBrCGCG. *J. biol. Chem.* **257**, 14686–14707.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Fuller, F. B. 1978 Decomposition of the linking number of a closed ribbon: a problem from molecular biology. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3557–3561.

Fuller, R. S., Funnell, B. E. & Kornberg, A. 1984 The dnaA protein complex with the *E. coli* chromosomal origin (*oriC*) and other DNA sites. *Cell* **38**, 889–900.

Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. 1975 Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843–1847.

Gotoh, O. & Tagashira, Y. 1981 Stabilities of nearest-neighbour doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles. *Biopolymers* **20**, 1033–1042.

Gough, G. W. & Lilley, D. M. J. 1985 DNA bending induced by cruciform formation. *Nature, Lond.* **313**, 154–156.

Greaves, D. R., Patient, R. K. & Lilley, D. M. J. 1985 Facile cruciform formation by an $(A\text{-}T)_{34}$ sequence from a *Xenopus* globin gene. *J. molec. Biol.* **185**, 461–478.

Griffith, J., Bleyman, M., Rauch, C. A., Kitchin, P. A. & Englund, P. T. 1986 Visualisation of the bent helix in kinetoplast DNA by electron microscopy. *Cell* **46**, 717–724.

Grindley, N. D. F., Lassick, M. R., Wells, R. G., Wityk, R. J., Salvo, J. J. & Reed, R. R. 1982 Transposon-mediated site-specific recombination: identification of three binding sites for resolvase at the *res* sites of γ δ and Tn3. *Cell* **30**, 19–27.

Hagerman, P. J. 1984 Evidence for the existence of stable curvature of DNA in solution. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4632–4636.

Hagerman, P. J. 1985 Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. *Biochemistry, Wash.* **24**, 7033–7037.

Hagerman, P. J. 1986 Sequence-directed curvature of DNA. *Nature, Lond.* **321**, 449–450.

Haniford, D. B. & Pulleyblank, D. E. 1985 Transition of a cloned $d(AT)_n$–$d(AT)_n$ tract to a cruciform *in vivo*. *Nucl. Acids Res.* **13**, 4343–4363.

Hentschel, C. C. 1982 Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease. *Nature, Lond.* **295**, 714–716.

Hochshild, A. & Ptashne, M. 1986 Cooperative binding of λ repressors to sites separated by integral turns of the DNA helix. *Cell* **44**, 681–687.

Horowitz, D. J. & Wang, J. C. 1984 Torsional rigidity of DNA and length dependence of the free energy of supercoiling. *J. molec. Biol.* **173**, 75–91.

Hörz, W. & Altenburger, W. 1981 Sequence specific cleavage of DNA by micrococcal nuclease. *Nucl. Acids Res.* **9**, 2643–2658.

Htun, H., Lund, E. & Dahlberg, J. E. 1984 Human U1 RNA genes contain an unusual nuclease S1 cleavage site within the conserved 3′ flanking region. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7288–7292.

Irani, M. H., Orosz, L. & Adhya, S. 1983 A control element within a structural gene: the *gal* operon of *Escherichia coli*. *Cell* **32**, 783–788.

Kirchhausen, T., Wang, J. C. & Harrison, S. C. 1985 DNA gyrase and its complexes with DNA: direct observation by electron microscopy. *Cell* **41**, 933–943.

Kirkegaard, K. & Wang, J. C. 1981 Mapping the topography of DNA wrapped around gyrase by nucleolytic and chemical probing of complexes of unique DNA sequences. *Cell* **23**, 721–729.

Kitchin, P. A., Klein, V. A., Ryan, K. A., Gann, K. L., Rauch, C. A., Kang, D. S., Wells, R. D. & Englund, P. T. 1986 A highly bent fragment of *Crithidia fasciculata* kinetoplast DNA. *J. biol. Chem.* **261**, 11302–11309.

Klug, A., Jack, A., Viswamitra, M. A. Kennard, O., Shakked, Z. & Steitz, T. A. 1979 A hypothesis on a specific sequence-dependent conformation of DNA and its relation to the binding of the *lac*-repressor protein. *J. molec. Biol.* **131**, 669–680.

Klug, A. & Lutter, L. C. 1981 The helical periodicity of DNA on the nucleosome. *Nucl. Acids Res.* **9**, 4267–4283.

Koo, H. S., Wu, H.-M. & Crothers, D. M. 1986 DNA bending at adenine-thymine tracts. *Nature, Lond.* **320**, 501–506.

Kunkel, G. R. & Martinson, H. G. 1981 Nucleosomes will not form on double-stranded RNA or over poly(dA).poly(dT) tracks in recombinant DNA. *Nucl. Acids Res.* **9**, 6869–6888.

Levene, S. D., Wu, H-M. & Crothers, D. M. 1986 Bending and flexibility of kinetoplast DNA. *Biochemistry, Wash.* **25**, 3988–3995.

Levitt, M. 1978 How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations. *Proc. natn. Acad. Sci. U.S.A.* **75**, 640–644.

Liu-Johnson, H-N., Gartenberg, M. R. & Crothers, D. M. 1986 The DNA binding domain and bending angle of *E. coli* CAP protein. *Cell* **47**, 995–1005.

Lomonossoff, G. P., Butler, P. J. G. & Klug, A. 1981 Sequence-dependent variation in the conformation of DNA. *J. molec. Biol.* **149**, 745–760.

Lutter, L. C. 1978 Kinetic analysis of deoxyribonuclease cleavages in the nucleosome core: evidence for a DNA superhelix. *J. molec. Biol.* **124**, 391–420.

Lutter, L. C. 1979 Precise location of DNase I cutting sites in the nucleosome core determined by high resolution gel electrophoresis. *Nucl. Acids Res.* **6**, 41–56.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Lutter, L. C. 1981 DNase II digestion of the nucleosome core: precise locations and relative exposure of sites. *Nucl. Acids Res.* **9**, 4251–4265.

Mace, H. A. F., Pelham, H. R. B. & Travers, A. A. 1983 Association of an S1-nuclease sensitive structure with short direct repeats 5′ of *Drosophila* heat shock genes. *Nature, Lond.* **304**, 555–557.

Marini, J. C., Levene, S.D., Crothers, D. M. & Englund, P. T. 1982 Bent helical structure in kinetoplast DNA. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7664–7668. Marini, J. C. & Englund, P. T. 1983 Correction *Proc. natn. Acad. Sci. U.S.A.* **80**, 7678.

McCall, M., Brown, T. & Kennard, O. 1985 The crystal structure of d(GGGGCCCC). A model for poly(dG).poly(dC). *J. molec. Biol.* **183**, 385–396.

McCall, M., Brown, T., Hunter, W. N. & Kennard, O. 1986 The crystal structure of d(GGATGGGAG) forms an essential part of the binding site for transcription factor IIIA. *Nature, Lond.* **322**, 661–664.

McKay, D. B. & Steitz, T. A. 1981 Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left handed DNA. *Nature, Lond.* **290**, 744–749.

Mengeritsky, G. & Trifonov, E. N. 1983 Nucleotide sequence-directed mapping of the nucleosomes. *Nucl. Acids Res.* **11**, 3833–3851.

Morrison, A. & Cozzarelli, N. R. 1981 Contacts between DNA gyrase and its binding site of DNA: features of symmetry and asymmetry revealed by protection from nucleases. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1416–1420.

Mossing, M. C. & Record, M. T., Jr 1986 Upstream operators enhance repression of the *lac* promoter. *Science, Wash.* **233**, 892–899.

Neubauer, B., Linxweiler, W. & Hörz, W. 1986 DNA engineering shows that nucleosome phasing on African green monkey α-satellite is the result of multiple additive histone–DNA interactions. *J. molec. Biol.* **190**, 639–645.

Nelson, H. C. M., Finch, J. T., Luisi, B. F. & Klug, A. 1987 The structure of an oligo(da)–oligo(dt) tract and its biological implications. *Nature, Lond.* **330**. (In the press.)

Palen, T. E. & Cech, T. R. 1984 Chromatin structure at the replication origins and transcription-initiation regions of the ribosomal RNA genes of *Tetrahymena*. *Cell* **36**, 933–942.

Patel, D. J., Kozlowski, S. A., Ikuta, S., Itakura, K., Bhatt, R. & Hare, D. R. 1982 NMR studies of DNA conformation and dynamics in solution. *Cold Spring Harb. Symp. quant. Biol.* **47**, 197–206.

Peck, L. & Wang, J. C. 1981 Sequence dependence of the helical repeat of DNA in solution. *Nature, Lond.* **292**, 375–378.

Prunell, A. 1982 Nucleosome reconstitution on plasmid-inserted poly (dA)·poly(dT). *EMBO J.* **1**, 173–179.

Prunell, A. & Kornberg, R. D. 1982 Variable center to center distance of nucleosomes in chromatin. *J. molec. Biol.* **154**, 515–523.

Prunell, A., Kornberg, R. D., Lutter, L., Klug, A., Levitt, M. & Crick, F. H. C. 1979 Periodicity of deoxyribonuclease I digestion of chromatin. *Science, Wash.* **204**, 855–858.

Pulleyblank, D. E., Haniford, D. B. & Morgan, A. R. 1985 A structural basis for S1 nuclease sensitivity of double stranded DNA. *Cell* **42**, 271–280.

Ramsay, N. 1986 Deletion analysis of a DNA sequence that positions itself precisely on the nucleosome core. *J. molec. Biol.* **189**, 179–188.

Ramsay, N., Felsenfeld, G., Rushton, B. M. & McGhee, J. D. 1984 A 145-base pair DNA sequence that positions itself precisely and asymmetrically on the nucleosome core. *EMBO J.* **3**, 2605–2611.

Rhodes, D. 1979 Nucleosome cores reconstituted from poly(dA–dT) and the octamer of histones. *Nucl. Acids Res.* **6**, 1805–1816.

Rhodes, D. 1985 Structural analysis of a triple complex between the histone octamer, a *Xenopus* gene for 5S RNA and transcription factor IIIA. *EMBO J.* **4**, 3473–3482.

Rhodes, D. & Klug, A. 1980 Helical periodicity of DNA determined by enzyme digestion. *Nature, Lond.* **286**, 573–578.

Rhodes, D. & Klug, A. 1981 Sequence-dependent helical periodicity of DNA. *Nature, Lond.* **292**, 378–380.

Rhodes, D. & Klug, A. 1986 An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA. *Cell* **46**, 123–132.

Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. 1984 Structure of the nucleosome core particle at 7 Å resolution. *Nature, Lond.* **311**, 532–537.

Ross, W., Landy, A., Kukuchi & Nash, H. 1979 Interaction of int protein with specific sites on λ*att* DNA. *Cell* **18**, 297–307.

Salmon, G. & Rogers, R. A. P. 1914 *A treatise on the analytic geometry of three dimensions*. (228 pages.) London: Longmans, Green.

Satchwell, S. C., Drew, H. R. & Travers, A. A. 1986 Sequence periodicities in chicken nucleosome core DNA. *J. molec. Biol.* **191**, 659–675.

Selsing, E., Wells, R. D., Alden, C. J. & Arnott, S. 1979 Bent DNA: visualisation of a base-paired and stacked A–B conformational junction. *J. biol. Chem.* **254**, 5417–5422.

Shakked, Z., Rabinovich, D., Kennard, O., Cruse, W. B. T., Salisbury, S. A. & Viswamitra, M. A. 1983 Sequence-dependent conformation of an A-DNA double helix: the crystal structure of the octamer d(GGTATACC). *J. molec. Biol.* **166**, 183–201.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Sherratt, D., Dyson, P., Boocock, M., Brown, L., Summers, D., Stewart, G. & Chan, P. 1984 Site-specific recombination in transposition and plasmid stability. *Cold Spring Harb. Symp. quant. Biol.* **59**, 227–233.

Shore, D. & Baldwin, R. L. 1983 Energetics of DNA twisting. II. Topoisomer analysis. *J. molec. Biol.* **170**, 983–1007.

Shuey, D. J. & Parker, C. S. 1986 Bending of promoter DNA on binding of heat shock transcription factor. *Nature, Lond.* **323**, 459–461.

Siebenlist, U., Simpson, R. B. & Gilbert, W. 1980 *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell* **20**, 269–281.

Silver, L. E., Torri, A. F. & Hajduk, S. L. 1986 Organised packaging of kinetoplast DNA networks. *Cell* **47**, 537–543.

Simpson, L. 1979 Isolation of maxicircle component of kinetoplast DNA from hemoflagellate protozoa. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1585–1588.

Simpson, R. T. & Kunzler, P. 1979 Chromatin and core particles formed from the inner histones and synthetic polydeoxyribonucleotides of defined sequence. *Nucl. Acids Res.* **6**, 1387–1485.

Simpson, R. T. & Stafford, D. W. 1983 Structural features of a phased nucleosome core particle. *Proc. natn. Acad. Sci. U.S.A.* **80**, 51–55.

Simpson, R. T., Thoma, F. & Brubaker, J. M. 1985 Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: a model system for the study of higher order structure. *Cell* **42**, 799–808.

Snyder, M., Buchman, A. R. & Davis, R. W. 1986 Bent DNA at a yeast autonomously replicating sequence. *Nature, Lond.* **324**, 87–89.

Spassky, A., Busby, S. & Buc, H. 1984 On the action of the cyclic AMP-cyclic AMP receptor protein complex at the *Escherichia coli* lactose and galactose promoter regions. *EMBO J.* **3**, 43–50.

Spassky, A., Kirkegaard, K. & Buc, H. 1985 Changes in the DNA structure of the *lac*UV5 promoter during formation of an open complex with RNA polymerase. *Biochemistry* **24**, 2723–2731.

Struhl, K. 1985 Naturally occurring (dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8419–8423.

Suck, D. & Oefner, C. 1986 Structure of DNase I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature, Lond.* **321**, 620–625.

Suck, D., Oefner, C. & Kabsch, W. 1984 Three-dimensional structure of bovine pancreatic DNase I at 2.5 Å resolution. *EMBO J.* **3**, 2423–2430.

Swartz, M. N., Trautner, T. A. & Kornberg, A. 1962 Enzymatic synthesis of deoxyribonucleic acid. XI. Further studies on nearest neighbour base sequences in deoxyribonucleic acids. *J. biol. Chem.* **237**, 1961–1967.

Thoma, F. 1986 Protein-DNA interactions and nuclease-sensitive regions determine nucleosome positions on yeast plasmid chromatin. *J. molec. Biol.* **190**, 177–190.

Thoma, F., Koller, Th. & Klug, A. 1979 Involvement of histone H1 in the organisation of the nucleosome and of the salt dependent superstructure of chromatin. *J. cell. Biol.* **83**, 403–427.

Thoma, F. & Simpson, R. T. 1985 Local protein-DNA interactions may determine nucleosome positions on yeast plasmids. *Nature, Lond.* **315**, 250–252.

Trifonov, E. N. 1980 Sequence-dependent deformational anisotropy of chromatin DNA. *Nucl. Acids Res.* **8**, 4041–4053.

Trifonov, E. N. 1985 Curved DNA. *CRC Crit. Rev. Biochem.* **19**, 89–106.

Trifonov, E. N. & Sussman, J. L. 1980 The pitch of chromatin DNA is reflected in its nucleotide sequence. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3816–3820.

Ulanovsky, L., Bodner, M., Trifonov, E. N. & Choder, M. 1986 Curved DNA: design, synthesis and circularisation. *Proc. natn. Acad. Sci. U.S.A.* **83**, 862–866.

Valentin-Hansen, P., Albrectsen, B. & J. E. Løve Larsen. 1986 DNA-protein recognition: demonstration of three genetically separated operator elements that are required for repression of the *Escherichia coli deo*CABD promoters by the DeoR repressor. *EMBO J.* **5**, 2015–2021.

Van Dyke, M. W. & Dervan, P. B. 1983 Chromomycin, mithramycin and oligomycin binding sites on heterogeneous deoxyribonucleic acid. Footprinting with (Methidiumpropyl-EDTA) iron (II). *Biochemistry, Wash.* **22**, 2373-2377.

Van Dyke, M. W., Hertzberg, R. P. & Dervan, P. B. 1982 Map of distamycin, netropsin and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA.Fe (II). *Proc. natn. Acad. Sci. U.S.A.* **79**, 5470–5474.

Viswamitra, M. A., Kennard, O., Jones, P. G., Sheldrick, G. M., Salisbury, S., Falvello, L. & Shakked, Z. 1978 DNA double helical fragment at atomic resolution. *Nature, Lond.* **273**, 687–688.

Wang, A. H-J., Fujii, S., van Boom, J. H. & Rich, A. 1982 Molecular structure of the octamer d (GGCCGGCC): modified A-DNA. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3968–3972.

Wang, J. C. 1979 Helical repeat of DNA in solution. *Proc. natn. Acad. Sci. U.S.A.* **76**, 200–203.

Widom, J. 1985 Bent DNA for gene regulation and DNA packaging. *Bioessays* **2**, 11–14.

Widom, J. & Klug, A. 1985 Structure of the 300 Å chromatin filament: X-ray diffraction from oriented samples. *Cell* **43**, 207–213.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Wu, H-M. & Crothers, D. M. 1984 The locus of sequence-directed and protein-induced DNA bending. *Nature, Lond.* **308**, 509–513.

Zahn, K. & Blattner, F. R. 1985 Binding and bending of the λ replication origin by the phage O protein. *EMBO J.* **4**, 3605–3616.

Zhurkin, V. B. 1983 Specific alignment of nucleosomes on DNA correlates with periodic distribution of purine–pyrimidine and pyrimidine–purine dimers. *FEBS Lett.* **158**, 293–297.

Zhurkin, V. B. 1985 Sequence-dependent bending of DNA and phasing of nucleosomes. *J. biomol. Struct. Dynamics* **2**, 785–804.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES